Tilburg University

On Markov Chains with Uncertain Data

Blanc, J.P.C.; den Hertog, D.

Publication date:
2008

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):

Blanc, J. P. C., & den Hertog, D. (2008). *On Markov Chains with Uncertain Data*. (CentER Discussion Paper; Vol. 2008-50). Operations research.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

No. 2008–50

ON MARKOV CHAINS WITH UNCERTAIN DATA

By Hans Blanc, Dick den Hertog

May 2008

ISSN 0924-7815

On Markov chains with uncertain data

Hans Blanc*

Dick den Hertog†

CentER, Tilburg University
P.O. Box 90153, 5000 LE Tilburg, The Netherlands

May 23, 2008

Abstract

In this paper, a general method is described to determine uncertainty intervals for performance measures of Markov chains given an uncertainty region for the parameters of the Markov chains. We investigate the effects of uncertainties in the transition probabilities on the limiting distributions, on the state probabilities after n steps, on mean sojourn times in transient states, and on absorption probabilities for absorbing states. We show that the uncertainty effects can be calculated by solving linear programming problems in the case of interval uncertainty for the transition probabilities, and by second order cone optimization in the case of ellipsoidal uncertainty. Many examples are given, especially Markovian queueing examples, to illustrate the theory.

Jel code: C61

Keywords: Markov chain; Interval uncertainty; Ellipsoidal uncertainty; Linear Programming; Second Order Cone Optimization.

1 Introduction

In practice the transition probabilities for a finite-state Markov chain are often estimated. This raises the following relevant question for practice: What are the effects of uncertainties in transition probabilities on e.g. the steady state distribution and on all kinds of performance measures based on the steady state distribution? This paper investigates the effects of the uncertainties in the transition probabilities on the limiting distributions, on the state probabilities after n steps, on mean sojourn times in transient states, and on absorption probabilities for absorbing states. We show for example that finding the minimal or maximal value for the limiting probability for a certain state can be done by solving LPs, when the uncertainty on the transition probabilities is ‘column-wise’ and characterized by linear constraints. When the uncertainty region is defined by ‘column-wise’ ellipsoidal constraints, the resulting problem is a Second Order Cone Optimization (SOCO) problem, which can be solved efficiently nowadays. Column-wise refers in the Markov chain context to probabilities or rates of leaving a certain state. If the uncertainty regions are not column-wise, the structure of the problem is shown to be more complex.

We now discuss related research. Markov chains with non-exact values for the transition probabilities have been considered in the Markov decision theory by e.g. [11] and [14]. Sensitivity of stationary distributions to perturbations in the transition matrices have been studied by e.g. [12], [19]. Stanford [16] proved some properties of the set of possible limiting distributions for a finite-state non-stationary Markov chain with uncertain transition probabilities, characterized by upper and lower bounds for each transition probability. Note that Stanford uses the terminology fuzzy transition probabilities, but that in fact he is just using intervals for the transition probabilities. Markov chains with fuzzy parameters, and more particular in Markov decision processes and queues with the arrival rate and service rate being fuzzy numbers, have been investigated by e.g. [2], [7], [8], [9], [10].

The closest research to our paper, however, is that of Smith [15]. He described the set of stationary probability vectors arising when the transition probabilities of an N -state Markov chain lie in an interval. This set is

*email: blanc@uvt.nl

†email: D.DenHertog@uvt.nl

shown to lie in a convex cone bounded by at most $3N$ hyperplanes. The difference with our paper is threefold. First, we give the exact characterization of the set of feasible stationary probability vectors. Note, that Walton and Poore [17] incorrectly interpreted the $3N$ hyperplanes of Smith as if they give the exact characterization. Second, our results also hold for more general uncertainty regions for the transition probabilities, e.g. for ellipsoidal uncertainty regions. Third, we also apply our methods to other interesting measures often used in Markov chains, as state probabilities after n steps, on mean sojourn times in transient states, and on absorption probabilities for absorbing states.

To determine the effect of uncertain data in Markov chains, one has to solve nonlinear optimization problems. In this paper we will use transformations to transform these nonconvex problems into tractable optimization problems (linear programming and second-order cone programming). Dantzig [3] already used such transformations for certain kinds of linear programming problems. He showed that linear programming problems with uncertain data described by an interval for each element, can be reformulated into an other pure linear programming problem. This was rediscovered by Serafini [13], who does not refer to Dantzig's work. This result is also described in [4], again without reference to Dantzig's results. Note that in the case of Markov chains we never have pure interval uncertainty, since we always have the constraints that the sum of all transition probabilities out of a state equals one. Hence, even for the case of interval uncertainties on the transition probabilities we cannot use the results given above.

In Section 2.1, uncertainty intervals are considered for the stationary state probabilities of finite-state Markov chains with discrete or continuous time parameter and rather general, column-wise uncertainty constraints on the transposed of the transition matrix or generator, respectively. Section 2.2 deals with general properties of the optimization problems. In Section 3.1, uncertainty intervals are considered for the stationary state probabilities of finite-state, discrete-time Markov chains with box constraints. In Section 3.2, uncertainty intervals are considered for the mean sojourn times in transient states and for the absorption probabilities of finite-state Markov chains with absorbing states and with box constraints on the transition probabilities. Section 3.3 deals with uncertainty intervals for the state probabilities after n steps of finite-state Markov chains. In Section 4, uncertainty intervals are considered for Markov chains with ellipsoidal constraints for the uncertainty region. In Section 5, we apply the theory to Markovian queueing models.

Throughout this paper, $\mathbf{1}$ denotes a column vector of ones and $\mathbf{0}$ denotes a column vector of zeros, both of a dimension that will be clear from the context. Further, I denotes the identity matrix and O denotes the zero matrix, also of a dimension that will be clear from the context.

2 Markov chains with general constraints

This section is concerned with the characterization of the region of a performance measure related to a finite-state Markov chain of which the transition probabilities lie in some uncertainty region.

2.1 Problem formulation

The stationary distribution of a discrete-time Markov chain with finite state space $\mathcal{S}$ (with N states) is determined as the solution of a set of linear equations of the form, cf. e.g. Wolff [18, Sect. 3-4],

$$\boldsymbol{\pi}^T = \boldsymbol{\pi}^T P, \quad \boldsymbol{\pi}^T \mathbf{1} = 1, \quad \text{with } P\mathbf{1} = \mathbf{1}, \quad P \geq O, \quad (1)$$

here, $\boldsymbol{\pi}^T$ stands for the row vector of stationary state probabilities, P is the matrix of one-step transition probabilities, with entry $p_{ij} \geq 0$ being the probability that the process jumps from state i to state j , $i, j \in \mathcal{S}$. Alternatively, this set of equations can be written in the transposed form

$$\boldsymbol{\pi} = P^T \boldsymbol{\pi}, \quad \boldsymbol{\pi}^T \mathbf{1} = 1, \quad \text{with } \mathbf{1}^T P^T = \mathbf{1}^T, \quad P^T \geq O. \quad (2)$$

The stationary distribution of a continuous-time Markov chain with finite state space $\mathcal{S}$ (N states) is determined as the solution of a set of linear equations of the form, cf. e.g. Wolff [18, Sect. 4-3],

$$\boldsymbol{\nu}^T G = \mathbf{0}^T, \quad \boldsymbol{\nu}^T \mathbf{1} = 1, \quad \text{with } G\mathbf{1} = \mathbf{0}, \quad g_{ij} \geq 0, \quad i \neq j, \quad i, j \in \mathcal{S}, \quad (3)$$

here, $\boldsymbol{\nu}^T$ stands for the row vector of stationary state probabilities, the matrix G is the generator of the transition function, with entry g_{ij} being the transition rate from state i to state j , $i \neq j$, $i, j \in \mathcal{S}$, and with entry g_{ii} being minus the transition rate out of state i , $i \in \mathcal{S}$. Alternatively, this set of equations can be written in the transposed form

$$G^T \boldsymbol{\nu} = \mathbf{0}, \quad \boldsymbol{\nu}^T \mathbf{1} = 1, \quad \text{with } \mathbf{1}^T G^T = \mathbf{0}^T, \quad g_{ij} \geq 0, \quad i \neq j, \quad i, j \in \mathcal{S}. \quad (4)$$

