Topics in Markov Chain Theory and Simulation Optimisation

Berkhout, J.

2016

document version
Publisher's PDF, also known as Version of record

4. PERTURBATION BOUNDS

This chapter is based on [2].

The topic of this chapter is perturbation analysis of Markov chains. Perturbation analysis provides bounds on the effect that a change in a Markov transition matrix has on the corresponding stationary distribution. This chapter compares and analyses bounds found in the literature for Markov chains and introduces new bounds. We provide for the first time an analysis of the relative error of these bounds. Specifically, we show that condition number bounds have typically a non-vanishing relative error as the size of the perturbation tends to zero. Our new perturbation bound will have the desirable feature that the relative error vanishes as the size of the perturbation tends to zero. We discuss a series of examples to illustrate the applicability of the various bounds. Specifically, we address the question on how the bounds developed for finite Markov chains behave as the size of the system grows. Furthermore, it is shown how perturbation bounds can be made fruitful for stability analysis where the objective is to gain insight into a parameter range for which a Markov chain remains stable.

The chapter is organized as follows. Section 4.1 gives an introduction. In Section 4.2 the perturbation bounds are presented. Examples are discussed in Section 4.3. Section 4.4 is devoted to perturbation bounds for the $M/G/1$ queue with breakdowns. Other than the small numerical examples reported in the literature, the queuing system will be analysed for the case of a large but finite state-space and for the infinite dimensional case.

4.1 Introduction

Perturbation analysis of Markov chains (PAMC) studies the effect a perturbation of a Markov transition matrix has on the stationary distribution of the chain. Consider a Markov chain with discrete state-space S, transition probability matrix P, and unique stationary distribution π_P. Furthermore, let R be an alternative Markov transition matrix on S with unique stationary distribution π_R.

PAMC addresses the following question: what is the effect of switching from P to R on the stationary distribution of the chain? More formally, PAMC theory studies bounds of the type

$$||\pi_R^\top - \pi_P^\top|| \leq \Delta(R, P), \tag{4.1}$$

where $||\cdot||$ denotes a suitable vector norm (details will be provided later in the text) and $\Delta(R, P)$ is a scalar function of P and R. The study of the effect of perturbing a Markov transition matrix on its stationary distribution dates back to Schweitzer's pioneering paper [166]. Best to our knowledge, the first paper putting this perturbation question into the framework of (4.1) is [145].

Specifically, [145] proposed bounds of the form

$$\Delta(R, P) = \kappa||R - P||, \tag{4.2}$$

for some appropriate matrix norm, where κ is the so-called *condition number*. While the condition number is typically applied to bounding the effect in terms of $R - P$, Theorem 3.2 in [150] provides a condition number for $\|R^m - P^m\|$. In the remainder of this chapter we will refer to any instance of the bound in (4.1) with $\Delta(R, P)$ as in (4.2) as *condition number bound* (CNB).

PAMC is a field of active research [16, 119, 153, 156, 167, 150, 48, 106, 18] and various CNBs have been proposed in the literature [57, 106]. An alternative type of bound derived via the strong stability method, called *strong stability bound* (SSB), bounds the weighted supremum norm of $||\pi_R^\top - \pi_P^\top||$ by an expression that is non-linear as function of $||R - P||$. For early references see [111, 110] and recent references are [140, 133, 162].

Perturbation bounds are of interest in a wide area of applications. For example, in mathematical physics [177] and climate modelling [50], in Bayesian statistics [15, 8], and in Bioinformatics [151, 159]. Perturbation bounds have also been applied in robustness analysis of social networks and of Google's PageRank algorithm [60].

A fruitful model for PAMC is that of a scaled perturbation. More specifically, let R, P be two Markov transition matrices defined on the same state-space. Then the convex combination of both transition matrices

$$P(\theta) = (1 - \theta)P + \theta R, \quad \theta \in [0, 1], \tag{4.3}$$

is a well-defined Markov transition matrix. Note that $P(0) = P$ and $P(1) = R$. In perturbation analysis of $P(\theta)$ we are interested in the effect of changing θ from 0 to some value $0 < \theta \leq 1$. By linearity of norms,

$$||P(\theta) - P|| = \theta||R - P||,$$

for $\theta \in [0, 1]$. This allows to scale the size of the perturbation via control parameter θ. Letting

$$\eta(R, P) = \frac{\Delta(R, P) - ||\pi_R^\top - \pi_P^\top||}{||\pi_R^\top - \pi_P^\top||}$$

denote the *relative error of perturbation bound* $\Delta(R, P)$, scaled perturbations, i.e., when choosing R equal to $P(\theta)$, allow for analysing the behaviour of the relative error $\eta(\theta) = \eta(P(\theta), P)$ as θ tends to zero.

The analysis of scaled perturbation is of particular interest if $P(\theta)$, for $\theta \in [0, 1]$, has a clear interpretation. We will illustrate this by a queueing model with denumerable state-space and breakdowns, where θ models the probability of a breakdown. An interesting observation is that in the parametrized model we establish conditions for stability of a mixture of a stable (no breakdowns) and an unstable (only breakdowns) Markov chain modelling a pure birth process. More specifically, we apply PAMC techniques to provide a lower bound for the domain of stability of $P(\theta)$. The contributions of this chapter are the following:

- We provide a unified approach to PAMC for finite and denumerable Markov chains. Our analysis covers CNBs and SSB.

- We introduce new bounds that do have the desirable property that the relative error of the bound tends to zero as the size of the perturbation tends to zero. These new bounds are derived by a series expansion approach.
- We will provide sufficient conditions under which the convergence of the series expansion already constitutes existence of a stationary distribution. By introducing the new concept of bias term, we are able to treat the case of Markov multi-chains (i.e., chains with several ergodic classes) and uni-chains in a unified framework.
- We will show that techniques derived in PAMC can be applied to stability analysis. A worked out example from queuing theory will illustrate the fruitfulness of PAMC methods for this type of problem.

4.2 Perturbation Analysis

Throughout this chapter, unless indicated otherwise, we will consider aperiodic Markov chains defined on an at most denumerable state-space S and consisting of one closed communicating class of states with possible transient states, also known as Markov uni-chains. We follow the convention that vectors are column vectors.

4.2.1 Preliminaries and Basic Definitions

If $P = (P_{ij})_{i,j \in S}$ is a Markov transition matrix of some Markov chain $\{X_k\}$, then $P_{ij} = \mathbb{E}[\mathbb{1}_{\{X_{k+1}=j\}}|X_k = i]$ for $i, j \in S$ and $k \in \mathbb{N}$, where $\mathbb{1}_{\{X_{k+1}=j\}}$ is one if $X_{k+1} = j$ and zero otherwise, $i, j \in S$. Sometimes $P(i, j) := P_{ij}$ is used instead for notation clarity. Further, let $f \in \mathbb{R}^S$ be a reward vector where f_i is the reward for being in state $i \in S$. With these definitions, one obtains

$$\mu^\top P f = \sum_{i,j \in S} \mu_i P_{ij} f_j = \sum_{i \in S} \mathbb{E}[f_{X_1}|X_0 = i]\mu_i \tag{4.4}$$

as the expected reward after one transition provided the Markov chain is started with initial distribution μ. For more details we refer to [115, 118].

In the following denote the ergodic projector of P by Π_P, i.e., the matrix with rows identical to $\pi_P^\top$, and we let D_P denote the deviation matrix of P, which is given by

$$D_P = \sum_{k=0}^{\infty} (P^k - \Pi_P) = (I - P + \Pi_P)^{-1} - \Pi_P, \tag{4.5}$$

provided that it exists, which we will assume throughout this chapter. The matrix $(I - P + \Pi_P)^{-1}$ is called the *fundamental matrix* (potential) of P, [115]. Letting $A^\#$ denote the group inverse or generalized potential of the matrix $A = I - P$, see [147, 144], it holds that $D_P = A^\#$ if the deviation matrix exists. Conditions for existence of the deviation matrix and its related properties have been extensively studied in the literature, see [115, 176]. For finite Markov chains, the deviation matrix is an instance of the generalized inverse of $I - P$; see [144] for an early reference. As Hunter demonstrates in [101] for finite Markov chains, the generalized inverse plays a major role in perturbation analysis.

Appropriate norms used throughout this chapter are discussed in the following. For column vector $x \in \mathbb{R}^S$, we denote by

- $||x||_\infty$ the maximum absolute value, also referred to as infinity-norm (∞-norm) or sup-norm,
- $||x||_1$ the sum of absolute values, also known as L_1-norm or 1-norm.
- $\|x\|_v$ the weighted supremum norm, or v-norm, defined as

$$\|x\|_v = \sup_{i \in S} \frac{|x_i|}{v(i)}, \tag{4.6}$$

for v such that $v(0) = 1$ and $v(i) \geq 1$ for all $i \in S$. Throughout this chapter we let

$$v(i) = \alpha^i, \quad i \in S,$$

with α some unspecified constant $\alpha \in [1, \infty)$.

Note that $v \geq 1$ implies for $x \in \mathbb{R}^S$ that $\|x\|_v \leq \|x\|_\infty$. Via the above norm definitions for column vectors we obtain the corresponding matrix norms via the operator norm (a.k.a. induced norm). For a given *column* vector norm, the operator norm of a matrix $A \in \mathbb{R}^{S \times S}$ is defined as

$$\|A\| = \sup\{\|Ax\| \;:\; x \in \mathbb{R}^S \text{ with } \|x\| = 1\}.$$

For $A \in \mathbb{R}^{S \times S}$, the operator norms corresponding to the three norm definitions above lead to the following, respectively.

- $||A||_\infty$ denotes the maximum absolute row sum.
- $||A||_1$ denotes the maximum absolute column sum.
- The corresponding operator norm for the weighted supremum norm is

$$\|A\|_v = \sup_i \frac{1}{v(i)} \sum_{j \in S} |A(i,j)| v(j).$$

Note that plugging a column vector in the operator norms leads to the original norm definitions for column vectors again. In contrast, filling in row vectors in the operator norms gives for $x \in \mathbb{R}^S$ the following, respectively.

- $||x^\top||_\infty$ denotes the sum of absolute values.
- $||x^\top||_1$ denotes the maximum absolute value.
- For the weighted supremum norm it holds that

$$\|x^\top\|_v = \sum_{k \in S} v(k)|x_k|. \tag{4.7}$$

Indeed, $||x||_\infty = ||x^\top||_1$ and vice versa. The distinction between row and column vectors is motivated by the application to Markov chains, where probability vectors are row vectors to which, e.g., norm (4.7) applies, and reward functions are column vectors to which, e.g., norm (4.6) applies. Specifically, applying the v-norm to (4.4) one readily obtains

$$|\mu^\top P f| \leq ||\mu^\top||_v \, ||P||_v \, ||f||_v,$$

and similarly

$$|\mu^\top P f| \leq ||\mu^\top||_\infty \, ||P||_\infty \, ||f||_\infty \quad \text{and} \quad |\mu^\top P f| \leq ||\mu^\top||_1 \, ||P||_1 \, ||f||_1.$$

Throughout this chapter an unspecified norm $\|\cdot\|$ can be replaced by any appropriate norms such as discussed above.

To illustrate the efficiency of perturbation bounds we will use throughout the chapter three different types of finite Markov chains introduced in the following. An example of a Markov chain on a denumerable state-space will be discussed in detail in Subsection 4.4.3.