In practice, it will happen that the entries of the matrices P or G are based on measurements or expert opinions, and hence, there will be some uncertainty about these entries. Consequently, there is uncertainty in the stationary distribution and we would like to analyze this. This means that we may want to solve problems of the form

$$\min_{\mathbf{x}, A} \{ \mathbf{c}^T \mathbf{x} \mid A \in \mathcal{U}, \mathbf{1}^T A = \mathbf{0}^T, A\mathbf{x} = \mathbf{0}, \mathbf{1}^T \mathbf{x} = 1, \mathbf{x} \geq \mathbf{0} \}; \quad (5)$$

here, $A = I - P^T$, cf. (2), or $A = -G^T$, cf. (4), $\mathcal{U}$ is the uncertainty region of the matrix A , and $\mathbf{c}$ is a vector of objective coefficients. In most applications we will take for $\mathbf{c}$ the unit vector $\mathbf{e}_i$ (consisting of all zeros except a 1 as i th coordinate) or its opposite to determine the smallest or largest possible value of the steady-state probability $x_i = \pi_i$ or $x_i = \nu_i$ for some i , $i \in \mathcal{S}$. But $\mathbf{c}$ could also be some other vector, e.g. to compute the smallest or largest possible value of a moment of the stationary distribution $\mathbf{x} = \boldsymbol{\pi}$ or $\mathbf{x} = \boldsymbol{\nu}$. Due to the logical constraints $\mathbf{1}^T P^T = \mathbf{1}^T$ in (2) and $\mathbf{1}^T G^T = \mathbf{0}^T$ in (4), the column elements of the matrix A are connected by the conditions $\mathbf{1}^T A = \mathbf{0}^T$. The off-diagonal elements $a_{ij} = -p_{ji}$, cf. (2), or $a_{ij} = -g_{ji}$, cf. (4), should be nonpositive:

$$a_{ij} \leq 0, \quad i \neq j, \quad i, j \in \mathcal{S}, \quad (6)$$

and, in the case of a discrete time Markov chain, the diagonal elements $a_{jj} = 1 - p_{jj}$ should be bounded by 1:

$$a_{jj} \leq 1, \quad j \in \mathcal{S}. \quad (7)$$

Throughout, we will suppose that $\mathcal{U}$ is a general column-wise uncertainty region defined by inequalities of the form

$$f_{jk}(\mathbf{a}_j) \leq 1, \quad k = 1, \dots, M_j, \quad j \in \mathcal{S}; \quad (8)$$

here, $\mathbf{a}_j$ denotes the j th column vector of the matrix A , M_j is the number of constraints on the elements of $\mathbf{a}_j$, and $f_{jk}(\mathbf{a}_j)$, $k = 1, \dots, M_j$, is a positively homogeneous function of degree 1, possibly after a shift by a common vector $\mathbf{a}_j^0$ which may represent a vector with nominal values for the j th column vector of the matrix A . Hence, for all positive scalars γ the functions $f_{jk}(\mathbf{a}_j)$ satisfy the relation

$$f_{jk}(\gamma \mathbf{a}_j + \mathbf{a}_j^0) = \gamma f_{jk}(\mathbf{a}_j + \mathbf{a}_j^0), \quad k = 1, \dots, M_j, \quad j \in \mathcal{S}. \quad (9)$$

Clearly, the logical constraints (6) and (7) can be written in this form, with, respectively,

$$f_{ji}(\mathbf{a}_j) = a_{ij} + 1, \quad i \neq j, \quad i, j = 1, \dots, N; \quad f_{jj}(\mathbf{a}_j) = a_{jj}, \quad j = 1, \dots, N. \quad (10)$$

Next, we discuss several general properties of optimization problem (5). The following proposition gives sufficient conditions for convexity of the feasible region with respect to $\mathbf{x}$.

Proposition 1 *Suppose that $\mathcal{U}$ is a convex and column-wise uncertainty region defined by (8), then the feasible region of optimization problem (5) is convex with respect to $\mathbf{x}$ and with respect to A .*

Proof We first prove convexity with respect $\mathbf{x}$. Let us take two feasible solutions for optimization problem (5): $(\mathbf{x}^1, A^1)$ and $(\mathbf{x}^2, A^2)$. We show that each convex combination $\mathbf{x}^3 = \lambda \mathbf{x}^1 + (1 - \lambda) \mathbf{x}^2$, $0 \leq \lambda \leq 1$, is also feasible. We define

$$\mu_j \doteq \frac{\lambda x_j^1}{\lambda x_j^1 + (1 - \lambda) x_j^2}, \quad j \in \mathcal{S},$$

and column vectors of a matrix A^3

$$\mathbf{a}_j^3 \doteq \mu_j \mathbf{a}_j^1 + (1 - \mu_j) \mathbf{a}_j^2, \quad j \in \mathcal{S}.$$

It is easy to see that $\mathbf{1}^T \mathbf{x}^3 = 1$, $\mathbf{x}^3 \geq \mathbf{0}$, $\mathbf{1}^T A^3 = \mathbf{0}^T$ and

$$A^3 \mathbf{x}^3 = \sum_{j \in \mathcal{S}} \mathbf{a}_j^3 x_j^3 = \lambda A^1 \mathbf{x}^1 + (1 - \lambda) A^2 \mathbf{x}^2 = \mathbf{0}.$$

This proves that $\mathbf{x}^3$ is feasible for problem (5), and hence proves the first part of the theorem.

We now prove convexity with respect to A . Again, let us take two feasible solutions for optimization problem (5): $(\mathbf{x}^1, A^1)$ and $(\mathbf{x}^2, A^2)$. We have to show that each convex combination $A^3 = \lambda A^1 + (1 - \lambda) A^2$, $0 \leq \lambda \leq 1$,

is also feasible. This easily follows from the fact that also A^3 corresponds to a Markov chain (with a limiting distribution). $\square$

Note, that the above proposition only ensures convexity with respect to $\mathbf{x}$ and not with respect to A . Indeed, it is easy to construct an example for which $\mathcal{U}$ is convex and column-wise, but for which the feasible region of (5) is not convex in A .

It is easy to construct simple examples that show that if the uncertainty region $\mathcal{U}$ is not convex, then the feasible region of (5) is not convex with respect to $\mathbf{x}$. E.g., the feasible region of the following example

$$\{x \mid ax = 1, a \in \{1, 2\}, x \geq 0\}$$

consists of two points $\{\frac{1}{2}, 1\}$, which is not convex.

If $\mathcal{U}$ is not column-wise then the feasible region is not necessarily convex. It is, e.g., easy to verify that the feasible region of the following example

$$\{(x_1, x_2) \mid ax_1 + bx_2 = 0, a + b = 1, a \geq 0, b \geq 0, x_1 \geq 0, x_2 \geq 0\}$$

consists of the union of the nonnegative x_1 -axis and the nonnegative x_2 -axis, which is not convex.

The previous example can be easily extended to show that (5) may be an NP-hard problem when $\mathcal{U}$ is not column-wise:

$$\{(x_1, x_2) \mid x_1 + x_2 = 1, ax_1 + bx_2 = 0, a + b = 1, a \geq 0, b \geq 0, x_1 \geq 0, x_2 \geq 0\}.$$

The feasible set for this problem is $\{(0, 1), (1, 0)\}$. Hence, in this way (0, 1) variables can be modeled.

2.2 Transformation of the optimization problems

With the assumptions on $\mathcal{U}$ made in the previous subsection, cf. (8), optimization problem (5) can be formulated as a nonlinear program:

$$\min_{\mathbf{x}, A} \{ \mathbf{c}^T \mathbf{x} \mid \mathbf{1}^T A = \mathbf{0}^T, A\mathbf{x} = \mathbf{0}, \mathbf{1}^T \mathbf{x} = 1, \mathbf{x} \geq \mathbf{0}, f_{jk}(\mathbf{a}_j) \leq 1, k = 1, \dots, M_j, j \in \mathcal{S} \}. \quad (11)$$

Beside the possibly nonlinear constraints (8) the constraint $A\mathbf{x} = \mathbf{0}$ is nonlinear and nonconvex. In many cases, this problem can be transformed into another problem with better structural properties. To this end, introduce a new matrix Ξ with column vectors

$$\xi_j \doteq x_j(\mathbf{a}_j - \mathbf{a}_j^0), \quad j \in \mathcal{S}. \quad (12)$$

Then, the nonlinear constraints $A\mathbf{x} = \mathbf{0}$ become linear constraints $\Xi \mathbf{1} + A_0 \mathbf{x} = \mathbf{0}$, with A_0 the matrix with columns the vectors $\mathbf{a}_j^0$, $j \in \mathcal{S}$. The logical constraints $\mathbf{1}^T A = \mathbf{0}^T$ remain linear constraints $\mathbf{1}^T \Xi + D \mathbf{x} = \mathbf{0}^T$, with $D \doteq \text{diag}(\mathbf{1}^T A_0)$ the diagonal matrix with the vector $\mathbf{1}^T A_0$ on its diagonal. The interval uncertainty constraints (8) translate into the inequalities $f_{jk}(\xi_j + \mathbf{a}_j^0) \leq x_j$, $k = 1, \dots, M_j$, $j \in \mathcal{S}$. In this way, the optimization problem (11) is transformed into the following optimization problem

$$\min_{\mathbf{x}, \Xi} \{ \mathbf{c}^T \mathbf{x} \mid \mathbf{1}^T \Xi + \mathbf{x}^T D = \mathbf{0}^T, \Xi \mathbf{1} + A_0 \mathbf{x} = \mathbf{0}, \mathbf{1}^T \mathbf{x} = 1, \mathbf{x} \geq \mathbf{0}, f_{jk}(\xi_j + \mathbf{a}_j^0) \leq x_j, k = 1, \dots, M_j, j \in \mathcal{S} \}. \quad (13)$$

Note that if the matrix A_0 satisfies $\mathbf{1}^T A_0 = \mathbf{0}^T$, then $D = O$ and the first constraint set reduces to $\mathbf{1}^T \Xi = \mathbf{0}^T$. Sections 3.1 and 4 consider special cases of constraints of the form (8), (9).