Example 4.1. Small Network: Let $S = \{0, 1\}$ and

$$P^s = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix},$$

with $p, q \in (0, 1)$. It is easily checked that

$$\pi_{P^s} = \frac{1}{p+q}(q, p)^\top$$

is the stationary distribution of P^s. The deviation matrix is given by:

$$D_{P^s} = \frac{1}{(p+q)^2} \begin{pmatrix} p & -p \\ -q & q \end{pmatrix}.$$

Ring Network: The next example that we will discuss is that of a ring, introduced in the following. Let $S = \{0, \dots, n-1\}$ and for any $n \geq 2$,

$$P^\circ(n) = \begin{pmatrix} 1-2b & b & 0 & 0 & \dots & b \\ b & 1-2\,b & b & 0 & \dots & 0 \\ 0 & b & 1-2\,b & b & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & b & 1-2\,b & b \\ b & 0 & \dots & 0 & b & 1-2b \end{pmatrix},$$

with $b \in (0, 1/2]$. We get the stationary distribution:

$$\pi_i^\circ(n) = \frac{1}{n}, \quad \text{for } i \in S.$$

For the deviation matrix, we obtain:

$$D^\circ(n) := D_{P^\circ(n)} = \begin{pmatrix} d_0 & d_1 & d_2 & \dots & d_{n-1} \\ d_{n-1} & d_0 & d_1 & \dots & d_{n-2} \\ d_{n-2} & d_{n-1} & d_0 & \dots & d_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_1 & d_2 & d_3 & \dots & d_0 \end{pmatrix},$$

where

$$d_i = \frac{(n-1)(n+1)}{12\,b\,n} - \frac{(n-i)i}{2\,b\,n} \ \text{ for } \ i \in S.$$

Furthermore, $\sum_{i=0}^{n-1} d_i = 0$. Equivalently, $D^\circ(n)$ can be expressed as

$$D^\circ(n) = \left(\widetilde{D}_{ij}(n)\right)_{i,j\in S},$$

where

$$\widetilde{D}_{ij}(n) = d_{(j-i)(\bmod n)+1} = \frac{(n-1)(n+1)}{12\,b\,n} - \frac{\{n-(j-i)(\bmod n)\}\{(j-i)(\bmod n)\}}{2\,b\,n}.$$

Star Network: The third example considered is the Star Network with state-space $S = \{0, \dots, n-1\}$. For $n \geq 2$ let

$$P^\star(n) = \begin{pmatrix} 1-\beta & \frac{\beta}{n-1} & \frac{\beta}{n-1} & \frac{\beta}{n-1} & \dots & \frac{\beta}{n-1} \\ 1-\gamma & \gamma & 0 & 0 & \dots & 0 \\ 1-\gamma & 0 & \gamma & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 1-\gamma & 0 & \dots & 0 & \gamma & 0 \\ 1-\gamma & 0 & \dots & 0 & 0 & \gamma \end{pmatrix},$$

for $\beta \in (0,1]$ and $\gamma \in [0,1)$. Following [77], the stationary distribution is given by

$$\pi_i^\star(n) = \begin{cases} \frac{1-\gamma}{1-\gamma+\beta} & \text{for } \ i = 0, \\ \frac{\beta}{(n-1)(1-\gamma+\beta)} & \text{for } i > 0. \end{cases}$$

For the deviation matrix, we obtain:

$$D^\star(n) := D_{P^\star(n)} = \left(\begin{array}{c|c} \frac{\beta}{(1-\gamma+\beta)^2} & -\frac{\beta}{(1-\gamma+\beta)^2(n-1)}\bar{1}^\top \\ \hline -\frac{(1-\gamma)}{(1-\gamma+\beta)^2}\bar{1} & \frac{1}{(1-\gamma)}I - \frac{\beta\{(1-\gamma)+(1-\gamma+\beta)\}}{(1-\gamma)(1-\gamma+\beta)^2(n-1)}\bar{1}\bar{1}^\top \end{array} \right),$$

where $\bar{1} = [1, \dots, 1]^\top$ of size $n-1$ and I denotes the $(n-1)\times(n-1)$ identity matrix. ◁

In our analysis we will frequently work with the *taboo matrix* of a Markov transition matrix P. In [110] a very elegant and flexible way for obtaining a taboo matrix is described. For this let h be a non-negative vector and σ a probability vector on S, such that $\pi_P^\top h > 0$ and $P - h\sigma^\top$ is a matrix with non-negative values, where $h\sigma^\top$ denotes a matrix product of vectors h and $\sigma^\top$, i.e., $h\sigma^\top$ is a square matrix. Condition $\pi_P^\top h > 0$ ensures that mass is removed from a recurrent part of the Markov chain described by P (note that this condition is satisfied for finite aperiodic Markov uni-chains without transient states and when at least one element of h is > 0). Then, the taboo matrix of P with respect to h and σ is defined as

$$T := P - h\sigma^\top. \tag{4.8}$$

For example, let

$$h = (P(0,0), P(1,0), P(2,0), ...)^\top$$

denote the first column of P, and let $\sigma = (1,0,0,\dots)^\top$, then

$$T = P - h\sigma^\top = \begin{cases} P(i,j) & j > 0 \\ 0 & \text{otherwise} \end{cases} .$$

In words, T is a degenerate transition matrix that avoids entering state zero which is obtained by setting the first column of P to zero. Alternatively, letting $h = (1,0,0,\dots)^\top$ and

$$\sigma = (P(0,0), P(0,1), P(0,2), ...)^\top,$$

then $T = P - h\sigma^\top$ is a degenerate transition matrix that never leaves state zero, which is obtained by setting the first row of P to zero. The taboo kernel is also known as the *residual matrix* in the literature, see [153].

In the following we write ${}_iP$ for the degenerate transition matrix that avoids entering state i which is obtained by setting the i-th column of P to zero, i.e., letting $\sigma = (0,\dots 0,1,0,\dots)^\top$, where the entry 1 is at the i-th position, and h the i-th column of P. The taboo matrix ${}_iP$ provides a convenient sufficient condition for positive recurrence of P on a denumerable state-space. The precise statement is provided in the following theorem.

Theorem 4.1. *Let P be irreducible. If there exists a matrix norm $\|\cdot\|$ such that for at least one $i \in S$ it holds that $||{}_iP|| < 1$, then P is positive recurrent.*

Proof. First note that the (j,k)-th element of $\sum_{n=0}^{\infty}({}_iP)^n$ gives the expected number of visits to state k before jumping to state i when starting in state j. The mean recurrence time at state i is thus given by summing the i-th row of $\sum_{n=0}^{\infty}({}_iP)^n$, which is finite due to the norm condition. Therefore, state i is positive recurrent. From irreducibility of P it follows that all states are positive recurrent. □

Remark 4.1. *The applicability of Theorem 4.1 highly depends on the chosen matrix norm. To give an example consider a birth-death chain with the following transition probabilities*

$$P(0,0) = 1-p, \quad P(0,1) = p,$$

and for $i = 1,2,\dots,$

$$P(i,i-1) = 1-q, \quad P(i,i+1) = q,$$

for $p, q \in (0,1)$. For this example it easily follows that

$$\|{}_0P\|_\infty = 1, \quad \forall p,q \in (0,1),$$

so that Theorem 4.1 does not apply in case of the ∞-norm. For ease of presentation let $q \geq p$, then

$$\|{}_0P\|_v = \frac{1-q}{\alpha} + q\alpha.$$

It can be shown that $\|{}_0P\|_v < 1$ with an appropriate choice of $\alpha \geq 1$ in the v-norm. In particular,

$$\|{}_0P\|_v < 1 \quad \textit{iif} \quad \alpha \in \left(1, \frac{1-q}{q}\right).$$

Note that the right-bound for the feasible region of α such that $\|{}_0P\|_v < 1$ approaches the left bound for $q \uparrow 1/2$. This means that choosing α sufficiently close to 1 ensures that $\|{}_0P\|_v < 1$ for any given $q < 1/2$. Note that this corresponds to birth-death chain theory which states that the chain is positive recurrent when $P(i, i+1) < 1/2$ for all $i = 1, 2, \ldots$.

Elaborating on the taboo matrix of P with the conditions as stated above, the deviation matrix can alternatively be written as

$$D_P = (I - \Pi_P) \sum_{n=0}^{\infty} T^n (I - \Pi_P), \tag{4.9}$$

when $\|T\| < 1$, see Theorem C.1 in Appendix C for a proof and for more details see [99, 111]. Note that

$$(I - T)^{-1} = \sum_{n=0}^{\infty} T^n,$$

under the assumption that $||T|| < 1$. We call taboo matrix T *proper* if $||T|| < 1$. Provided that T defined in (4.8) is proper in case of the v-norm, the v-norm of $\pi_P^\top$ can be bounded by

$$\|\pi_P^\top\|_v \leq \frac{\pi_P^\top h \, \|\sigma^\top\|_v}{1 - ||T||_v}, \tag{4.10}$$

see [110]. The above bound also follows almost directly from Lemma C.1 in Appendix C.

The idea behind considering T rather than P, is that T might be constructed in such a way that the norm of T is strictly less than one. The following example illustrates the effect on $||T||$ from either removing the first column or first row. Note that removing the second column or second row would lead to again other values of $||T||$.

Example 4.2. For the Small Network, i.e., $P = P^s$, we find after removing the first column

$$||T||_v = \max\{\alpha p, 1 - q\}.$$

Removing the first row leads to

$$||T||_v = \max\left\{0, \frac{(1-\alpha)q}{\alpha} + 1\right\}.$$

For the Ring and the Star networks we present the resulting norms for $||T||_v$ (including the ∞-norm by letting α tend to one) and $||T||_\infty$, respectively, in Table 4.1 and Table 4.2. ◁

Removing:	Ring (i.e., $P = P^\circ(n)$)	Star (i.e., $P = P^\star(n)$)
1st row of P	$\frac{b}{\alpha} + 1 - 2\,b + \alpha\, b$	$\frac{b}{\alpha} + 1 - 2\,b + \alpha\, b$
1st column of P	$\max\{\alpha\, b + \alpha^{n-1}\, b, \frac{b}{\alpha} + 1 - 2\,b + b\,\alpha\}$	$\max\left\{\gamma, \frac{\alpha\,\beta}{n-1}\frac{1-\alpha^{n-1}}{1-\alpha}\right\}$

Tab. 4.1: The v-norm for different choices for T (including the ∞-norm).

In the following we discus a general way of choosing T. Let $P_{\bullet j}$ denote the j-th column of P. For a column vector x we let $\|x\|_{\inf} = \inf_i |x_i|$. We denote the j-th unit vector by e_j, i.e., e_j has all elements zero except for the j-th element which is equal to 1.

Removing:	Ring (i.e., $P = P^\circ(n)$)	Star (i.e., $P = P^\star(n)$)
1st row of P	1	$\max\{\gamma, (n-1)(1-\gamma)\}$
1st column of P	1	$\gamma + \frac{\beta}{n-1}$

Tab. 4.2: The infinity norm for different choices for T.

Lemma 4.1. *Let P be a Markov transition matrix on S. Let j^* be the column index with maximal value $\|P_{\bullet j}\|_{\inf}$. If $||P_{\bullet j^*}||_{\inf} > 0$, let $h = P_{\bullet j^*}$ and $\sigma = e_{j^*}$, then for T defined as in* (4.8) *it holds that $||T||_v < 1$, where $v \equiv 1$.*

Proof. Without loss of generality assume that after appropriate relabelling of the states $j^* = 0$. Let $||P_{\bullet 0}||_v = q > 0$. Removing the first column from P thus decreased the row sum of each row of P by at least q, which implies the desired result. □

4.2.2 Condition Number Perturbation Bounds for Finite Chains

Several condition numbers have been proposed in the literature for finite Markov chains for which $S = \{0, 1, \ldots, n-1\}$, see [57] for an overview. We keep the numbering as in [57], where seven different condition numbers were discussed. Moreover it is shown in [57] and [119] that condition numbers κ_3 and κ_6, to be defined presently, outperform the other condition numbers, while the choice between κ_3 and κ_6 depends on the choice of norms. Condition number κ_3 is given by [90, 122]:

$$\kappa_3(P) = \frac{\max_j (D_P(j,j) - \min_i D_P(i,j))}{2}$$

and leads to the following bound:

$$||\pi_R^\top - \pi_P^\top||_1 \le \kappa_3(P)||R - P||_\infty.$$

Alternatively, condition number κ_6 in [167] is given by:

$$\kappa_6(P) = \frac{1}{2} \max_{i,j} \sum_{k=0}^{n-1} |D_P(i,k) - D_P(j,k)|,$$

and gives the following bound:

$$||\pi_R^\top - \pi_P^\top||_\infty \le \kappa_6(P)||R - P||_\infty.$$

Example 4.3. The condition numbers for the Markov chains introduced in Example 4.1 are as follows:

$$\kappa_3(P^s) = \frac{1}{2(p+q)} \quad \text{and} \quad \kappa_6(P^s) = \frac{1}{p+q},$$

$$\kappa_3(P^\circ(n)) = \frac{\lfloor \frac{n}{2} \rfloor (n - \lfloor \frac{n}{2} \rfloor)}{4\,b\,n},$$

$$\kappa_6(P^\circ(n)) = \frac{1}{2} \sum_{k=0}^{n-1} \left| D_{P^\circ(n)} \left(\left\lfloor \frac{n}{2} \right\rfloor + 1, k-1 \right) - D_{P^\circ(n)}(1, k-1) \right|,$$

and

$$\kappa_3(P^\star(n)) = \frac{1}{2(1-\gamma)} \quad \text{and} \quad \kappa_6(P^\star(n)) = \frac{1}{1-\gamma}.$$

It is worth noting that $\kappa_3(P^\circ(n))$ grows linearly in n. As the condition number applies to the 1-norm of $\pi_R^\top - \pi_P^\top$, which is bounded by 1, the bound becomes thus trivial for large n. For the Star Network, κ_3 and κ_6 do not depend on n but become trivial for γ close to 1.

The fact that κ_3 and κ_6 behave so different for the Ring and the Star networks stems from the fact that both condition numbers are defined via the deviation matrix. The elements of the deviation matrix are related to mean recurrence times of the corresponding Markov chain, see [144, 101]. Specifically, in the Ring Network the length of a path from, say, node 0 to node $\lfloor n/2 \rfloor$ grows with n, whereas in the Star Network any node can reached from any other node in 2 steps. $\lhd$

It is known that $\kappa_3(P) < \kappa_6(P)$ (in fact it holds that $2\kappa_3(P) \leq \kappa_6(P)$), see [122]). Note that this inequality implies for the Ring Network that $\kappa_6(P^\circ(n))$ tends to infinity as well. In [122] it is shown that $\kappa_3(P) \geq (n-1)/(2n)$, with n being the size of transition matrix, and a Markov chain is provided for which equality is reached. As we will discuss in the subsequent section, $\kappa_6(P)$ may be preferable to $\kappa_3(P)$ in case bounds on perturbations of expected rewards are considered.

4.2.3 The Choice of Norms in Perturbation Analysis

In bounding perturbations it is important to understand how a perturbation of the Markov chain affects the steady-state reward. Put differently, using the notation as already introduced in the introduction, relating a perturbation bound for $||\pi_R^\top - \pi_P^\top||$ to that of $|\pi_R^\top f - \pi_P^\top f|$ is of importance in applications. The following lemma formalizes how the steady-state reward can be bounded via perturbation bounds for $||\pi_R^\top - \pi_P^\top||$ in case of different norms.

Lemma 4.2. *For arbitrary vectors $\widetilde{\mu}$ and μ on $\mathbb{R}^S$ and cost function $f \in \mathbb{R}^S$ such that $|\widetilde{\mu}^\top f - \mu^\top f| < \infty$ it holds*

$$|\widetilde{\mu}^\top f - \mu^\top f| \leq \begin{cases} ||\widetilde{\mu}^\top - \mu^\top||_\infty \, ||f||_\infty \\ \\ ||\widetilde{\mu}^\top - \mu^\top||_1 \, ||f||_1 \\ \\ ||\widetilde{\mu}^\top - \mu^\top||_v \, ||f||_v \end{cases}.$$

Proof. By simple algebra,

$$|\widetilde{\mu}^\top f - \mu^\top f| \leq \sum_i |\widetilde{\mu}_i - \mu_i| \, |f_i| \leq \sup_j |f_j| \sum_i |\widetilde{\mu}_i - \mu_i| = ||\widetilde{\mu}^\top - \mu^\top||_\infty \, ||f||_\infty.$$

For the last inequality, which coincides with the second inequality in case of $\alpha = 1$, note that

$$|\widetilde{\mu}^\top f - \mu^\top f| \leq \sum_i |\widetilde{\mu}_i - \mu_i| \, |f_i|$$

$$= \sum_i |\widetilde{\mu}_i - \mu_i| v_i \frac{|f_i|}{v_i}$$
$$\leq \left(\sup_j \frac{|f_j|}{v_j} \right) \sum_i |\widetilde{\mu}_i - \mu_i| v_i$$
$$= ||\widetilde{\mu}^\top - \mu^\top||_v \, ||f||_v,$$

which concludes the proof. □

In this chapter we study the case that μ in Lemma 4.2 is a stationary distribution. Lemma 4.2 illustrates that there is a trade-off in the choice of norms. Indeed, since $\|\pi_R^\top - \pi_P^\top\|_1 \leq \|\pi_R^\top - \pi_P^\top\|_\infty$ it seems attractive to ask for perturbation bounds on $\|\pi_R^\top - \pi_P^\top\|_1$. The downside is that this choice affects the norm of the reward vector, in particular, it holds that $\|f\|_\infty \leq \|f\|_1$. As an illustration, consider the following example of a finite Markov chain. Let P be the transition matrix of a $M/M/1/N$ queue, where N is the size of the buffer of the queue including the service place, and suppose that we are interested in the effect that replacing P by R has on the stationary queue length. More specifically, let $f_l(s) = s$, for $s \in S = \{0, 1, \ldots, N\}$, and note that

$$\|f_l\|_1 = \frac{N(N+1)}{2} > N = \|f_l\|_\infty.$$

In the light of Lemma 4.2, in bounding $|\pi_P^\top f_l - \pi_R^\top f_l|$ the smaller bound on the norm distance of $\pi_R^\top - \pi_P^\top$ by applying the 1-norm might be outweighed by the increase in norm for the reward. If, on the other side, one is only interested in an overflow probability, i.e., $f_p(s) = 0$ for $s < N$ and $f_p(N) = 1$, then $\|f_p\|_1 = \|f_p\|_\infty = 1$ and the 1-norm bound for $\pi_R^\top - \pi_P^\top$ is appropriate. Another example where this norm trade-off is relevant is in the analysis of the 'wisdom of crowds' phenomenon in social networks, [77]. Here, f represents a belief vector with bounded support, i.e., $f(s) \in [a, b]$ for $a < b \in \mathbb{R}$, and $\pi_P^\top f$ is the consensus reached in the social network modelled by P. From the above discussion it is clear that the choice of the norm for evaluating $\pi_R^\top - \pi_P^\top$ depends on the application.

In the light of the above discussion it is worth noting that the v-norm can be adjusted to the problem under consideration. To see this, recall that we have assumed that v is of the form $v(i) = \alpha^i$, $i \in S$, with α some unspecified constant. Let us express this dependency of v on α here by writing v_α. Hence, the best bound for $|\widetilde{\mu}^\top f - \mu^\top f|$ by means of the v-norm is given by the solution of the following minimization problem

$$|\widetilde{\mu}^\top f - \mu^\top f| \leq \min_\alpha ||\widetilde{\mu}^\top - \mu^\top||_{v_\alpha} \, ||f||_{v_\alpha}. \tag{4.11}$$

The upside of this minimization is that it trades off the effect the norm has on the reward and the vector distance. The downside is of course that the minimization itself can be rather demanding as $||\widetilde{\mu}^\top - \mu^\top||_{v_\alpha}$ or a bound thereof typically has a complex form. For denumerable Markov chains, v can be constructed via a Lyapunov-type of drift condition; see [140] for details.

4.2.4 Perturbation Bounds

In perturbation analysis, D_P occurs in conjunction with a perturbation matrix $\Delta = R - P$ which has row sums equal to zero. From $\Delta(I - \Pi_P) = \Delta$ and (4.9) it follows that

$$\Delta(I - T)^{-1}(I - \Pi_P) = \Delta D_P \tag{4.12}$$

and instead of D_P for perturbation bounds it suffices to consider

$$(I - T)^{-1}(I - \Pi_P).$$

Note that due to the fact that $\Delta(I - T)^{-1}$ fails to have row sums equal to zero, the term $I - \Pi_P$ on the LHS in (4.12) cannot be disregarded. In other words, $\Delta(I - T)^{-1} \neq \Delta(I - T)^{-1}(I - \Pi_P)$, except for special cases. By simple algebra, it holds for Markov transition matrices R and P that

$$\pi_R^\top = \pi_P^\top + \pi_R^\top (R - P) D_P \tag{4.13}$$

$$= \pi_P^\top + \pi_R^\top (R - P)(I - T)^{-1}(I - \Pi_P). \tag{4.14}$$

Remark 4.2. *The above formula is called* update formula *and allows for deriving a perturbation bound. Using the fact that* $||\pi_R^\top||_\infty = 1$, (4.14) *yields*

$$||\pi_R^\top - \pi_P^\top||_\infty \leq ||R - P||_\infty \, ||(I - T)^{-1}(I - \Pi_P)||_\infty,$$

which provides a first perturbation bound. Put differently $||(I - T)^{-1}(I - \Pi_P)||_\infty$ *yields a condition number.*

Repeated insertion of the expression for $\pi_R^\top$ in (4.13) on the RHS of (4.13), yields

$$\pi_R^\top = \pi_P^\top \sum_{k=0}^{N} ((R - P) D_P)^k + \pi_R^\top ((R - P) D_P)^{N+1}. \tag{4.15}$$

We call

$$B(R, P) = \lim_{N \to \infty} \pi_R^\top ((R - P) D_P)^N$$

the *bias term*, provided that the limit exists. Letting N tend to infinity in (4.15) we arrive at

$$\pi_R^\top = \pi_P^\top \sum_{k=0}^{\infty} ((R - P) D_P)^k \; + B(R, P) \tag{4.16}$$

$$= \pi_P^\top (I - (R - P) D_P)^{-1} + B(R, P),$$

provided the series exists and the bias term is finite. As we will explain in the following, the bias term is typically zero in case that R and P are uni-chain. The series in (4.16) already appears without the bias term in [166]. It has been rediscovered in [47] and extended to Markov chains on a general state-space in [91], both references study problem classes where the bias term is zero.

In deriving the series expansion in (4.16) we required that the stationary distribution π_R exists. As the next theorem shows, convergence of the series already implies existence of π_R. Moreover, we provide sufficient conditions for the bias term to be equal to the zeros vector.

Theorem 4.2. *Let P be irreducible, aperiodic and positive recurrent. Suppose that the series in* (4.16) *converges to some finite limit $\mu^\top$, i.e., let*

$$\mu^\top = \pi_P^\top (I - (R - P) D_P)^{-1}.$$

(i) If $\mu_i \geq 0$, for $i \in S$, then μ is a stationary distribution of R.

(ii) If R is irreducible and aperiodic and there exists $i \in S$ such that $||_i R|| < 1$, then μ is the unique stationary distribution of R and $B(R, P)$ is the zeros vector.

Proof. To see that μ is invariant with respect to R, note that,

$$\Pi_P + (I - P)D_P = I.$$

Multiplying the above equation from the left by $\mu^\top$, yields

$$\pi_P^\top + \mu^\top (I - P)D_P = \mu^\top. \tag{4.17}$$

By simple algebra,

$$\begin{aligned}
\mu^\top &= \pi_P^\top \sum_{k=0}^{\infty} ((R - P)D_P)^k \\
&= \pi_P^\top + \pi_P^\top \sum_{k=1}^{\infty} ((R - P)D_P)^k \\
&= \pi_P^\top + \pi_P^\top \sum_{k=0}^{\infty} ((R - P)D_P)^k (R - P)D_P \\
&= \pi_P^\top + \mu^\top (R - P)D_P.
\end{aligned} \tag{4.18}$$

Subtracting (4.17) from (4.18) yields

$$\mu^\top (I - R)D_P = 0.$$

Existence of D_P implies that $D_P = (I - P + \Pi_P)^{-1} - \Pi_P$, see (4.5). Since $(I - R)\Pi_P = 0$, it holds that

$$\mu^\top (I - R)(I - P + \Pi_P)^{-1} = 0.$$

Multiplying the above equation from the right with $(I - P + \Pi_P)$ yields $\mu = \mu R$, which shows that μ is invariant to R. Further, multiplying (4.17) from the right with an appropriate column vector of ones, i.e., $\bar{1}$, shows

$$\pi_P^\top \bar{1} + \mu^\top (I - P)D_P \bar{1} = \mu^\top \bar{1} \Leftrightarrow \mu^\top \bar{1} = 1$$

since $(I - P)D_P \bar{1} = (I - \Pi_P)\bar{1} = 0$. This shows that μ sums up to 1. Provided that μ is component-wise a non-negative vector, μ is a stationary distribution, which proves part (i).

For part (ii), note that by Theorem 4.1 it follows that R is positive recurrent. This together with the assumption that R is irreducible and aperiodic implies that R is ergodic and

$$\lim_{n \to \infty} R^n = \Pi_R, \tag{4.19}$$

where Π_R is a matrix with all rows equal to $\pi_R^\top$ and π_R is the unique stationary distribution of R. Since all rows of Π_R are identical to $\pi_R^\top$ and $\mu^\top \bar{1} = 1$, it holds that

$$\mu^\top \Pi_R = \pi_R^\top. \tag{4.20}$$

We have already shown that μ is an invariant distribution of R. This together with (4.19) and (4.20) yields

$$\mu^\top = \lim_{n\to\infty} \mu^\top R^n = \mu^\top \Pi_R = \pi_R^\top.$$

Uniqueness of the solution follows from ergodicity of R and the bias term is consequently the zeros vector, which concludes the proof. □

Remark 4.3. *Part (i) of Theorem 4.2 applies in case that R is a multi-chain with transient states. In this case the stationary distribution is not unique. This can be nicely explained via the bias term. As the bias term depends on P, it carries information on the Markov chain that is used in approximating π_R. Letting P tend to R, the limit of $B(R,P)$ typically will not tend to zero if R is a multi-chain. This phenomenon is studied in the literature on singular perturbations, see, for example, [117, 195, 196, 18].*

Note that in order to prove uniqueness of the stationary distribution we need the conditions put forward in part (ii) of Theorem 4.2.