Above we have shown that the nonconvex problem (5) can be reformulated as (13) in case all f_{jk} are positively homogeneous functions of degree 1. Hence, if all f_{jk} are convex, we have reformulated the original nonconvex problem as a convex problem. If all f_{jk} are linear, then problem (13) is a linear programming problem. A special case, namely box constraints, is treated in Section 3. If all f_{jk} are ellipsoidal then problem (13) is a second-order cone optimization (SOCO) problem. This case is treated in Section 4. There are many other examples for f_{jk} satisfying the positively homogeneous of degree 1 property. A convex example is

$$\left[\sum_{i \in \mathcal{S}} \beta_{ij} (a_{ij} - a_{ij}^0)^{\alpha_j} \right]^{1/\alpha_j},$$

in which $\alpha_j \geq 1$, and $\beta_{ij} \geq 0$, and a nonconvex example is

$$\prod_{i \in \mathcal{S}} (a_{ij} - a_{ij}^0)^{\alpha_i},$$

in which $\sum_{i \in \mathcal{S}} \alpha_i = 1$. Finally, we note that the same reformulation can be used when f_{jk} is positively homogeneous of degree 0 (e.g. the previous example, but now with $\sum_{i \in \mathcal{S}} \alpha_i = 0$).

Remark 1 Since we are dealing with stationary distributions, we should, in principle, take care that the uncertainty region $\mathcal{U}$ in (5) is such that there exists a unique, strictly positive solution $\mathbf{x}$ to $A\mathbf{x} = \mathbf{0}$ for each $A \in \mathcal{U}$ with $\mathbf{1}^T A = \mathbf{0}^T$. However, this may be difficult to verify, especially if the functions $f_{jk}(\mathbf{a}_j)$ in (8) are nonlinear. Fortunately, it turns out that the optimization problems (13) are robust against violations of the stationarity assumption. If it happens that a transition matrix is feasible which corresponds to a Markov chain in which one or more states are transient, then (13) yields a lower bound of 0 for the probabilities of these states — which should then be interpreted as limiting probabilities. Only the inverse transformation to (12), $\mathbf{a}_j = \xi_j/x_j + \mathbf{a}_j^0$, $j \in \mathcal{S}$, which allows one to find a transition matrix corresponding to an extreme value of a stationary state probability, is not well defined for states with $x_j = 0$. Even if the uncertainty region $\mathcal{U}$ is such that (only) transition matrices are feasible that correspond to a Markov chain with two (or more) disjunct closed subsets of states, (13) yields a lower bound of 0 for all state probabilities (the limiting probability if the chain starts in another closed subset) and an upper bound that corresponds to the maximum attainable value of the limiting probability if the chain starts in e.g. this state. See also Example 1 and Remark 10. $\square$

Remark 2 If one is interested in the matrix A that minimizes (5), the inverse transformation to (12), $\mathbf{a}_j = \xi_j/x_j + \mathbf{a}_j^0$, $j \in \mathcal{S}$, can be used. However, this transformation is not well defined for states with $x_j = 0$. In that case one could solve (5) for A (for a given solution $\mathbf{x}$):

$$\{A \mid A \in \mathcal{U}, \mathbf{1}^T A = \mathbf{0}^T, A\mathbf{x} = \mathbf{0}\}. \quad (14)$$

If all f_{jk} are linear, this feasibility problem is linear in A , and if all f_{jk} are ellipsoidal, this feasibility problem is quadratic. Note, that in many cases there are multiple solutions for A that minimizes (5). Therefore, one might solve the above given feasibility problem with an objective function added, e.g., to minimize the deviation to the nominal value A^0 . $\square$

Remark 3 The shift vectors $\mathbf{a}_j^0$ in (9) do not need to represent nominal values for the vectors $\mathbf{a}_j$, $j \in \mathcal{S}$. They do not need to satisfy $\mathbf{1}^T \mathbf{a}_j^0 = 0$, and their elements do not need to satisfy inequalities like (6) and (7), while the optimization programs (13) still have feasible solutions. See further Remark 10. $\square$

3 Markov chains with box constraints

This section is concerned with the characterization of the uncertainty region of a performance measure related to a finite-state discrete-time Markov chain of which the transition probabilities lie in some uncertainty region determined by linear boundaries. Section 3.1 deals with stationary distributions of recurrent Markov chains, Section 3.2 with mean sojourn times and absorption probabilities in Markov chains with absorbing states, and Section 3.3 with transient distributions.

3.1 Stationary Markov chains with box constraints

Suppose that $\mathcal{U}$ is an interval uncertainty region, that is, for each element of the matrix A there exists an uncertainty interval independent of the values of the other elements of A :

$$a_{ij}^- \leq a_{ij} \leq a_{ij}^+, \quad i, j \in \mathcal{S}. \quad (15)$$

The lower bound and the upper bound are assumed to satisfy the following logical constraints due to the interpretation as transition probabilities. The off-diagonal elements $a_{ij} = -p_{ji}$, cf. (2), should be nonpositive and cannot be smaller than -1 :

$$-1 \leq a_{ij}^- \leq a_{ij}^+ \leq 0, \quad i \neq j, i, j \in \mathcal{S}, \quad (16)$$

and, avoiding absorbing states with $p_{ii} = 1$, the diagonal elements $a_{ii} = 1 - p_{ii}$ should be positive and cannot be larger than 1:

$$0 < a_{ii}^- \leq a_{ii}^+ \leq 1, \quad i \in \mathcal{S}. \quad (17)$$

We restrict the bounds of the uncertainty intervals (15) to those satisfying (16) and (17). Then, the logical constraints (6) and (7) are not explicitly needed in the problem formulation. Further note that since $\mathbf{1}^T A = \mathbf{0}^T$, it is necessary for a feasible uncertainty region that

$$\sum_{i \in \mathcal{S}} a_{ij}^- \leq 0, \quad \sum_{i \in \mathcal{S}} a_{ij}^+ \geq 0, \quad j \in \mathcal{S}. \quad (18)$$

The problem of finding the uncertainty intervals for the stationary probabilities can be written in the general form (11) by defining

$$f_{j,2i-1}(\mathbf{a}_j) = 2(a_{ij}^0 - a_{ij})/(a_{ij}^+ - a_{ij}^-), \quad f_{j,2i}(\mathbf{a}_j) = 2(a_{ij} - a_{ij}^0)/(a_{ij}^+ - a_{ij}^-), \quad i, j \in \mathcal{S}; \quad (19)$$

here, $a_{ij}^0 \doteq \frac{1}{2}(a_{ij}^- + a_{ij}^+)$, $i, j \in \mathcal{S}$. However, in the present case it is simpler to let A_- be the matrix of lower bounds a_{ij}^- , $i, j \in \mathcal{S}$, and A_+ be the matrix of upper bounds a_{ij}^+ , $i, j \in \mathcal{S}$. Then, we can describe the interval uncertainty region $\mathcal{U}$ by $A_- \leq A \leq A_+$, cf. (15), and the optimization problem (11) by

$$\min_{\mathbf{x}, A} \{\mathbf{c}^T \mathbf{x} \mid A_- \leq A \leq A_+, \mathbf{1}^T A = \mathbf{0}^T, A\mathbf{x} = \mathbf{0}, \mathbf{1}^T \mathbf{x} = 1, \mathbf{x} \geq \mathbf{0}\}. \quad (20)$$

The above optimization problem in the variables $\mathbf{x}$ and A is nonlinear due to the product $A\mathbf{x}$. To transform this problem into a linear one, assume that the Markov chain is ergodic for all matrices A with $A_- \leq A \leq A_+$ so that $\mathbf{x} > \mathbf{0}$, and introduce a new matrix Ξ with column vectors (avoiding the shift in (12) and the introduction of a vector $\mathbf{a}_j^0$):

$$\xi_j \doteq x_j \mathbf{a}_j, \quad j \in \mathcal{S}. \quad (21)$$

Then, the nonlinear constraints $A\mathbf{x} = \mathbf{0}$ become linear constraints $\Xi \mathbf{1} = \mathbf{0}$. The logical constraints $\mathbf{1}^T A = \mathbf{0}^T$ simply become $\mathbf{1}^T \Xi = \mathbf{0}^T$ since the column vector $\mathbf{a}_j$ of the matrix A is multiplied by the scalar x_j , $j \in \mathcal{S}$. The interval uncertainty constraints (15) translate into the inequalities $x_j a_{ij}^- \leq \xi_{ij} \leq x_j a_{ij}^+$, $i, j \in \mathcal{S}$. In this way, the nonlinear optimization problem (20) is transformed into the LP problem

$$\min_{\mathbf{x}, \Xi} \{\mathbf{c}^T \mathbf{x} \mid x_j a_{ij}^- \leq \xi_{ij} \leq x_j a_{ij}^+, i, j \in \mathcal{S}, \mathbf{1}^T \Xi = \mathbf{0}^T, \Xi \mathbf{1} = \mathbf{0}, \mathbf{1}^T \mathbf{x} = 1, \mathbf{x} \geq \mathbf{0}\}. \quad (22)$$

This LP problem consists of $N^2 + N$ variables, with $2N$ equality constraints (one of the constraints $\mathbf{1}^T \Xi = \mathbf{0}^T$, $\Xi \mathbf{1} = \mathbf{0}$ is redundant) and N^2 inequality constraints.