The series in (4.16) can be facilitated for deriving perturbation bounds by

$$\begin{aligned}
\pi_R^\top - \pi_P^\top &= \pi_P^\top \sum_{k=1}^{\infty} ((R-P)D_P)^k + B(R,P) \\
&= \pi_P^\top (R-P)D_P \sum_{k=0}^{\infty} ((R-P)D_P)^k + B(R,P) \\
&= \pi_P^\top (R-P)D_P (I-(R-P)D_P)^{-1} + B(R,P). \qquad (4.21)
\end{aligned}$$

Following the above line of equations, bounding $\pi_R^\top - \pi_P^\top$ requires bounding $(I-(R-P)D_P)^{-1}$. We will show that the conditions put forward in the following lemma not only imply norm bounds for $(I-(R-P)D_P)^{-1}$ but also imply that $B(R,P)$ is the zeros vector.

Lemma 4.3. *For any matrix norm it holds with the above notation that:*

(i) If $||(R-P)D_P|| < 1$, then

$$||(I-(R-P)D_P)^{-1}|| \leq \frac{1}{1-||(R-P)D_P||},$$

(ii) if $||R-P||\,||D_P|| < 1$, then

$$||(I-(R-P)D_P)^{-1}|| \leq \frac{1}{1-||R-P||\,||D_P||},$$

(iii) if $\|T\| + ||R-P||(1+||\pi_P^\top||) < 1$, then

$$||(I-(R-P)D_P)^{-1}|| \leq \frac{1-||T||}{1-||T||-||R-P||(1+||\pi_P^\top||)}.$$

In addition, any of the conditions (i), (ii) or (iii) implies that the bias term $B(R,P)$ equals the zeros vector.

Proof. We only provide a proof of part (iii) as the proofs of (i) and (ii) can be obtained from a similar (and simpler) line of arguments. Recall from (4.12) that

$$(R-P)D_P = (R-P)\sum_{k=0}^{\infty} T^k (I - \Pi_P).$$

By the condition it follows that $||T|| < 1$ and thus applying norms yields

$$||(R-P)D_P|| \leq ||R-P|| \frac{1 + ||\pi_P^\top||}{1 - ||T||}. \tag{4.22}$$

Our condition $\|T\| + ||R-P||(1+||\pi_P^\top||) < 1$ is equivalent to the expression on the above RHS being strictly less than 1. This implies that the Neumann series $\sum_{k=0}^{\infty}((R-P)D_P)^k$ converges. Consequently $(I-(R-P)D_P)$ is invertible with norm bounded by

$$\begin{aligned} ||(I-(R-P)D_P)^{-1}|| &\leq \sum_{k=0}^{\infty} ||(R-P)D_P||^k \\ &= \frac{1}{1 - ||(R-P)D_P||}. \end{aligned}$$

Inserting the bound in (4.22) in the expression on the above RHS concludes the proof of the statement.

For the proof of the last part of the lemma, note that $||\pi_R^\top((R-P)D_P)^N|| \leq ||\pi_R^\top||\,||(R-P)D_P||^N$, so that $||(R-P)D_P|| < 1$ implies convergence of $||\pi_R^\top((R-P)D_P)^n||$ to zero as n tends to infinity. □

Remark 4.4. *It is worth noting that* $||(R-P)D_P|| < 1$ *typically fails in case* R *is a multi-chain. Put differently, while in principle the results in the remainder of this chapter apply to* R *being a multi-chain, we have found no example of a pair* R, P *with* R *a multi-chain and* P *a uni-chain such that* $||(R-P)D_P|| < 1$. *We conjecture that* $||(R-P)D_P|| < 1$ *rules out the case that* R *is a multi-chain but we have not been able to prove this so far.*

Note that

$$||(R-P)D_P|| \leq ||R-P||\,||D_P|| \leq \frac{||R-P||(1+||\pi_P^\top||)}{1-||T||}$$

implies that the bounds put forward in Lemma 4.3 are increasingly limited in their applicability, while the evaluation of the bounds becomes simpler. In fact, computing $||(R-P)D_P||$ is often not feasible as D_P is either not known in closed form or is prohibitively complex in general, see [61, 92, 125]. For the Markov chains in Example 4.1, D_P is known in explicit form. For this type of problems it makes sense to apply the norm bounds put forward in Lemma 4.3 to (4.21). More specifically, assuming $||(R-P)D_P|| < 1$ let

$$\Delta_{\text{DB}}(R,P) := \frac{||\pi_P^\top(R-P)D_P||}{1 - ||(R-P)D_P||},$$

then

$$||\pi_R^\top - \pi_P^\top|| \leq \Delta_{\text{DB}}(R,P), \tag{4.23}$$

which we will call the *direct bound* (DB).

Remark 4.5. *The bound in* (4.23) *has the following nice feature. Let P and R be two Markov chains with $P \neq R$ but with the same stationary distribution. Then,* (4.23) *detects this and yields the correct value 0, whereas condition number type bounds yield a non-zero bound.*

The next bound can serve as alternative in case D_P is difficult to find. It follows from replacing $(R-P)D_P$ in (4.23) with the taboo matrix representation and bounding the result via (4.22). Specifically, this leads to

$$||\pi_R^\top - \pi_P^\top|| \leq ||\pi_P^\top|| \, ||R-P|| \frac{1+||\pi_P^\top||}{1-||T||} \frac{1-||T||}{1-||T||-||R-P||(1+||\pi_P^\top||)}.$$

Let

$$\Delta_{\text{SSB}}(R,P) := ||\pi_P^\top|| \, ||R-P|| \frac{1+||\pi_P^\top||}{1-||T||-||R-P||(1+||\pi_P^\top||)}, \tag{4.24}$$

provided that $||T|| + ||R-P||(1+||\pi_P^\top||) < 1$. Then,

$$||\pi_R^\top - \pi_P^\top|| \leq \Delta_{\text{SSB}}(R,P) \tag{4.25}$$

and the bound $\Delta_{\text{SSB}}(R,P)$ in (4.24) is called *Strong Stability Bound* (SSB) in the literature [111]. For applications of SSB, we refer to [1, 6, 7, 40, 41, 133].

Based on the SSB we get the robust sensitivity bound for π_P stated in Lemma 4.4, it is a bound for perturbation derivatives in all directions.

Lemma 4.4. *Provided that $||T|| < 1$, it holds that*

$$\frac{d||\pi_P||}{d||P||} = \lim_{P,R:||R-P||\to 0} \frac{||\pi_R^\top - \pi_P^\top||}{||R-P||} \leq \frac{||\pi_P^\top||(1+||\pi_P^\top||)}{1-||T||}.$$

Proof. Provided that $||T|| < 1$ we may take $||R-P||$ sufficiently small such that (4.24) holds. Dividing inequality (4.25) by $||R-P||$ and letting $||R-P||$ tend to zero proves the claim. □

An obvious improvement of the bound in (4.24) is to replace $||\pi_P^\top|| \, ||R-P||$ by $||\pi_P^\top(R-P)||$; see Remark 4.5.

While P and π_P are fixed, and T offers in practice only limited flexibility, R is a free variable of perturbation bounds. Essentially, the direct bound and SSB only apply if R is not too far away from P, i.e., if $||R-P||$ is small. This is the major drawback of this type of perturbation bounds compared to condition number bounds. To overcome this drawback, we may scale the perturbation such that the perturbation bounds do apply. To see this, consider the scaled model in (4.3), where the static perturbation is replaced by a scaled one, i.e., we perturb P by $\theta(R-P)$ and denote the resulting transition matrix by $P(\theta)$. Now, θ can be chosen such that the norm bounds apply to $\theta||R-P||$. For example, the condition on the applicability for SBB in (4.24) translates to

$$||T|| + \theta||R-P||(1+||\pi_P^\top||) < 1 \quad \text{iff} \quad 0 \leq \theta < \frac{1-||T||}{||R-P||(1+||\pi_P^\top||)}.$$

We call the upper bound for θ on the RHS above the *domain of SBB with respect to R*.

In the following we take an alternative route for obtaining a perturbation bound. Starting point is (4.13) but other than for deriving (4.16) we now only perform the insertion operation K times, leading to

$$\pi^\top_{P(\theta)} = \pi^\top_P \sum_{k=0}^{K} (\theta(R-P)D_P)^k + \pi^\top_{P(\theta)}(\theta(R-P)D_P)^{K+1}. \tag{4.26}$$

For $K \geq 1$, equation (4.26) yields the following bound:

$$\|\pi^\top_{P(\theta)} - \pi^\top_P\| \leq \left\| \pi^\top_P \sum_{k=1}^{K} (\theta(R-P)D_P)^k \right\| + \|\pi^\top_{P(\theta)}(\theta(R-P)D_P)^{K+1}\|.$$

Obviously, $\pi^\top_{P(\theta)}$ is not known for the actual bound, and we use the fact that

$$\begin{aligned}\|\pi^\top_{P(\theta)}(\theta(R-P)D_P)^{K+1}\| &\leq \|\pi^\top_{P(\theta)}\| \|(\theta(R-P)D_P)^{K+1}\| \\ &\leq c_{\|\cdot\|} \|(\theta(R-P)D_P)^{K+1}\|,\end{aligned}$$

where we define the norm dependent upper bound $c_{\|\cdot\|}$ for $\|\pi^\top_{P(\theta)}\|$ as follows

$$c_{\|\cdot\|} = \sup_{Q \in \mathbb{P}(S)} \|\pi^\top_Q\|, \tag{4.27}$$

where $\mathbb{P}(S)$ represents all stochastic matrices defined on S. In case the ∞-norm (resp., 1-norm) is applied to $\pi^\top_{P(\theta)}$ we thus have

$$\|\pi^\top_{P(\theta)}(\theta(R-P)D_P)^{K+1}\| \leq \|(\theta(R-P)D_P)^{K+1}\|.$$

For the general v-norm, a bound $c_{\|\cdot\|}$ can be obtained from (4.10).

The *series expansion perturbation bound of order* K (SEB(K)) is now introduced by

$$\Delta_{\mathrm{SEB}(K)}(P(\theta), P) := \left\| \pi^\top_P \sum_{k=1}^{K} (\theta(R-P)D_P)^k \right\| + c_{\|\cdot\|} \|(\theta(R-P)D_P)^{K+1}\|, \tag{4.28}$$

where $c_{\|\cdot\|}$ is as defined in (4.27), and it holds that

$$\|\pi^\top_{P(\theta)} - \pi^\top_P\| \leq \Delta_{\mathrm{SEB}(K)}(P(\theta), P),$$

for $\theta \in [0, 1]$.

Remark 4.6. *Note that we may bound* (4.28) *as follows*

$$\|\pi^\top_{P(\theta)} - \pi^\top_P\| \leq \sum_{k=1}^{K} \|\pi^\top_P((R-P)D_P)^k\| \theta^k + c_{\|\cdot\|} \|((R-P)D_P)^{K+1}\| \theta^{K+1},$$

so that the polynomial terms only have to be calculated once and can be used for evaluating the bound for different values of θ. This is allows for fast computation and memory efficiency but, due to the additional bounding, the numerical quality of the bound decreases.

From
$$\|((R-P)D_P)^{K+1}\| \leq \|(R-P)D_P\|^{K+1}$$
it follows that the series in (4.16) converges for $P(\theta) = P + \theta(R-P)$ at least for $\theta < (\|(R-P)D_P\|)^{-1}$. Hence, for θ sufficiently small
$$\pi_P^\top \sum_{k=0}^{K} (\theta(R-P)D_P)^k \tag{4.29}$$
provides an approximation of $\pi_{P(\theta)}$, where the error is bounded by some constant times $\theta^{K+1}\|((R-P)D_P)^{K+1}\|$. The series put forward in (4.29) is called series expansion approximation of order K. Letting K tend to infinity in (4.29) we obtain for θ sufficiently small that
$$\pi_{P(\theta)}^\top = \pi_P^\top \sum_{k=0}^{\infty} \theta^k((R-P)D_P)^k.$$
Note that the above series expansion implies that $\pi_{P(\theta)}$ tends to π_P as θ tends to zero; for more details we refer to [93, 92].

To test the performance of the different bounds in the scaled perturbation setting (i.e., (4.3)) we will investigate the relative error of the perturbation bounds. Clearly, a better bound results in a smaller relative error. Consider a condition number bound for $||\pi_{P(\theta)}^\top - \pi_P^\top||$. The following reasoning only uses the basic definition of a CNB in (4.1) so that the arguments apply to the condition number bounds for finite chains discussed in Section 4.2.2 and the CNB in Remark 4.2 as well. Generally speaking, let $\Delta_{\mathrm{CNB}(\kappa)}(P(\theta), P) = \theta||R-P||\kappa$ denote a condition number bound for $\|\pi_{P(\theta)}^\top - \pi_P^\top\|$. In the following let $||\pi_{P(\theta)}^\top - \pi_P^\top|| > 0$. The relative error inferred by using $\theta||R-P||\kappa$ rather than $||\pi_{P(\theta)}^\top - \pi_P^\top||$ is given by
$$\begin{aligned} \eta_{\mathrm{CNB}}(\theta) &:= \frac{\Delta_{\mathrm{CNB}(\kappa)}(P(\theta), P) - ||\pi_{P(\theta)}^\top - \pi_P^\top||}{||\pi_{P(\theta)}^\top - \pi_P^\top||} \\ &= \frac{\theta||R-P||\kappa - ||\pi_{P(\theta)}^\top - \pi_P^\top||}{||\pi_{P(\theta)}^\top - \pi_P^\top||} \\ &= \frac{\theta||R-P||\kappa}{||\pi_{P(\theta)}^\top - \pi_P^\top||} - 1. \end{aligned} \tag{4.30}$$
Note that this relative error is by definition ≥ 0. In the same vein, if $||T|| + \theta||R - P||(1 + ||\pi_P^\top||) < 1$, we define the relative error of SSB by
$$\eta_{\mathrm{SSB}}(\theta) := \frac{\Delta_{\mathrm{SSB}}(P(\theta), P) - ||\pi_{P(\theta)}^\top - \pi_P^\top||}{||\pi_{P(\theta)}^\top - \pi_P^\top||}.$$
If $\|(R-P)D_P\| < 1$, we define the relative error of DB by
$$\eta_{\mathrm{DB}}(\theta) := \frac{\Delta_{\mathrm{DB}}(P(\theta), P) - ||\pi_{P(\theta)}^\top - \pi_P^\top||}{||\pi_{P(\theta)}^\top - \pi_P^\top||},$$
and the relative error of SEB of order K, for $K \geq 1$, by
$$\eta_{\mathrm{SEB}(K)}(\theta) := \frac{\Delta_{\mathrm{SEB}(K)}(P(\theta), P) - ||\pi_{P(\theta)}^\top - \pi_P^\top||}{||\pi_{P(\theta)}^\top - \pi_P^\top||}.$$

Note that all relative errors are by definition greater than or equal to 0. The following theorem analyses the relative error of the discussed bounds. It shows that in general the relative error of a condition number bound converges for $\theta \downarrow 0$ to a finite non-zero value. This means that even for a small perturbation this bound has a significant relative error. The same holds true for the SSB, while the SEB(K)-based bounds have the desirable property that the relative error vanishes. Moreover, the rate of convergence of the relative error of SEB(K) can be explicitly computed. Before analysing the relative errors we first give an overview of the considered bounds.

Summary of all considered perturbation bounds:

$$\begin{aligned}
\Delta_{\mathrm{CNB}(\kappa)}(P(\theta), P) &= \theta ||R-P||\kappa, \\
\Delta_{\mathrm{SSB}}(P(\theta), P) &= \frac{||\pi_P^\top||(1+||\pi_P^\top||)}{\frac{1-||T||}{\theta||R-P||} - 1 - ||\pi_P^\top||}, \\
\Delta_{\mathrm{DB}}(P(\theta), P) &= \frac{\theta||\pi_P^\top(R-P)D_P||}{1-\theta||(R-P)D_P||}, \\
\Delta_{\mathrm{SEB}(K)}(P(\theta), P) &= \left\| \pi_P^\top \sum_{k=1}^{K} (\theta(R-P)D_P)^k \right\| + c_{||\cdot||} \|(\theta(R-P)D_P)^{K+1}\|.
\end{aligned}$$

Theorem 4.3 (Relative Errors)**.** *Let* $\|\pi_{P(\theta)}^\top - \pi_P^\top\| > 0$*, for all* $\theta \in (0,1]$.

(i) The relative error of the condition number bound (CNB) is given by

$$\eta_{CNB(\kappa)}(\theta) = \frac{||R-P||\kappa}{||\pi_{P(\theta)}^\top (R-P)D_P||} - 1,$$

and it holds that

$$\lim_{\theta\downarrow 0} \eta_{CNB(\kappa)}(\theta) = \frac{||R-P||\kappa}{||\pi_P^\top(R-P)D_P||} - 1 \geq 0,$$

where equality is only reached in the special case when $||R-P||\kappa$ *equals* $||\pi_P^\top(R-P)D_P||$.

(ii) Provided that $||T|| + \theta||R-P||(1+||\pi_P^\top||) < 1$*, the relative error of the strong stability bound (SSB) is given by*

$$\eta_{\mathrm{SSB}}(\theta) = \frac{||R-P||\,||\pi_P^\top||(1+||\pi_P^\top||)}{||\pi_{P(\theta)}^\top(R-P)D_P||(1-||T||-\theta||R-P||(1+||\pi_P^\top||))} - 1,$$

and it holds that

$$\lim_{\theta\downarrow 0} \eta_{\mathrm{SSB}}(\theta) = \frac{||R-P||\,||\pi_P^\top||(1+||\pi_P^\top||)}{||\pi_P^\top(R-P)D_P||(1-||T||)} - 1 \geq 0,$$

where equality is only reached in the special case when the numerator equals the denominator in the fraction.

(iii) Provided that $\theta\|(R-P)D_P\| < 1$*, the relative error of the direct bound (DB) is given by*

$$\eta_{\mathrm{DB}}(\theta) = \frac{\frac{\|\pi_P^\top (R-P)D_P\|}{1-\theta\|(R-P)D_P\|}}{||\pi_{P(\theta)}^\top (R-P)D_P||} - 1,$$

and it holds that $\lim_{\theta\downarrow 0} \eta_{\mathrm{DB}}(\theta) = 0$*.*

(iv) Provided that $\theta\|(R-P)D_P\| < 1$*, the relative error of the series expansion bound of order* $K \geq 1$ *(i.e., SEB(K)) is given by*

$$\eta_{SEB(K)}(\theta) = \frac{2c_{||\cdot||}\|((R-P)D_P)^{K+1}\|\theta^K}{||\pi_{P(\theta)}^\top (R-P)D_P||},$$

and it holds that $\eta_{SEB(K)}(\theta)$ *is of order* $O(\theta^{K-1})$*.*

Proof. All relative error expressions follow by simply inserting the different bound and using the result that

$$\pi_{P(\theta)}^\top - \pi_P^\top = \theta\pi_{P(\theta)}^\top (R-P)D_P,$$

in the denominator (see also (4.13)). E.g., for the CNB it holds that

$$\eta_{\mathrm{CNB}(\kappa)}(\theta) = \frac{\theta||R-P||\kappa}{||\pi_{P(\theta)}^\top - \pi_P^\top||} - 1 = \frac{||R-P||\kappa}{||\pi_{P(\theta)}^\top (R-P)D_P)||} - 1,$$

where we simplified the expression in the second equation. For the limit, we use that $\pi_{P(\theta)}$ tends to π_P as θ tends to zero, which follows from (4.28) for $K = 1$.

We now turn to the computing the relative error for the K-th order SE. Following (4.28) we can write

$$\eta_{\mathrm{SEB}(K)}(\theta) = \frac{\overbrace{\|\pi_P^\top \sum_{k=1}^{K}(\theta(R-P)D_P)^k\|}^{=:H} + c_{||\cdot||}\|(\theta(R-P)D_P)^{K+1}\|}{||\theta\pi_{P(\theta)}^\top (R-P)D_P||} - 1. \tag{4.31}$$

For H it holds that

$$H = \|\pi_P^\top \sum_{k=0}^{K-1}(\theta(R-P)D_P)^k \theta(R-P)D_P\|.$$

After some algebra,

$$H = \left\|\pi_P^\top \sum_{k=0}^{\infty}(\theta(R-P)D_P)^k \left[I - (\theta(R-P)D_P)^K\right]\theta(R-P)D_P\right\|$$

and using condition $\|\theta(R-P)D_P\| < 1$ together with Theorem 4.2 we arrive at

$$H = \left\|\pi_{P(\theta)}^\top \left[I - (\theta(R-P)D_P)^K\right]\theta(R-P)D_P\right\|,$$

which can be straightforwardly bounded by

$$H \leq \|\pi_{P(\theta)}^\top \theta(R-P)D_P\| + c_{||\cdot||}\|(\theta(R-P)D_P)^{k+1}\|.$$

Inserting the above bound for H into (4.31) yields for the relative error

$$\eta_{\mathrm{SEB}(K)}(\theta) \leq \frac{2c_{||\cdot||}\|((R-P)D_P)^{K+1}\|}{||\pi_{P(\theta)}^\top(R-P)D_P||}\theta^K.$$

We now turn to establishing the rate of convergence of $\eta_{\mathrm{SEB}(K)}(\theta)$. First, note that the nominator on the above RHS is of order $O(\theta^K)$. We now turn to the denominator. Evoking the update formula (4.16), it follows that

$$\pi_{P(\theta)}^\top(R-P)D_P = \pi_{P(\theta)}^\top - \pi_P^\top = \sum_{k=1}^{\infty}\theta^k((R-P)D_P)^k,$$

which shows that $\pi_{P(\theta)}^\top(R-P)D_P$ can be written as power series with leading term $\theta(R-P)D_P$. Finiteness of the matrices and their norms therefore implies that $\|\pi_{P(\theta)}^\top(R-P)D_P\|$ is of order $O(\theta)$. Hence, $\eta_{\mathrm{SEB}(K)}(\theta)$ is of order $O(\theta^{K-1})$.

□

Remark 4.7. *Theorem 4.3 illustrates a conceptual limitation of condition number bounds since the relative error of a condition number bound fails to tend to zero as θ tends to zero. The same holds for SSB.*

As Theorem 4.3 shows, perturbation bounds have the intrinsic drawback that in general the relative error does not vanish for small perturbations. For the condition number bounds their applicability is questionable due to the fact that those bounds grow linearly with respect to the perturbation size θ whereas SEB(K) shows that the dependence of $\pi_{P(\theta)}$ on θ is non-linear. SSB, even though not a linear type of bound, suffers from the problem that the domain of applicability is so small that the non-linearity of the functional form of SSB does not come into play. SEB(K) is of a polynomial type and has an asymptotic relative error with specified rate of convergence.

For an illustration of Theorem 4.3 we generated two random transition matrices P and R with 40 states. The random generation is done by drawing random numbers from $(0,1)$ and normalizing the rows so that they sum up to 1. Then we considered in case of the ∞-norm all perturbation bounds on the interval $\theta \in (0,1]$ together with the true perturbation effect $\|\pi_{P(\theta)}^\top - \pi_P^\top\|_\infty$ (the true effect was calculated numerically). The results can be found in Figure 4.1. Figure 4.1 shows that in this experiment all bounds, except for CNB, are similar in performance on the interval $\theta \in [0, 0.1]$. For $\theta > 0.1$ there arises a difference in performance, where the SEB of order $K = 3$ performs best. DB performs similar to SEB(1) on the interval $\theta \in (0, 0.3]$ but for $\theta > 0.3$ SEB(1) outperforms DB. This simple example illustrates that the CNB is apparently too general to be competitive compared to the other bounds. The differences become more apparent if we look at the relative errors for the different bounds plotted in Figure 4.2. The results for SSB are not plotted because the condition in part (ii) of Lemma 4.3 is not met.

Remark 4.8. *Provided that θ_0 exists such that $\theta_0||(R-P)D_P|| < 1$, it holds that*

$$\eta_{SEB(K)}(\theta) = O(\theta_0^{K-1}),$$

for $0 \leq \theta \leq \theta_0$.