Remark 4 Since the constraints $\mathbf{1}^T \Xi = \mathbf{0}^T$ imply $\xi_{kj} = -\sum_{i \neq k, i \in \mathcal{S}} \xi_{ij}$, the following inequalities should hold in (22):

$$-x_j \sum_{i \neq k, i \in \mathcal{S}} a_{ij}^+ \leq \xi_{kj} \leq -x_j \sum_{i \neq k, i \in \mathcal{S}} a_{ij}^-, \quad k, j \in \mathcal{S}. \quad (23)$$

These constraints are redundant if

$$a_{kj}^- \geq -\sum_{i \neq k, i \in \mathcal{S}} a_{ij}^+, \quad a_{kj}^+ \leq -\sum_{i \neq k, i \in \mathcal{S}} a_{ij}^-, \quad k, j \in \mathcal{S}. \quad (24)$$

Moreover, if the conditions (24) hold, then also conditions (18) are satisfied. It should be noted that the inequalities (24) read in terms of the one-step transition probabilities of a discrete-time Markov chain

$$p_{jk}^- \geq 1 - \sum_{i \neq k, i \in \mathcal{S}} p_{ji}^+, \quad p_{jk}^+ \leq 1 - \sum_{i \neq k, i \in \mathcal{S}} p_{ji}^-, \quad k, j \in \mathcal{S}. \quad (25)$$

If these inequalities do not hold, the uncertainty intervals for some individual elements of the transition matrix are larger than the total uncertainty in the remaining part of the row of such an element in the transition matrix. Hence, it is no loss of generality if the uncertainty intervals are first reduced to satisfy (25) before the matrices A_- and A_+ are constructed and the LP problem (22) is solved. $\square$

Remark 5 Summing the inequalities for ξ_{ij} in the LP problem (22) over j , for each i , $i \in \mathcal{S}$, yields with the constraints $\Xi \mathbf{1} = \mathbf{0}$ ($A\mathbf{x} = \mathbf{0}$) the constraints $A_-\mathbf{x} \leq \mathbf{0} \leq A_+\mathbf{x}$. Hence, each $\mathbf{x}$ for which there exists a matrix A such that $(\mathbf{x}, A)$ is a feasible solution of (20), is a feasible solution of the reduced LP problem

$$\min_{\mathbf{x}} \{\mathbf{c}^T \mathbf{x} \mid A_-\mathbf{x} \leq \mathbf{0}, A_+\mathbf{x} \geq \mathbf{0}, \mathbf{1}^T \mathbf{x} = 1, \mathbf{x} \geq \mathbf{0}\}. \quad (26)$$

This LP problem with only N variables captures in many cases the solution of the full LP problem (22), for instance for $N = 2$ (see Example 1 below) and $N = 3$ for which this property can be shown with more effort; in other cases, it *may* yield too wide uncertainty intervals because it includes vectors $\mathbf{x}$ for which there is no corresponding matrix A that satisfies the omitted constraint $\mathbf{1}^T A = \mathbf{0}^T$, cf. (20). See Example 2 for a case where the solution to this LP problem does not yield the correct uncertainty interval. $\square$

Example 1 Consider a Markov chain with two states and an interval uncertainty region for the transition probabilities. For $N = 2$, we show that the reduced LP problems (26) yield the same uncertainty intervals as the full LP problems (22) provided that the conditions (25) hold, and derive expressions for the uncertainty intervals. The conditions (25) imply in this case that it must hold that

$$p_{11}^+ + p_{12}^- = 1, \quad p_{11}^- + p_{12}^+ = 1, \quad p_{21}^+ + p_{22}^- = 1, \quad p_{21}^- + p_{22}^+ = 1.$$

This means that if $p_{i,3-i}^- = \eta_i - \delta_i$ and $p_{i,3-i}^+ = \eta_i + \delta_i$, then $p_{ii}^- = 1 - \eta_i - \delta_i$ and $p_{ii}^+ = 1 - \eta_i + \delta_i$, $i = 1, 2$, so that the matrices of lower bounds and of upper bounds become

$$A_- = \begin{pmatrix} \eta_1 - \delta_1 & -\eta_2 - \delta_2 \\ -\eta_1 - \delta_1 & \eta_2 - \delta_2 \end{pmatrix}, \quad A_+ = \begin{pmatrix} \eta_1 + \delta_1 & -\eta_2 + \delta_2 \\ -\eta_1 + \delta_1 & \eta_2 + \delta_2 \end{pmatrix}.$$

The feasible region of the LP problem (26) is determined by the inequalities

$$-\delta_1 x_1 - \delta_2 x_2 \leq \eta_1 x_1 - \eta_2 x_2 \leq \delta_1 x_1 + \delta_2 x_2,$$

together with the normalization $x_1 + x_2 = 1$. This implies that the uncertainty intervals for the stationary probabilities are contained in the intervals

$$\frac{\eta_2 - \delta_2}{\eta_1 + \delta_1 + \eta_2 - \delta_2} \leq x_1 \leq \frac{\eta_2 + \delta_2}{\eta_1 - \delta_1 + \eta_2 + \delta_2}, \quad \frac{\eta_1 - \delta_1}{\eta_1 - \delta_1 + \eta_2 + \delta_2} \leq x_2 \leq \frac{\eta_1 + \delta_1}{\eta_1 + \delta_1 + \eta_2 - \delta_2}.$$

The constraints $\Xi \mathbf{1} = \mathbf{0}$ and $\mathbf{1}^T \Xi = \mathbf{0}^T$ of (22) imply that $\xi_{11} = -\xi_{12} = -\xi_{21} = \xi_{22} \doteq \xi$, while the inequalities in (22) reduce to

$$(\eta_1 - \delta_1)x_1 \leq \xi \leq (\eta_1 + \delta_1)x_1, \quad (\eta_2 - \delta_2)x_2 \leq \xi \leq (\eta_2 + \delta_2)x_2.$$

With the normalization $x_1 + x_2 = 1$ this leads to the same intervals for x_1 and x_2 as above. Hence, the uncertainty intervals for the stationary probabilities are

$$\frac{\eta_2 - \delta_2}{\eta_1 + \delta_1 + \eta_2 - \delta_2} \leq \pi_1 \leq \frac{\eta_2 + \delta_2}{\eta_1 - \delta_1 + \eta_2 + \delta_2}, \quad \frac{\eta_1 - \delta_1}{\eta_1 - \delta_1 + \eta_2 + \delta_2} \leq \pi_2 \leq \frac{\eta_1 + \delta_1}{\eta_1 + \delta_1 + \eta_2 - \delta_2}.$$

These bounds are bilinear functions of the absolute uncertainties δ_1 and δ_2 . Only if $\delta_1 = \delta_2 \doteq \delta$, these bounds become linear in δ . This result holds provided that $0 < \eta_i - \delta_i \leq \eta_i + \delta_i \leq 1$, $i = 1, 2$, cf. (16), (17). However, if e.g. $\delta_1 = \eta_1$, corresponding to $p_{12}^- = 0$ and $p_{11}^+ = 1$, state 1 can be absorbing and state 2 transient, which is well reflected by the above inequalities yielding an upper bound of 1 for π_1 and a lower bound of 0 for π_2 , cf. Remark 1. Even if $\delta_1 = \eta_1 = \delta_2 = \eta_2 = 0$, which means that only $P = I$ is feasible and the Markov chain consists of two absorbing states, the solutions of (22), yielding $0 \leq \pi_1, \pi_2 \leq 1$, make sense as possible limiting probabilities (depending on the initial distribution). $\square$