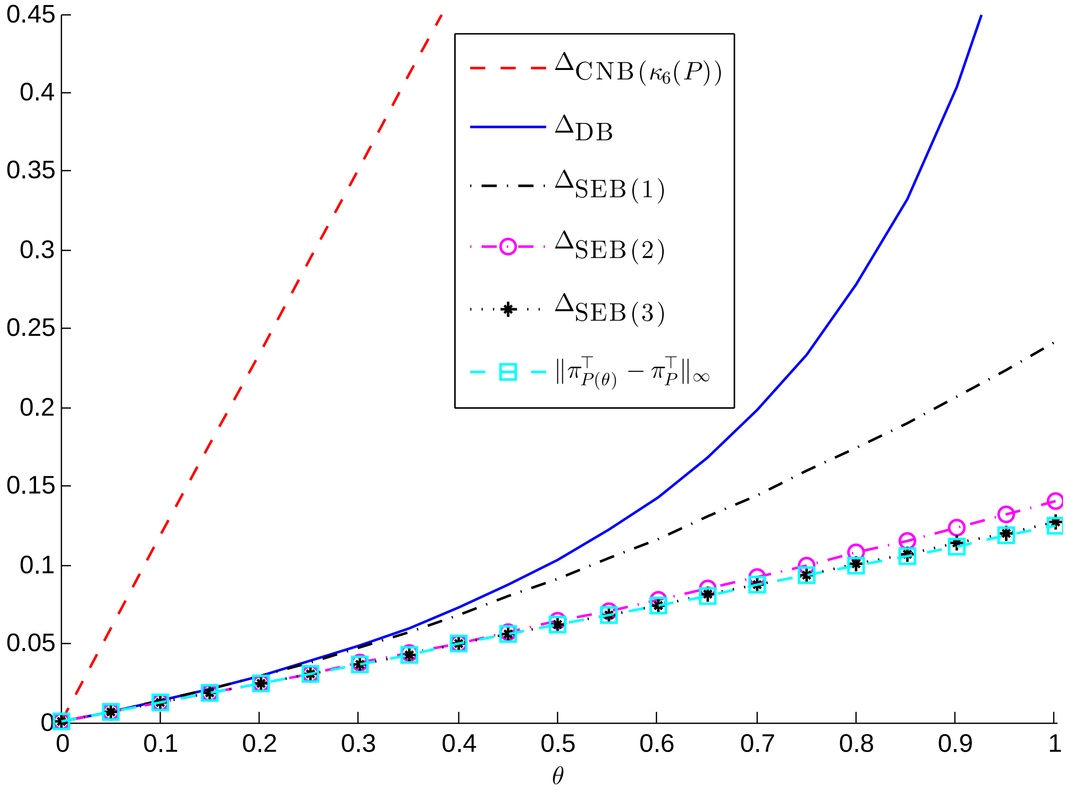

Fig. 4.1: Perturbation bounds for $\|\pi^\top_{P(\theta)} - \pi^\top_P\|_\infty$ with $\theta \in (0, 1]$, where $P(\theta) = (1 - \theta)P + \theta R$ for randomly generated P and R consisting of 40 states.

Remark 4.9. *The result put forward in Theorem 4.3 seems to contradict the fact that for finite Markov chains it holds that*

$$\left| \frac{(\pi_R)_i - (\pi_P)_i}{(\pi_R)_i} \right| \leq 2\eta n + O(\eta^2),\ i \in S = \{0, \ldots, n-1\}, \tag{4.32}$$

where η is bounded by $||R - P||$ and n denotes the size of the state-space, which indicates that the relative element-wise error in using π_P as a substitute for π_R tends to zero as P approaches R, see [109, 105, 158] for details. Note that above equation is equivalent to

$$|(\pi_R)_i - (\pi_P)_i| \leq (\pi_R)_i \left(2\eta n + O(\eta^2)\right),\ i \in S = \{0, \ldots, n-1\},$$

and reads in norm-version, using, for example, the ∞-norm (or 1-norm),

$$||\pi^\top_R - \pi^\top_P|| \leq 2\eta n + O(\eta^2).$$

Hence, the element-wise relative error result in (4.32) *is a statement about continuity of finite Markov chains and does not imply that the relative error in predicting the true norm distance between π_R and π_P by a CNB becomes small; for details compare the definition of the relative error in* (4.30) *with that in* (4.32)*.*

We conclude this section by presenting an interesting result for stability theory in which the objective is to find an upper bound for θ for which Markov chain with $P(\theta)$ remains stable.

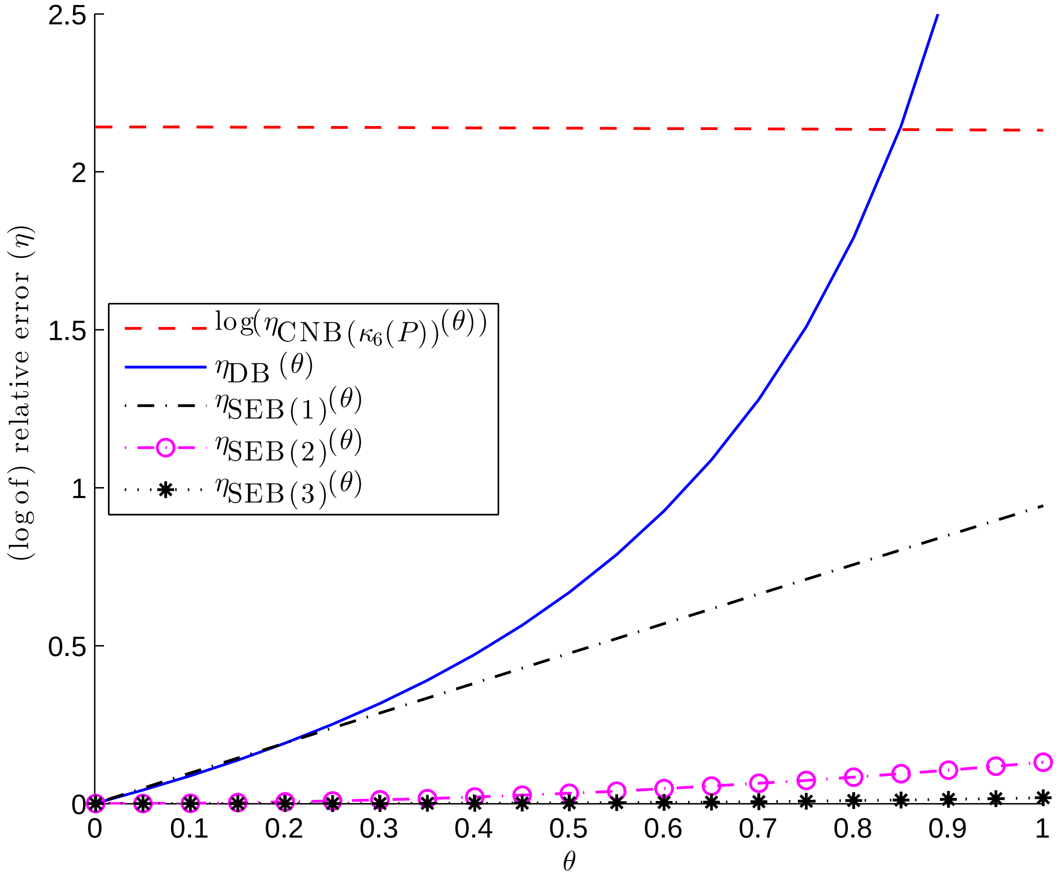

Fig. 4.2: Relative errors of the perturbation bounds for $\|\pi^\top_{P(\theta)} - \pi^\top_P\|_\infty$ with $\theta \in (0, 1]$, where $P(\theta) = (1-\theta)P + \theta R$ for randomly generated P and R consisting of 40 states.

Corollary 4.1. *Consider the model $P(\theta) = (1-\theta)P + \theta R$, $\theta \in [0, 1)$, with P aperiodic, irreducible and positive recurrent. If there exists an $i \in S$ such that*

$$\theta < \frac{1 - ||_i P||}{||R - P||},$$

then $P(\theta)$ has a unique stationary distribution.

Proof. Note that $P(\theta)$ is aperiodic and irreducible for $\theta \in [0, 1)$. It remains to be shown that $P(\theta)$ is positive recurrent. By computation,

$$\begin{aligned} ||_i(P(\theta))|| &= ||_i((1-\theta)P + \theta R)|| \\ &\leq ||_i P + \theta(R - P)|| \\ &\leq ||_i P|| + \theta||R - P||. \end{aligned}$$

Hence, provided that θ satisfies $||_i P|| + \theta||R - P|| < 1$, it follows $||_i(P(\theta))|| < 1$ and by Theorem 4.1 we conclude that $P(\theta)$ is positive recurrent. Solving θ out of $||_i P|| + \theta||R - P|| < 1$ concludes the proof. □

Remark 4.10. *Note that from Corollary 4.1 it follows that if condition (ii) in Theorem 4.3 for the SSB with $T = {}_iP$, for some $i \in S$, is satisfied, then $P(\theta)$ is stable, i.e., has a unique stationary distribution.*

4.3 Explicit Perturbation Bounds for the Small Network (Finite State Space)

In this section we explicitly compute the bounds from Theorem 4.3, i.e., CNB, SSB, DB and SEB(K) (for $K = 0, 1$), for the Small Network from Example 4.2. The following convex combination is considered

$$P(\theta) = (1-\theta)\underbrace{\begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}}_{=P^s} + \theta \underbrace{\begin{pmatrix} 1-\widetilde{p} & \widetilde{p} \\ \widetilde{q} & 1-\widetilde{q} \end{pmatrix}}_{:=\widetilde{P}^s}.$$

We are interested in perturbing $P(0)$ by choosing $\theta > 0$. Note that for the difference in Markov transition matrices it holds

$$P(\theta) - P(0) = \theta(\widetilde{P}^s - P^s) = \theta \begin{pmatrix} p-\widetilde{p} & \widetilde{p}-p \\ \widetilde{q}-q & q-\widetilde{q} \end{pmatrix}.$$

which gives

$$||P(\theta) - P(0)||_v = \theta(1+\alpha)\max\left\{|p-\widetilde{p}|, \frac{1}{\alpha}|q-\widetilde{q}|\right\}.$$

In the following the explicit perturbation bounds are presented for the v-norm. Using (4.13) in the calculation for CNB we get

$$||\pi^\top_{P(\theta)} - \pi^\top_{P^s}||_v \leq ||\pi^\top_{P(\theta)}||_v ||P(\theta) - P^s||_v ||D_{P^s}||_v.$$

It holds that (see also Example 4.2)

$$||\pi^\top_{P(\theta)}||_v \leq \alpha \qquad \text{and} \qquad ||D_{P^s}||_v = \frac{1+\alpha}{(p+q)^2}\max\left\{p, \frac{q}{\alpha}\right\}$$

so that we obtain for the CNB

$$||\pi^\top_{P(\theta)} - \pi^\top_{P^s}||_v \leq \theta\left(\frac{1+\alpha}{p+q}\right)^2 \max\left\{\alpha|p-\widetilde{p}|, |q-\widetilde{q}|\right\}\max\left\{p, \frac{q}{\alpha}\right\}.$$

In the general framework of CNB, where the perturbation bound has the form as given in (4.2), it holds that $\kappa = \frac{1+\alpha}{(p+q)^2}\max\{\alpha p, q\}$ for this example.

For the SSB we compute

$$||\pi^\top_{P^s}||_v = \frac{q+p\alpha}{p+q}.$$

Next, the individual terms in (4.24) have to be computed. Here, we make use of the taboo matrix bound as provided in Example 4.2 after removing the first column. SSB can only be provided for small perturbations, i.e., small values of θ. More specifically, provided that

$$\theta < \frac{1-\max\{\alpha p, 1-q\}}{\left(1+\frac{q+p\alpha}{p+q}\right)(1+\alpha)\max\left\{|p-\widetilde{p}|, \frac{1}{\alpha}|q-\widetilde{q}|\right\}},$$

the SSB bound for $||\pi^\top_{P(\theta)} - \pi^\top_{P^s}||_v$ is given by

$$\frac{\left(\frac{q+p\alpha}{p+q}\right)\left(1+\frac{q+p\alpha}{p+q}\right)\theta(1+\alpha)\max\left\{|p-\widetilde{p}|, \frac{1}{\alpha}|q-\widetilde{q}|\right\}}{1-\min\{\max\{\alpha p, 1-q\}, \max\{\alpha(1-p), q\}\} - \left(1+\frac{q+p\alpha}{p+q}\right)\theta(1+\alpha)\max\left\{|p-\widetilde{p}|, \frac{1}{\alpha}|q-\widetilde{q}|\right\}}.$$

For example, letting $\alpha = 1$, which is possible, see Lemma 4.1, yields the simplified expression

$$\frac{4\theta \max\{|p-\widetilde{p}|, |q-\widetilde{q}|\}}{1-\min\{\max\{p,1-q\},\max\{1-p,q\}\}-4\theta\max\{|p-\widetilde{p}|,|q-\widetilde{q}|\}}$$

for SSB. By inspection of above, it is obvious that SSB behaves poorly for p and q close to one or close to zero as in this case the norm of the taboo matrix approaches one.

Calculations show that DB leads to

$$||\pi^\top_{P(\theta)} - \pi^\top_{P^s}||_v \leq \frac{\theta|p\widetilde{q}-\widetilde{p}q|(1+\alpha)}{(p+q)\left(p+1-\theta(1+\alpha)\max\{|p-\widetilde{p}|, \frac{|q-\widetilde{q}|}{\alpha}\}\right)}$$

under the assumption that

$$\theta < \frac{p+1}{(1+\alpha)\max\{|p-\widetilde{p}|, \frac{|q-\widetilde{q}|}{\alpha}\}}.$$

For SEB(K) with $K = 0$ it holds

$$||\pi^\top_{P(\theta)} - \pi^\top_{P^s}||_v \leq \frac{\theta(1+\alpha)}{p+q}\max\{\alpha|p-\widetilde{p}|, |q-\widetilde{q}|\}$$

of which the construction is similar to CNB but with the difference that CNB requires an additional bounding on $||(P(\theta)-P^s)D_{P^s}||_v$ to obtain $||(P(\theta)-P^s)||_v||D_{P^s}||_v$, which stems from the fact that $||(P(\theta)-P^s)D_{P^s}||_v \leq ||(P(\theta)-P^s)||_v||D_{P^s}||_v$. More specifically, CNB is by factor

$$\frac{\text{CNB}}{\text{SEB}(0)} = \frac{1+\alpha}{p+q}\max\left\{p, \frac{q}{\alpha}\right\} \geq 1$$

larger than SEB(0). In case $\alpha = 1$ this factor is $2\max\{p,q\}/(p+q)$, which is greater than 1 for $p \neq q$. When α is chosen to be $>> 1$ this factor is likely to grow linearly in α. This illustrates that, although being more general, CNB loses on quality in contrast to SEB(0) since it does not utilize the contraction property of $(P(\theta) - P^s)D_{P^s}$.

After similar calculations it can be shown that SEB(K) with $K = 1$ results in

$$||\pi^\top_{P(\theta)} - \pi^\top_{P^s}||_v \leq \frac{\theta(1+\alpha)}{(p+q)^2}\left(|p\widetilde{q}-\widetilde{p}q| + \theta|p-\widetilde{p}+q-\widetilde{q}|\max\{\alpha|p-\widetilde{p}|, |q-\widetilde{q}|\}\right).$$

4.4 An Extensive Perturbation Analysis of a Queueing System

To illustrate the application of perturbation bounds in a setting where the deviation matrix is not available in a closed-form, we discuss in this section the $M/G/1$ queue with breakdowns. In addition, we consider the finite version of the queue, i.e., the $M/G/1/N$ queue with breakdowns and we illustrate SEB(K). The breakdown model will have the special feature that we perturb the system with no breakdowns by an unstable chain modelling a pure birth process.

The basic model of the $M/G/1$ queue with breakdowns is introduced in Section 4.4.1 and in Section 4.4.2 a discussion of the literature is provided. The perturbation bounds for both models are presented in Section 4.4.3 and Section 4.4.4, respectively.

4.4.1 The Basic Model

Consider a single server queue. Customers arrive at the queue according to a Poisson-λ-arrival process. Service times are identically distributed with mean $1/\mu$ and we denote the service time distribution by $\mathcal{S}(x)$. Throughout this section we assume that $\lambda/\mu < 1$. At the beginning of each service, there is a probability θ that the server breaks down (and the customer is send back to the queue) and enters a repair state, the length of which is exponentially distributed with rate r and which is independent of everything else, and with probability $(1-\theta)$ the server is operational and serves the customer (if any, according to first come first served). The only points in time where a possible server breakdown can occur is right at the beginning of a service. This system is modelled by the jump chain embedded at service completions and completions of a repair, and it has state space $S = \{0, 1, \dots\}$. The transition probabilities from $i \in S$ to $j \in S$, denoted as $P_\theta(i,j)$, are given as follows:

For $i = 0$, the process jumps to $j \geq 0$ if a customer arrives and the server is operational and during the service of this customer there are j additional arrivals. This probability is given by

$$(1-\theta)\int_0^\infty e^{-\lambda x}\frac{(\lambda x)^j}{j!}\, d\mathcal{S}(x).$$

Alternatively, a customer arrives at the empty queue and the server breaks down at service initiation and during the repair time of the server there are $j-1$ additional arrivals, so that at the end of the repair time there are in total j customers at the server. This probability is given by

$$\theta\int_0^\infty e^{-\lambda x}\frac{(\lambda x)^{j-1}}{(j-1)!}\, re^{-rx}dx = \theta\frac{r}{\lambda+r}\left(\frac{\lambda}{\lambda+r}\right)^{j-1},$$

for $j \geq 1$ and zero for $j = 0$, where we make use of the convention that $0! = 1$. Combining these results, for $i = 0$, we arrive at

$$P_\theta(0,j) = (1-\theta)\int_0^\infty e^{-\lambda x}\frac{(\lambda x)^j}{j!}\, d\mathcal{S}(x) + \theta\frac{r}{\lambda+r}\left(\frac{\lambda}{\lambda+r}\right)^{j-1}\mathbb{1}_{\{j\geq 1\}}.$$

For $i \geq 1$, the process jumps to state $j \geq i-1$ if the server remains operationally, so that service of the subsequent customer in the queue may begin, and during the service of this customer there are $j-i+1 \geq 0$ additional arrivals. This probability is given by

$$(1-\theta)\int_0^\infty e^{-\lambda x}\frac{(\lambda x)^{j-i+1}}{(j-i+1)!}\, d\mathcal{S}(x).$$

Alternatively, there is a server breakdown and during the exponential repair time there are $j-i \geq 0$ arrivals from the outside. This probability is given by

$$\theta\frac{r}{\lambda+r}\left(\frac{\lambda}{\lambda+r}\right)^{j-i}.$$

Combining these results, we arrive at

$$P_\theta(i,j) = (1-\theta)\int_0^\infty e^{-\lambda x}\frac{(\lambda x)^{j-i+1}}{(j-i+1)!}\, d\mathcal{S}(x) + \theta\frac{r}{\lambda+r}\left(\frac{\lambda}{\lambda+r}\right)^{j-i}\mathbb{1}_{\{j\geq i\}},$$

for $1 \leq i$ and $i-1 \leq j$. All other entries of P_θ are set to zero.

Observe that for $\theta = 1$, P_1 models a pure birth process and the queue is not stable, whereas P_0 models a stable $M/G/1$ queue with no breakdowns. The transition matrix P_θ is given through the convex combination $\theta P_1 + (1-\theta)P_0$ of the two transition matrices.

4.4.2 Discussion of Literature

Since the pioneering work of Thiruvengadam [179] and Avi-Itzhak and Naor [17], there has been a considerable interest in the study of queues with server breakdowns, see for example [46, 137, 187] and references therein. However, the majority of results is expressed in terms of systems of equations the solution of which is rather challenging, or have solutions which are not easily interpretable in practice. For instance, Baccelli and Znati [23] provide the generating function of the number of customers in the $M/G/1$ system with dependent breakdowns. Also, results are given in terms of the inverse of Laplace transforms, see, e.g., [23], which require numerical inversion for solving a given system. To overcome these difficulties, approximation methods are used where the complex (real) system is replaced by one which is "close" to it in some sense but which has a simpler structure (resp., components) and for which analytical results are available.

4.4.3 The Infinite Capacity $M/G/1$ Queue with Breakdowns (Denumerable State Space)

In this section the $M/G/1$ queue with breakdowns is considered. Note that SSB is the only bound applicable as the size of the state-space is infinite and the deviation matrix is not known in explicit form. As next we provide auxiliary results for obtaining the overall SSB. Recall that P_0 is the transition matrix of the embedded jump chain of an $M/G/1$ queue and we consider the taboo matrix $T = {}_0(P_0)$, i.e., we set the first column of P_0 to zero.

For the taboo matrix T it holds that

$$\begin{aligned}
\|T\|_v &= \sup_{i\geq 0} \frac{1}{\alpha^i} \sum_{j\geq 1} \alpha^j \left| \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^{j-i+1}}{(j-i+1)!} d\mathcal{S}(x) \right| \mathbb{1}_{\{j-i+1\geq 0\}} \\
&= \sup_{i\geq 0} \frac{1}{\alpha^i} \sum_{j\geq 1} \alpha^j \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^{j-i+1}}{(j-i+1)!} d\mathcal{S}(x) \mathbb{1}_{\{j\geq i-1\}} \\
&= \sup_{i\geq 0} \frac{1}{\alpha^i} \sum_{j\geq \max(i-1,1)} \alpha^j \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^j}{j!} d\mathcal{S}(x)
\end{aligned}$$

In particular for $i = 0, 1$,

$$\begin{aligned}
\sup_{0\leq i\leq 1} \frac{1}{\alpha^i} \sum_{j\geq \max(i-1,1)} \alpha^j \left| \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^j}{j!} d\mathcal{S}(x) \right| &= \sum_{j\geq 1} \alpha^j \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^j}{j!} d\mathcal{S}(x) \\
&= \sum_{j\geq 1} \int_0^\infty e^{-\lambda x} \frac{(\lambda \alpha x)^j}{j!} d\mathcal{S}(x)
\end{aligned}$$

$$
\begin{aligned}
&= \int_0^\infty e^{-\lambda x} \sum_{j\geq 1} \frac{(\lambda\alpha x)^j}{j!} d\mathcal{S}(x) \\
&= \int_0^\infty e^{-\lambda x}(e^{\lambda\alpha x} - 1) d\mathcal{S}(x) \\
&= \int_0^\infty e^{-\lambda(1-\alpha)x} d\mathcal{S}(x) - \int_0^\infty e^{-\lambda x} d\mathcal{S}(x),
\end{aligned}
$$

and for $i > 1$

$$
\begin{aligned}
&\sup_{i\geq 2} \frac{1}{\alpha^i} \sum_{j\geq \max(i-1,1)} \alpha^{j-1} \left| \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^j}{j!} d\mathcal{S}(x) \right| \\
&= \sup_{i\geq 2} \frac{1}{\alpha^i} \sum_{j\geq i-1} \alpha^{j-1} \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^j}{j!} d\mathcal{S}(x) \\
&= \frac{1}{\alpha^3} \int_0^\infty e^{-\lambda x} \sum_{j\geq 0} \frac{(\lambda\alpha x)^j}{j!} d\mathcal{S}(x) - \frac{1}{\alpha^3} \int_0^\infty e^{-\lambda x} d\mathcal{S}(x) \\
&= \frac{1}{\alpha^3} \left(\int_0^\infty e^{-\lambda(1-\alpha)x} d\mathcal{S}(x) - \int_0^\infty e^{-\lambda x} d\mathcal{S}(x) \right).
\end{aligned}
$$

Denoting by $\mathcal{S}^*(z)$ the Laplace-Stieltjes transform of $\mathcal{S}(x)$ and using the fact that $\alpha \geq 1$ we arrive at

$$
\|T\|_v = \|_0(P_0)\|_v \leq b_1(\alpha) := \mathcal{S}^*(\lambda(1-\alpha)) - \mathcal{S}^*(\lambda),
$$

provided that α is such that

$$
\mathcal{S}^*(\lambda(1-\alpha)) < \infty.
$$

Furthermore, using (4.10) one obtains

$$
||\pi_0^\top||_v \leq b_2(\alpha) := \frac{\sum_i \pi_0(i) P_0(i,0)}{1 - b_1(\alpha)} = \frac{\pi_0(0)}{1 - b_1(\alpha)}.
$$

We now turn to computing a bound for $||P_1 - P_0||_v$. For $i = 0$:

$$
\begin{aligned}
&\sum_{j\geq 0} \alpha^j |P_1(0,j) - P_0(0,j)| \\
&= \sum_{j\geq 0} \alpha^j \left| \frac{r}{r+\lambda} \left(\frac{\lambda}{\lambda+r}\right)^{j-1} \mathbb{1}_{\{j\geq 1\}} - \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^j}{j!} d\mathcal{S}(x) \right| \\
&= \int_0^\infty e^{-\lambda x} d\mathcal{S}(x) + \sum_{j\geq 0} \alpha^{j+1} \left| \frac{r}{r+\lambda} \left(\frac{\lambda}{\lambda+r}\right)^{j} - \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^{j+1}}{(j+1)!} d\mathcal{S}(x) \right|.
\end{aligned}
$$

For $i \geq 1$:

$$
\begin{aligned}
&\frac{1}{\alpha^i} \sum_{j\geq 0} \alpha^j |P_1(i,j) - P_0(i,j)| \\
&= \frac{1}{\alpha^i} \sum_{j\geq 0} \alpha^j \left| \frac{r}{r+\lambda} \left(\frac{\lambda}{\lambda+r}\right)^{j-i} \mathbb{1}_{\{j\geq i\}} - \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^{j-i+1}}{(j-i+1)!} d\mathcal{S}(x) \right| \mathbb{1}_{\{j-i+1\geq 0\}} \\
&= \frac{1}{\alpha} \int_0^\infty e^{-\lambda x} d\mathcal{S}(dx) + \frac{1}{\alpha^i} \sum_{j\geq i} \alpha^{j+1} \left| \frac{r}{r+\lambda} \left(\frac{\lambda}{\lambda+r}\right)^{j-i} - \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^{j-i+1}}{(j-i+1)!} d\mathcal{S}(x) \right|
\end{aligned}
$$

$$\leq \int_0^\infty e^{-\lambda x} d\mathcal{S}(dx) + \sum_{j\geq 0} \alpha^{j+1} \left| \frac{r}{r+\lambda}\left(\frac{\lambda}{\lambda+r}\right)^j - \int_0^\infty e^{-\lambda x}\frac{(\lambda x)^{j+1}}{(j+1)!} d\mathcal{S}(x)\right|.$$

Combining the above results we let

$$b_3(\alpha) \;:=\; \int_0^\infty e^{-\lambda x} d\mathcal{S}(x) + \sum_{j\geq 0} \alpha^{j+1} \left| \frac{r}{r+\lambda}\left(\frac{\lambda}{\lambda+r}\right)^j - \int_0^\infty e^{-\lambda x}\frac{(\lambda x)^{j+1}}{(j+1)!} d\mathcal{S}(x)\right|$$

and obtain

$$||P_1 - P_0||_v \leq b_3(\alpha).$$

Inserting the above bounds into (4.24) we obtain as SSB

$$||\pi_\theta^\top - \pi_0^\top||_v \;\leq\; b_2(\alpha)\frac{\theta(1+b_2(\alpha))b_3(\alpha)}{1-b_1(\alpha)-\theta(1+b_2(\alpha))b_3(\alpha)},$$

provided that

$$\theta < \frac{1-b_1(\alpha)}{(1+b_2(\alpha))b_3(\alpha)}$$

and $1 \leq \alpha \leq \min(1/\lambda, z_\lambda)$, where z_λ denotes the right point of the domain of the values for α such that $\mathcal{S}^*(\lambda(1-\alpha))$ is finite (the case $z_\lambda = \infty$ is not excluded).