Example 2 Next, we discuss a case in which the LP problem (26) yields too wide uncertainty intervals. Consider a Markov chain with four states, with the following matrices of lower and upper bounds on the transition probabilities:

$$P_- = \begin{pmatrix} 0.00 & 0.50 & 0.30 & 0.05 \\ 0.05 & 0.00 & 0.50 & 0.30 \\ 0.30 & 0.05 & 0.00 & 0.50 \\ 0.50 & 0.30 & 0.05 & 0.00 \end{pmatrix}, \quad P_+ = \begin{pmatrix} 0.05 & 0.60 & 0.40 & 0.10 \\ 0.10 & 0.05 & 0.60 & 0.40 \\ 0.40 & 0.10 & 0.05 & 0.60 \\ 0.60 & 0.40 & 0.10 & 0.05 \end{pmatrix}. \quad (27)$$

These matrices are cyclically symmetric, and their entries satisfy the conditions (25). However, the solutions to the reduced LP problems (26) yield the intervals

$$0.2160 \leq x_k \leq 0.2840, \quad k = 1, 2, 3, 4,$$

while the solutions to the full size LP problems (22) yield the uncertainty intervals

$$0.2165 \leq \pi_k \leq 0.2834, \quad k = 1, 2, 3, 4.$$

The vectors $\mathbf{x}$, satisfying $A_- \mathbf{x} = (I - P_+^T) \mathbf{x} \leq \mathbf{0}$, $A_+ \mathbf{x} = (I - P_-^T) \mathbf{x} \geq \mathbf{0}$ and $\mathbf{x}^T \mathbf{1} = 1$, that yield the lower and upper bound for x_1 are, respectively:

$$\mathbf{x}_{\text{low}}^T = (0.2160 \ 0.2653 \ 0.2833 \ 0.2354), \quad \mathbf{x}_{\text{up}}^T = (0.2840 \ 0.2325 \ 0.2180 \ 0.2655).$$

These vectors turn out not to correspond to a transition matrix P in the prescribed range. The transition matrices P_{low} and P_{up} that yield the lower and upper bound for π_1 , respectively, are

$$P_{\text{low}} = \begin{pmatrix} 0.00 & 0.60 & 0.35 & 0.05 \\ 0.05 & 0.05 & 0.60 & 0.30 \\ 0.30 & 0.10 & 0.05 & 0.55 \\ 0.50 & 0.40 & 0.10 & 0.00 \end{pmatrix}; \quad P_{\text{up}} = \begin{pmatrix} 0.05 & 0.50 & 0.35 & 0.10 \\ 0.10 & 0.00 & 0.50 & 0.40 \\ 0.40 & 0.05 & 0.00 & 0.55 \\ 0.60 & 0.30 & 0.05 & 0.05 \end{pmatrix}.$$

The first of these matrices minimizes the transition probabilities into state 1 and tries to delay the return to state 1 as much as possible, the latter maximizes the transition probabilities out of state 1 and tries to hasten the return to state 1 as much as possible. The corresponding stationary distributions are, respectively:

$$\pi_{\text{low}}^T = (0.2165 \ 0.2675 \ 0.2742 \ 0.2419), \quad \pi_{\text{up}}^T = (0.2834 \ 0.2306 \ 0.2274 \ 0.2586).$$

We have found cases with wider box constraints on the transition probabilities for which the difference between the solutions of (26) and (22) is larger. $\square$

Remark 6 One could a priori use that the row sums of the transition matrix P are equal to 1, and write the set of balance equations (2) (for instance for $N = 3$) as

$$(I - P^T) \boldsymbol{\pi} = \mathbf{0}, \quad \boldsymbol{\pi}^T \mathbf{1} = 1, \quad \text{with } I - P^T = \begin{pmatrix} p_{12} + p_{13} & -p_{21} & -p_{31} \\ -p_{12} & p_{21} + p_{23} & -p_{32} \\ -p_{13} & -p_{23} & p_{31} + p_{32} \end{pmatrix}.$$

However, this reduction of the number of parameters also reduces the degrees of freedom in choosing the uncertainty intervals for $\mathcal{U}$. For instance, if the off-diagonal elements all have the same absolute uncertainty of $\pm\delta$, the diagonal elements inherit in this form a large absolute uncertainty of $\pm(N-1)\delta$. Further, one could argue that one of the balance equations is redundant, and replace it by the normalization condition. This yields a set of equations of the form

$$\hat{A} \boldsymbol{\pi} = \mathbf{e}_3, \quad \text{with } \hat{A} = \begin{pmatrix} p_{12} + p_{13} & -p_{21} & -p_{31} \\ -p_{12} & p_{21} + p_{23} & -p_{32} \\ 1 & 1 & 1 \end{pmatrix}, \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

This problem has six degrees of freedom, but with restrictions

$$p_{12} + p_{13} \leq 1, \quad p_{21} + p_{23} \leq 1, \quad p_{31} + p_{32} \leq 1, \quad p_{ij} \geq 0, \quad i, j = 1, 2, 3.$$

A problem in determining the matrices $\hat{A}_-$ and $\hat{A}_+$ in this form is that some parameters are present in different rows with opposite signs. $\square$

It is further possible to add column-wise linear constraints of the form, cf. (8), (9),

$$f_{j,2N+k}(\mathbf{a}_j) = \sum_{i=1}^N h_{ij}^{(k)}(a_{ij} - a_{ij}^0) \leq 1, \quad k = 1, \dots, M_j, \quad j \in \mathcal{S}, \quad (28)$$

to the interval constraints (15) to obtain a linear uncertainty region $\mathcal{U}$; here, M_j denotes the number of constraints on the elements of $\mathbf{a}_j$, $h_{ij}^{(k)}$ are constants, and $\mathbf{a}_j^0$ is the common nominal vector in (19). Under the transformation (21) these constraints remain linear:

$$\sum_{i=1}^N h_{ij}^{(k)} \xi_{ij} \leq x_j \left(1 + \sum_{i=1}^N h_{ij}^{(k)} a_{ij}^0 \right), \quad k = 1, \dots, M_j, \quad j \in \mathcal{S}. \quad (29)$$

Hence, the optimization problems (22) extended with constraints of this form remain LP problems.

Example 3 A constraint of the form $a_{21} \leq a_{31}$, which corresponds to the inequality $\pi_{12} \geq \pi_{13}$, is transformed by (21) into $\xi_{21} \leq \xi_{31}$. Formally, this linear inequality can be written in the form (29) by taking $h_{21} = -h_{31} = (a_{31}^0 - a_{21}^0)^{-1}$ for any vector $\mathbf{a}_1^0$. In general, the vectors $\mathbf{a}_j^0$, $j \in \mathcal{S}$, in (9) play no essential role as long as all constraints are linear. $\square$

3.2 Absorbing Markov chains with box constraints

In a Markov chain with transient states (the set $\mathcal{T}$) and absorbing states (the set $\mathcal{A}$), the structure of the transition matrix is, cf. e.g. Wolff [18, Sect. 3-7]:

$$P = \begin{pmatrix} Q & R \\ O & I \end{pmatrix}; \quad (30)$$

here, the matrix Q contains the transition probabilities between transient states, the matrix R contains the transition probabilities from transient states to absorbing states, the matrix O consists of zeros (no way back from absorbing to transient states), and matrix I denotes the identity matrix for the absorbing states. The row sums of the matrix P should be equal to 1 so that $Q\mathbf{1} + R\mathbf{1} = \mathbf{1}$. Performance measures of interest include the elements of the matrix $(I - Q)^{-1}$: the ij th element represents the expected sojourn time in state j before absorption given that the Markov chain starts in state i , $i, j \in \mathcal{T}$, and the elements of the matrix $(I - Q)^{-1}R$: the ij th element represents the probability of absorption in state j , $j \in \mathcal{A}$, given that the Markov chain starts in state i , $i \in \mathcal{T}$. These quantities are nonlinear functions of the elements of the matrix Q , and linear functions of the elements of the matrix R . However, we will show that the problems of determining the uncertainty regions of these quantities of interest given that the elements of the matrix Q (and R) lie in uncertainty intervals, can be reduced to LP problems.

First, consider the problem of determining the range of the expected sojourn time in state $k \in \mathcal{T}$, given some initial distribution $\mathbf{q}_0$ over the states in $\mathcal{T}$. For this purpose, we have to solve problems of the form

$$\min_{(Q, R) \in \mathcal{U}} \{ \mathbf{q}_0^T (I - Q)^{-1} \mathbf{c} \mid Q\mathbf{1} + R\mathbf{1} = \mathbf{1} \}; \quad (31)$$

here, e.g., $\mathbf{c} = \pm \mathbf{e}_k$ for some $k \in \mathcal{T}$ if the smallest or largest possible value of the expected sojourn time in state k has to be determined. Suppose that $\mathcal{U}$ is an interval uncertainty region of the following form:

$$0 \leq q_{ij}^- \leq q_{ij} \leq q_{ij}^+ \leq 1, \quad i, j \in \mathcal{T}; \quad (32)$$

$$0 \leq r_{ij}^- \leq r_{ij} \leq r_{ij}^+ \leq 1, \quad i \in \mathcal{T}, j \in \mathcal{A}. \quad (33)$$

These condition will be denoted by $Q_- \leq Q \leq Q_+$ and $R_- \leq R \leq R_+$. The row sums of the matrix P in (30) should be equal to 1. This implies that the lower and upper bounds should satisfy

$$Q_- \mathbf{1} + R_- \mathbf{1} \leq Q\mathbf{1} + R\mathbf{1} = \mathbf{1} \leq Q_+ \mathbf{1} + R_+ \mathbf{1}. \quad (34)$$

In the sequel, it will be assumed that these conditions are satisfied. For the optimization problems (31), the uncertainty intervals for the matrix R are only relevant as far as it must hold that $\mathbf{1} - R_+ \mathbf{1} \leq Q\mathbf{1} \leq \mathbf{1} - R_- \mathbf{1}$. To remove the nonlinearity of the objective function, introduce the vector $\mathbf{v}^T \doteq \mathbf{q}_0^T (I - Q)^{-1}$:

$$\min_{\mathbf{v}, Q} \{ \mathbf{v}^T \mathbf{c} \mid \mathbf{v}^T (I - Q) = \mathbf{q}_0^T, Q_- \leq Q \leq Q_+, \mathbf{1} - R_+ \mathbf{1} \leq Q\mathbf{1} \leq \mathbf{1} - R_- \mathbf{1} \},$$

or, by taking the transposed of the relations,

$$\min_{\mathbf{v}, Q} \{ \mathbf{c}^T \mathbf{v} \mid (I - Q^T) \mathbf{v} = \mathbf{q}_0, Q_-^T \leq Q^T \leq Q_+^T, \mathbf{1}^T R_-^T \leq \mathbf{1}^T (I - Q^T) \leq \mathbf{1}^T R_+^T \}. \quad (35)$$

In this form, there are constraints on the column sums of the matrix Q^T . The above optimization problems can be transformed into linear ones by introducing a new matrix Ξ with elements, cf. (21),

$$\xi_{ij} \doteq v_j(I_{ij} - q_{ji}), \quad i, j \in \mathcal{T}. \quad (36)$$

Then, the nonlinear constraints $(I - Q^T)\mathbf{v} = \mathbf{q}_0$ become linear constraints $\Xi\mathbf{1} = \mathbf{q}_0$. But more importantly, the conditions on the column sums remain linear:

$$v_i \sum_{l \in \mathcal{A}} r_{il}^- \leq \sum_{j \in \mathcal{T}} \xi_{ji} \leq v_i \sum_{l \in \mathcal{A}} r_{il}^+, \quad i \in \mathcal{T}. \quad (37)$$

Hence, the optimization problems (35) become LP problems:

$$\min_{\mathbf{v}, \Xi} \{\mathbf{c}^T \mathbf{v} \mid \Xi\mathbf{1} = \mathbf{q}_0, v_j(I_{ij} - q_{ji}^+) \leq \xi_{ij} \leq v_j(I_{ij} - q_{ji}^-), i, j \in \mathcal{T}, v_i \sum_{l \in \mathcal{A}} r_{il}^- \leq \sum_{j \in \mathcal{T}} \xi_{ji} \leq v_i \sum_{l \in \mathcal{A}} r_{il}^+, i \in \mathcal{T}\}. \quad (38)$$

These LP problems consist of $N^2 + N$ variables, with N equality constraints and $N^2 + N$ inequality constraints; here, N is the number of transient states.

Next, consider the problem of determining the range of the absorption probability in state $k \in \mathcal{A}$, given some initial distribution $\mathbf{q}_0$ over the states in $\mathcal{T}$. To this end, we have to solve problems of the form

$$\min_{(Q, R) \in \mathcal{U}} \{\mathbf{q}_0^T (I - Q)^{-1} R \mathbf{c} \mid Q\mathbf{1} + R\mathbf{1} = \mathbf{1}\}, \quad (39)$$

cf. (31), with $\mathbf{c} = \pm \mathbf{e}_k$ for some $k \in \mathcal{A}$ if the smallest or largest possible value of the absorption probability in state k has to be determined. Suppose again that $\mathcal{U}$ is an interval uncertainty region for which the inequalities (32) and (33) hold. To remove the nonlinearity of the objective function, use again the vector $\mathbf{v}^T \doteq \mathbf{q}_0^T (I - Q)^{-1}$ to obtain after transposition of the relations:

$$\min_{\mathbf{v}, Q, R} \{\mathbf{c}^T R^T \mathbf{v} \mid (I - Q^T)\mathbf{v} = \mathbf{q}_0, Q_-^T \leq Q^T \leq Q_+^T, R_-^T \leq R^T \leq R_+^T, \mathbf{1}^T Q^T + \mathbf{1}^T R^T = \mathbf{1}^T\}. \quad (40)$$

The above nonlinear optimization problems can be transformed into linear ones by introducing new matrices Ξ and Θ with elements, cf. (36),

$$\xi_{ij} \doteq v_j(I_{ij} - q_{ji}), \quad i, j \in \mathcal{T}, \quad \theta_{lj} \doteq v_j r_{jl}, \quad l \in \mathcal{A}, j \in \mathcal{T}. \quad (41)$$

Then, the nonlinear constraints $(I - Q^T)\mathbf{v} = \mathbf{q}_0$ become linear constraints $\Xi\mathbf{1} = \mathbf{q}_0$ and the nonlinear objective function $\mathbf{c}^T R^T \mathbf{v}$ becomes a linear objective function $\mathbf{c}^T \Theta \mathbf{1}$. Moreover, the conditions on the column sums remain linear: $\mathbf{1}^T \Xi - \mathbf{1}^T \Theta = \mathbf{0}^T$. Hence, the optimization problems (40) become LP problems:

$$\min_{\mathbf{v}, \Xi, \Theta} \{\mathbf{c}^T \Theta \mathbf{1} \mid \Xi\mathbf{1} = \mathbf{q}_0, \mathbf{1}^T \Xi = \mathbf{1}^T \Theta, v_j(I_{ij} - q_{ji}^+) \leq \xi_{ij} \leq v_j(I_{ij} - q_{ji}^-), i, j \in \mathcal{T}, v_j r_{jl}^- \leq \theta_{lj} \leq v_j r_{jl}^+, l \in \mathcal{A}, j \in \mathcal{T}\}. \quad (42)$$

These LP problems consist of $N^2 + N + NM$ variables, with $2N + M$ equality constraints and $N^2 + NM$ inequality constraints; here, N is the number of transient states and M is the number of absorbing states. In the examples below, we will use the elements of the vector $\mathbf{w} \doteq \Theta \mathbf{1}$ to denote the absorption probabilities.

Remark 7 Apart from the conditions (32), (33) and (34), the lower and upper bound matrices Q_- , R_- , Q_+ and R_+ should satisfy conditions to ensure that the matrix $I - Q$ has a positive inverse for all feasible Q (that is, to ensure that the states in $\mathcal{T}$ are really transient). A sufficient condition is that the matrix $I - Q_+$ has a positive inverse, but this is not necessary; see Example 5. Necessary conditions are that at least one element of the matrix R_- is positive (to enable absorption in $\mathcal{A}$) and that all diagonal elements of the matrix Q_+ are smaller than one (to prevent absorption in $\mathcal{T}$), but these are not sufficient (two states in $\mathcal{T}$ may still form a recurrent set). Another sufficient condition is that the matrix R_- has positive row sums (to enable direct absorption from each transient state), but also this is not necessary. If some states in $\mathcal{T}$ are not transient for a feasible Q , the upper bounds on their mean sojourn times are infinite, and the corresponding LP problems (38) are unbounded. $\square$

Remark 8 Since the constraints $\mathbf{1}^T \Xi = \mathbf{1}^T \Theta$ imply $\xi_{kj} = -\sum_{i \neq k, i \in \mathcal{T}} \xi_{ij} + \sum_{l \in \mathcal{A}} \theta_{lj}$, the following inequalities hold in the feasible region of the LP programs (42), cf. Remark 4:

$$v_j \left[\sum_{i \neq k, i \in \mathcal{T}} (q_{ji}^- - I_{ij}) + \sum_{l \in \mathcal{A}} r_{jl}^- \right] \leq \xi_{kj} \leq v_j \left[\sum_{i \neq k, i \in \mathcal{T}} (q_{ji}^+ - I_{ij}) + \sum_{l \in \mathcal{A}} r_{jl}^+ \right], \quad k, j \in \mathcal{T}. \quad (43)$$

These constraints are redundant if similar conditions as (25) are imposed:

$$q_{jk}^- \geq 1 - \sum_{i \neq k, i \in \mathcal{T}} q_{ji}^+ - \sum_{l \in \mathcal{A}} r_{jl}^+, \quad q_{jk}^+ \leq 1 - \sum_{i \neq k, i \in \mathcal{T}} q_{ji}^- - \sum_{l \in \mathcal{A}} r_{jl}^-, \quad j, k \in \mathcal{T}. \quad (44)$$