Example 4.4. If the service times are exponentially distributed with rate μ it holds that

$$\mathcal{S}^*(\lambda(1-\alpha)) = \frac{\mu}{\mu+\lambda(1-\alpha)}$$

and $z_\lambda = \frac{\mu+\lambda}{\lambda} - \epsilon$, for some $\epsilon > 0$. The above bounds can now be explicitly computed:

$$b_1(\alpha) = \frac{\mu}{\mu+\lambda(1-\alpha)} - \frac{\mu}{\mu+\lambda} = \frac{\lambda\mu\alpha}{(\mu+\lambda)(\mu+\lambda(1-\alpha))},$$

$$b_2(\alpha) = \frac{1-\lambda/\mu}{1-b_1(\alpha)}.$$

and

$$b_3(\alpha) = \frac{\mu}{\lambda+\mu} + \alpha\sum_{j\geq 0}\alpha^j \left| \frac{r}{r+\lambda}\left(\frac{\lambda}{\lambda+r}\right)^j - \left(\frac{\lambda}{\lambda+\mu}\right)^{j+1}\frac{\mu}{\mu+\lambda}\right|.$$

Note that in case $\mu = r$, $b_3(\alpha)$ simplifies to

$$b_3(\alpha) = \frac{\mu}{\lambda+\mu} + \alpha\sum_{j\geq 0}\alpha^j\left(\frac{\mu}{\mu+\lambda}\right)^2\left(\frac{\lambda}{\lambda+\mu}\right)^j = \frac{\mu}{\lambda+\mu}\left(1+\frac{\alpha\mu}{\mu+\lambda-\alpha\lambda}\right)$$

provided that $\alpha < \frac{\lambda+\mu}{\lambda}$. ◁

In the following, we let $\lambda = 0.5$, $\mu = 1$, $r = 1$ and $f(s) = 0$ for $s \leq 2$ and $f(s) = 1$ for $s > 2$, i.e., we are interested in the probability of having more than 2 customers at the queue in stationary regime, i.e.,

$$||f||_v = \frac{1}{\alpha^3}.$$

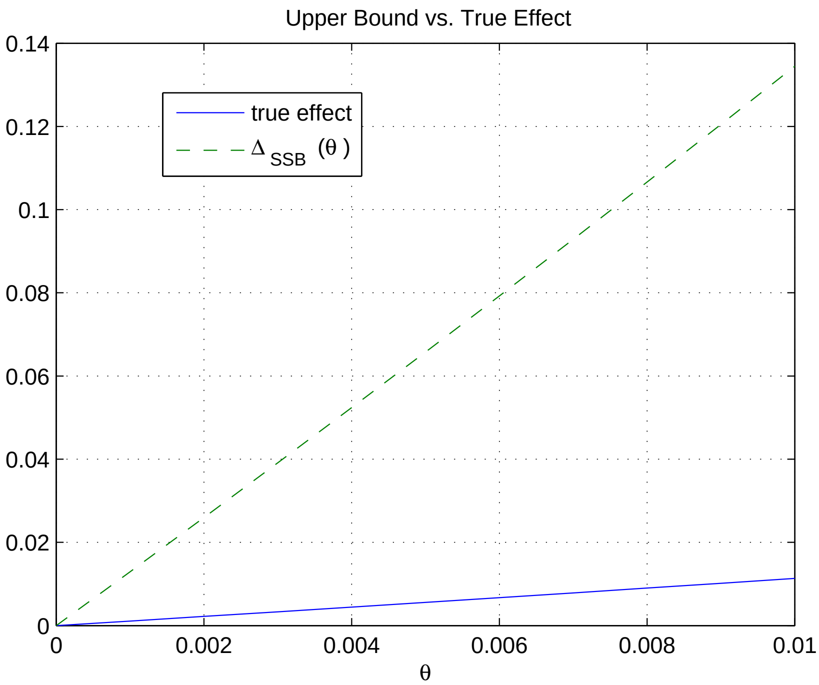

Fig. 4.3: The true change in probability of more than 2 customers in the system vs. the strong stability bound.

For ease of computation we assume that the service times are exponentially distributed.

We are now able to apply the bound provided in Lemma 4.2 to $|\pi_\theta f - \pi_0 f|$ in combination with the above SSB, where we let θ vary from 0 to 0.01, see Figure 4.3. The minimization with respect to α in (4.11) has been solved numerically.

As can be seen from Figure 4.3, SSB provides qualitative insight rather than numerically satisfying approximations.

Recall that $T = {}_0(P_0)$ and, by Remark 4.10, applicability of SSB implies stability of the system with breakdowns. SSB can thus be used as a means of establishing a lower bound for the domain of stability of the queue with breakdowns. More precisely, by Example 4.4, for $\mu = r = 1$ condition

$$||T||_v \leq b_1(\alpha) < 1$$

implies

$$\alpha \leq \frac{(\mu + \lambda)^2}{(2\mu + \lambda)\lambda},$$

which yields for the numerical setting of our example

$$\alpha \leq \frac{9}{5}.$$

In accordance with Corollary 4.1, a lower bound for the region of stability of $P(\theta)$ is

$$\frac{1 - ||T||_v}{||P_1 - P_0||_v} \geq \max_{1 \leq \alpha \leq 9/5} \frac{(\mu + \lambda)^2 - \lambda(2\mu + \lambda)\alpha}{\mu(\mu + \lambda + \alpha(\mu - \lambda))},$$

where we used the bounds provided in Example 4.4. For the numerical values of the example we obtain

$$\max_{1 \leq \alpha \leq 9/5} \frac{9 - 5\alpha}{6 + 2\alpha} = \frac{1}{2},$$

where the maximum is attained at $\alpha = 1$. Hence, the system remains stable for a breakdown probability up to $\approx 1/2$.

In the following section, we will show that the series expansion bound yields numerically better bounds. This comes, however, at the price of restricting the analysis to a finite version of the model.

4.4.4 The $M/G/1/N$ Queue with Breakdowns (Finite State Space)

In this section a $M/G/1/N$ queue is considered with finite size N (where N is not too large). We assume that a customer finding a full queue upon arrival is lost. In this case the state space is $S = \{0, 1, \ldots, N\}$, and D_θ (short for D_{P_θ}) as well as π_θ (short for π_{P_θ}) can be easily computed numerically so that SEB can be applied. In this case, SEB can be used for numerical computations. We illustrate the series expansion bound with some numerical examples. We choose $N = 50$ as the maximum number of jobs in the system. Like in the previous section, we let $\lambda = 0.5$, $\mu = 1$, $r = 1$, and assume that service times are exponentially distributed.

Remark 4.11. *Note that for large N the mean queue length of the finite system is (almost) identical to that of the infinite one. In this case one could use the strong stability bounds for approximate performance evaluation rather than computing SEB explicitly.*

We compute SEB for the v-norm with $\alpha = 1$. We have to check the condition put forward in (iv) of Theorem 4.3 numerically. For our numerical setting we obtain $||(P_1 - P_0)D_0||_v = 8$, which implies $\theta||(P_1 - P_0)D_0||_v < 1$ for $0 \leq \theta \leq \theta_0 < 1/8$. In the following we choose $\theta_0 = 0.1$.

In Figure 4.4 we plot the absolute relative error of SEB(K) for $K = 1, 2$ and 3, for the probability of having more than 2 customers in the systems. More specifically, we bound $|\pi_\theta^\top f - \pi_0^\top f|$ for $\theta \in [0, \theta_0]$, with $\theta_0 = 0.1$, using SEB(K), where $f(s) = 1$ if $s > 2$ and zero otherwise. It thus holds that $||f||_v = 1$. In line with Lemma 4.2, we obtain the bound

$$|\pi_\theta^\top f - \pi_0^\top f| \leq \Delta_{\mathrm{SEB}(K)}(P(\theta), P_0).$$

The absolute relative error as plotted in Figure 4.4 is given by

$$\frac{\left| \Delta_{\mathrm{SEB}(K)}(P(\theta), P_0) - |\pi_\theta^\top f - \pi_0^\top f| \right|}{|\pi_\theta^\top f - \pi_0^\top f|},$$

for $K = 1, 2, 3$ and $\theta \in [0, 0.1]$.

4.4.5 Discussion of Results

In this section we established numerical approximations for the single server queue with breakdowns. SSB has the advantage of providing bounds for infinite queues,

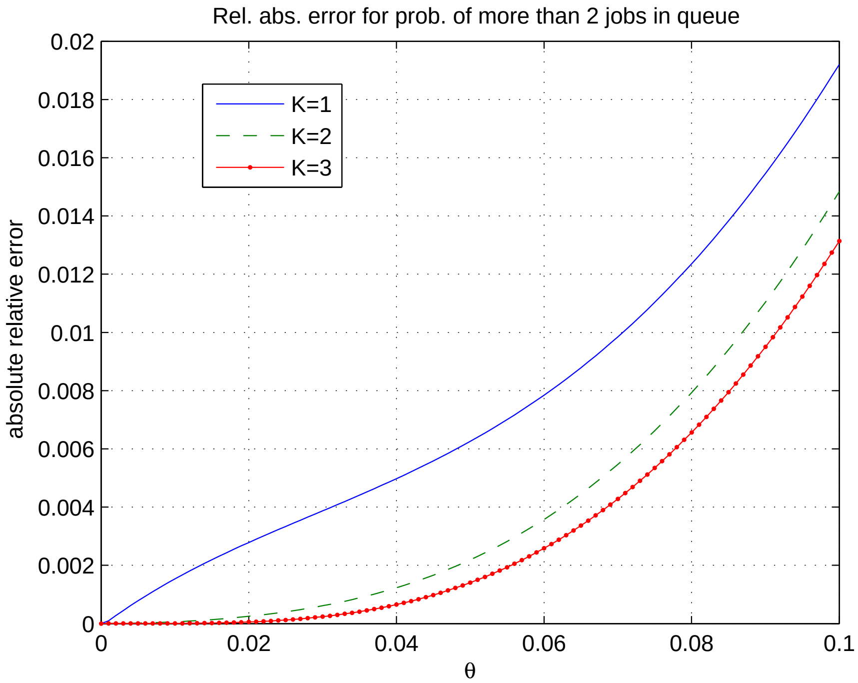

Fig. 4.4: The relative absolute error for approximating the $|\pi_\theta^\top f - \pi_0^\top f|$ with SEB(K) with $K = 1, 2$ and 3.

but unfortunately, the numerical quality of the bounds is rather poor. In light of Theorem 4.3, this comes as no surprise. SEB proved to be numerically very efficient for the model but required that a finite queue is studied. There is, however, an interesting link between the two approaches as the techniques developed for SSB lend themselves to establish lower bounds of convergence for series expansions.

4.5 Conclusion

Perturbation bounds for Markov chains have been intensively studied in the literature. Condition number bounds are attractive as they provide uniform perturbation bounds. Unfortunately, due to their simple structure they fail to capture the true non-linear dependence of the stationary distribution on the Markov transition matrix as was shown via examples and the introduced notion of relative error of perturbation bounds. A new family of bounds based on a series expansion approach overcomes this drawback. Furthermore, a realistic example from queueing theory illustrated the potential use of perturbation bounds in stability analysis.