Similar conditions can be imposed on the upper and lower bounds of the elements of the matrix R :

$$r_{jk}^- \geq 1 - \sum_{i \in \mathcal{T}} q_{ji}^+ - \sum_{l \neq k, l \in \mathcal{A}} r_{jl}^+, \quad r_{jk}^+ \leq 1 - \sum_{i \in \mathcal{T}} q_{ji}^- - \sum_{l \neq k, l \in \mathcal{A}} r_{jl}^-, \quad j \in \mathcal{T}, \quad k \in \mathcal{A}. \quad (45)$$

The above conditions can be imposed without loss of generality, since the uncertainty interval of an individual element of the transition matrix P in (30) cannot be larger than the total uncertainty in the remaining row of such an element in this matrix. $\square$

Example 4 First, consider a case in which the matrix $I - Q_+$ has a positive inverse. Let there be three transient states in $\mathcal{T}$ and two absorbing states in $\mathcal{A}$, while the lower and upper bound matrices are

$$Q_- = \begin{pmatrix} 0.4 & 0.3 & 0.05 \\ 0 & 0.4 & 0.3 \\ 0 & 0 & 0.4 \end{pmatrix}, \quad R_- = \begin{pmatrix} 0 & 0 \\ 0.1 & 0.05 \\ 0.45 & 0.05 \end{pmatrix}, \quad Q_+ = \begin{pmatrix} 0.5 & 0.4 & 0.15 \\ 0 & 0.5 & 0.4 \\ 0 & 0 & 0.5 \end{pmatrix}, \quad R_+ = \begin{pmatrix} 0.1 & 0.1 \\ 0.15 & 0.1 \\ 0.5 & 0.1 \end{pmatrix}.$$

The elements of these matrices satisfy the conditions (44) and (45). For initial state $i = 1$ ($\mathbf{q}_0 = \mathbf{e}_1$), the solutions to the LP problems (38) and (42) read

$$1.667 \leq v_1 \leq 2, \quad 0.833 \leq v_2 \leq 1.6, \quad 0.647 \leq v_3 \leq 1.618, \quad 0.565 \leq w_1 \leq 0.897, \quad 0.103 \leq w_2 \leq 0.435.$$

For initial state $i = 2$ ($\mathbf{q}_0 = \mathbf{e}_2$), the solutions to the LP problems (38) and (42) read

$$v_1 = 0, \quad 1.667 \leq v_2 \leq 2, \quad 0.909 \leq v_3 \leq 1.454, \quad 0.691 \leq w_1 \leq 0.856, \quad 0.144 \leq w_2 \leq 0.309.$$

For initial state $i = 3$ ($\mathbf{q}_0 = \mathbf{e}_3$), the solutions to the LP problems (38) and (42) read

$$v_1 = 0, \quad v_2 = 0, \quad 1.667 \leq v_3 \leq 2, \quad 0.818 \leq w_1 \leq 0.909, \quad 0.091 \leq w_2 \leq 0.181.$$

Clearly, for upper diagonal matrices Q_+ the Markov chain is acyclic, and the matrix $I - Q_+$ has a positive inverse provided only that the diagonal elements $q_{ii}^+ < 1$, which is necessary for state i to be transient, $i \in \mathcal{T}$. In these cases, it generally holds that $(1 - q_{ii}^-)^{-1} \leq v_i \leq (1 - q_{ii}^+)^{-1}$, if $\mathbf{q}_0 = \mathbf{e}_i$, $i \in \mathcal{T}$. Note, however, that the uncertainty intervals for the mean sojourn time v_2 if $\mathbf{q}_0 = \mathbf{e}_1$ and that of v_3 if $\mathbf{q}_0 = \mathbf{e}_2$ are not the same although the interval constraints on q_{11}, q_{12}, q_{22} , are the same as those on q_{22}, q_{23}, q_{33} , respectively; this is due to the constraints on the other transition probabilities. Further note that with two absorbing states the uncertainty intervals for the absorption probabilities w_1 and w_2 are complementary. $\square$

Example 5 Next, consider cases in which the matrix $I - Q_+$ does not have a positive inverse. Let there be three transient states in $\mathcal{T}$ and three absorbing states in $\mathcal{A}$ while the lower and upper bound matrices are fully symmetric:

$$Q_- = \begin{pmatrix} \delta^- & \delta^- & \delta^- \\ \delta^- & \delta^- & \delta^- \\ \delta^- & \delta^- & \delta^- \end{pmatrix}, \quad R_- = \begin{pmatrix} \theta^- & \theta^- & \theta^- \\ \theta^- & \theta^- & \theta^- \\ \theta^- & \theta^- & \theta^- \end{pmatrix}, \quad Q_+ = \begin{pmatrix} \delta^+ & \delta^+ & \delta^+ \\ \delta^+ & \delta^+ & \delta^+ \\ \delta^+ & \delta^+ & \delta^+ \end{pmatrix}, \quad R_+ = \begin{pmatrix} \theta^+ & \theta^+ & \theta^+ \\ \theta^+ & \theta^+ & \theta^+ \\ \theta^+ & \theta^+ & \theta^+ \end{pmatrix}.$$

Clearly, we must have $\delta^- < \frac{1}{3}$, cf. (34), but δ^+ could be larger than $\frac{1}{3}$. However, the matrix $I - Q_+$ does not possess a positive inverse for $\delta^+ \geq \frac{1}{3}$. Still, the LP problems (38) and (42) have finite solutions provided that $\theta^- > 0$, cf. Remark 7. For instance, for $\delta^- = 0.3$, $\delta^+ = 0.35$, $\theta^- = 0.01$, $\theta^+ = 0.05$, the following uncertainty intervals follow by solving (38) and (42) with initial state $i = 1$ ($\mathbf{q}_0 = \mathbf{e}_1$):

$$4 \leq v_1 \leq 12.667, \quad 3 \leq v_2, v_3 \leq 11.667, \quad 0.1 \leq w_1, w_2, w_3 \leq 0.714.$$

The lower bounds correspond to the case $Q = Q_-$ (provided this is feasible, i.e., if $3\delta^- \geq 1 - 3\theta^+$) and for w_k to the case that the elements of the k th column of R all equal to θ^- , and follow by inverting the matrix $I - Q_-$ as:

$$v_1 \geq \frac{1 - 2\delta^-}{1 - 3\delta^-} = 4, \quad v_2, v_3 \geq \frac{\delta^-}{1 - 3\delta^-} = 3, \quad w_1, w_2, w_3 \geq \frac{\theta^-}{1 - 3\delta^-} = \frac{1}{10}.$$

The upper bounds correspond to asymmetric matrices Q with for v_j the elements of the j th column of Q all equal to $\delta^+ = 0.35$ and $R = R_-$, but the other elements at various not uniquely determined values, e.g., at $(1 - \delta^+ - 3\theta^-)/2 = 0.31$, $j \in \mathcal{T}$, and for w_k the elements of the k th column of R all equal to $\theta^+ = 0.05$ and all other elements of R equal to $\theta^- = 0.01$, while the elements of Q are not uniquely determined but could all be $(1 - \theta^+ - 2\theta^-)/3 = 0.31$. This implies

$$v_1 \leq \frac{\delta^+ + 3\theta^-}{3\theta^-} = \frac{38}{3}, \quad v_2, v_3 \leq \frac{\delta^+}{3\theta^-} = \frac{35}{3}, \quad w_1, w_2, w_3 \leq \frac{\theta^+}{\theta^+ + 2\theta^-} = \frac{5}{7}.$$

Note that $(1 - \delta^+ - 3\theta^-)/2 \geq \delta^-$ is equivalent to $\delta^+ \leq 1 - 2\delta^- - 3\theta^-$ and that $(1 - \theta^+ - 2\theta^-)/3 \geq \delta^-$ is equivalent to $\theta^+ \leq 1 - 2\theta^- - 3\delta^-$, while the latter inequalities should hold by (44) and (45). Hence, the above expressions for the upper bounds hold for all possible combinations of $\delta^-, \theta^-, \delta^+, \theta^+$. $\square$

3.3 Transient Markov chains with box constraints

The uncertainty of the transition probabilities of a Markov chain may also play a role in determining the short-run properties of a process. In this section, it will be assumed that there is a fixed interval uncertainty region $\mathcal{U}$ such that, for each n , $P_- \leq P_n \leq P_+$, with P_n the n th step transition matrix, $n = 1, 2, \dots$. The successive transition matrices P_n can be chosen independently within these fixed bounds. Further, there may be uncertainty intervals for the initial state probabilities: $\pi_-^{(0)} \leq \pi^{(0)} \leq \pi_+^{(0)}$. We only impose the following logical conditions on the bounds of the uncertainty intervals, and allow both recurrent and transient states in the finite space $\mathcal{S}$:

$$0 \leq p_{ij}^- \leq p_{ij}^+ \leq 1, \quad i, j \in \mathcal{S}; \quad \mathbf{0} \leq \pi_-^{(0)} \leq \pi_+^{(0)} \leq \mathbf{1}. \quad (46)$$

To determine the uncertainty intervals of the distribution of the Markov chain after n steps, we have to solve optimization problems of the form, for $n = 1, 2, \dots$,

$$\min_{P_n, \dots, P_1, \pi^{(0)}} \{ \mathbf{c}^T \pi^{(n)} \mid \pi^{(r)} = P_r^T \pi^{(r-1)}, P_- \leq P_r \leq P_+, P_r \mathbf{1} = \mathbf{1}, r = 1, \dots, n, \pi_-^{(0)} \leq \pi^{(0)} \leq \pi_+^{(0)}, \mathbf{1}^T \pi^{(0)} = 1 \}; \quad (47)$$

here, $\mathbf{c} = \pm \mathbf{e}_k$ for some $k \in \mathcal{S}$ if the uncertainty interval of the probability that the Markov chain is in state k after n steps has to be determined. The above nonlinear optimization problem can be transformed into an LP problem by introducing matrices Ξ_r , $r = 1, \dots, n$, with elements

$$\xi_{ij}^{(r)} \doteq \pi_j^{(r-1)} p_{ji}^{(r)}, \quad i, j \in \mathcal{S}, \quad r = 1, \dots, n. \quad (48)$$

Then, the nonlinear constraints $\pi^{(r)} = P_r^T \pi^{(r-1)}$ become linear constraints $\Xi_r \mathbf{1} = \pi^{(r)}$. Further, the logical constraints $P_r \mathbf{1} = \mathbf{1}$ remain linear constraints $\Xi_r^T \mathbf{1} = \pi^{(r-1)}$. Hence, the optimization problem (47) is equivalent to the LP problem

$$\min_{\Xi_n, \dots, \Xi_1, \pi^{(0)}} \{ \mathbf{c}^T \Xi_n \mathbf{1} \mid \Xi_r^T \mathbf{1} = \pi^{(r-1)} = \Xi_{r-1} \mathbf{1}, r = 2, \dots, n, \Xi_1^T \mathbf{1} = \pi^{(0)}, \mathbf{1}^T \pi^{(0)} = 1, \pi_j^{(r-1)} p_{ji}^- \leq \xi_{ij}^{(r)} \leq \pi_j^{(r-1)} p_{ji}^+, i, j \in \mathcal{S}, r = 1, \dots, n, \pi_-^{(0)} \leq \pi^{(0)} \leq \pi_+^{(0)} \}. \quad (49)$$

Observe that the intermediate distributions $\pi^{(r)}$ only act as auxiliary variables since they are completely determined by the matrices Ξ_r , $r = 1, \dots, n - 1$. Hence, the above LP problem essentially consists of $n \cdot N^2 + N$ variables, with $n \cdot N + 1$ equality constraints and $n \cdot N^2 + N$ inequality constraints; here, N is again the number of states in the space $\mathcal{S}$.

Remark 9 If the matrices P_- and P_+ in (47) depend on the step number r , the same transformation (48) leads to an LP problem as (49) but with p_{ji}^- and p_{ji}^+ in the inequalities for $\xi_{ij}^{(r)}$ depending on r . If the Markov chain is time-homogeneous, i.e. $P_r = P$, then transformation (48) can not be used since problem (47) does not have column-wise uncertainty. $\square$

Table 1: Lower and upper bounds for $\pi_j^{(n)}$, $j = 1, 2, 3, 4$, for uncertainty region (27), starting in state 1.

n	$\pi_1^{(n)} \geq$	$\pi_1^{(n)} \leq$	$\pi_2^{(n)} \geq$	$\pi_2^{(n)} \leq$	$\pi_3^{(n)} \geq$	$\pi_3^{(n)} \leq$	$\pi_4^{(n)} \geq$	$\pi_4^{(n)} \leq$
0	1.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1	0.000000	0.050000	0.500000	0.600000	0.300000	0.400000	0.050000	0.100000
2	0.145000	0.270000	0.032500	0.130000	0.255000	0.400000	0.317500	0.462500
3	0.258500	0.424000	0.196625	0.360000	0.086125	0.225750	0.155000	0.315625
4	0.129531	0.298519	0.199863	0.378375	0.205500	0.383381	0.131338	0.299275
5	0.156922	0.340828	0.133962	0.311263	0.164878	0.350713	0.192222	0.377386
6	0.175060	0.364744	0.166762	0.356439	0.144911	0.330195	0.151716	0.339717
7	0.149746	0.339784	0.162708	0.355094	0.166654	0.358925	0.153314	0.344148
8	0.157607	0.351153	0.151840	0.343991	0.156584	0.350145	0.162915	0.356627
9	0.159275	0.353777	0.158809	0.353292	0.154361	0.348120	0.154737	0.348724
10	0.154791	0.349252	0.156969	0.351954	0.158319	0.353258	0.155960	0.350755

Example 6 Consider a four-state Markov chain with lower and upper bounds on the transition matrices given by (27). Table 1 shows the lower and upper bounds for $\pi_j^{(n)}$, $j = 1, 2, 3, 4$, for the case that the Markov chain starts with probability 1 in state 1 for $n = 0$. Initially, there is clear non-monotone pattern in the bounds due to the initial state, but after a few steps the bounds for all state probabilities seem to tend to the same values. Table 2 shows the equal lower and upper bounds for $\pi_j^{(n)}$, $j = 1, 2, 3, 4$, for the case that the Markov chain starts with probability $\frac{1}{4}$ in any of the states for $n = 0$ (first pair of columns) and for the case that the initial distribution of the Markov chain has lower bound $\pi_-^{(0)} = 0.2125 \cdot \mathbf{1}$ and upper bound $\pi_+^{(0)} = 0.2875 \cdot \mathbf{1}$ (second pair of columns). Note that the latter bounds are equal to the uncertainty bounds of $\pi_j^{(1)}$ of the former case that the Markov chain starts with probability $\frac{1}{4}$ in any of the states for $n = 0$. This table shows — compare the upper bounds of $\pi_j^{(n)}$ of the former case with those of $\pi_j^{(n-1)}$ of the latter case, $n = 2, 3, \dots$ — that ignoring the shape of the feasible region and only working with component-wise lower and upper bounds may lead to too wide uncertainty intervals, as could be expected. For comparison, Table 2 also shows (last pair of columns) the equal lower and upper bounds for $\pi_j^{(n)}$, $j = 1, 2, 3, 4$, that are obtained by repeatedly solving a one-step problem, with initial distribution intervals the resulting uncertainty intervals of the previous step (“Step-by-step approach”), as opposed to solving (49) for $n > 1$ (“Multi-step approach”). Again, it turns out that ignoring the full shape of the feasible region may yield too large intervals. Finally, Table 2 shows the equal lower and upper bounds for $\pi_j^{(n)}$, $j = 1, 2, 3, 4$, for the case that the initial distribution of the Markov chain has lower bound $\pi_-^{(0)} = 0.0 \cdot \mathbf{1}$ and upper bound $\pi_+^{(0)} = 0.5 \cdot \mathbf{1}$ (third pair of columns). Observe that in contrast to the other cases where the uncertainty intervals tend to widen with n the uncertainty intervals tend to narrow with n in this example with very large uncertainty in the initial distribution. Note that the bounds are generally not monotone functions of n ; e.g. with initial bounds

Table 2: Lower and upper bounds for $\pi_j^{(n)}$, $j = 1, 2, 3, 4$, for uncertainty region (27), with symmetric start.

n	Multi-step approach						Step-by-step approach	
	$\pi_j^{(n)} \geq$	$\pi_j^{(n)} \leq$	$\pi_j^{(n)} \geq$	$\pi_j^{(n)} \leq$	$\pi_j^{(n)} \geq$	$\pi_j^{(n)} \leq$	$\pi_j^{(n)} \geq$	$\pi_j^{(n)} \leq$
0	0.250000	0.250000	0.212500	0.287500	0.000000	0.500000	0.250000	0.250000
1	0.212500	0.287500	0.184375	0.319375	0.025000	0.500000	0.212500	0.287500
2	0.187500	0.315625	0.172688	0.332969	0.088750	0.431250	0.184375	0.319375
3	0.174063	0.331344	0.166038	0.340442	0.120563	0.392000	0.163094	0.344031
4	0.166558	0.339888	0.161139	0.346036	0.130434	0.380878	0.146964	0.362795
5	0.161996	0.345047	0.159513	0.347898	0.145442	0.364050	0.134735	0.377032
6	0.159612	0.347774	0.157917	0.349697	0.148313	0.360592	0.125463	0.387829
7	0.158105	0.349488	0.157119	0.350616	0.151530	0.357009	0.118433	0.396015
8	0.157237	0.350479	0.156783	0.350991	0.154212	0.353890	0.113102	0.402222
9	0.156796	0.350978	0.156459	0.351362	0.154549	0.353535	0.109060	0.406928
10	0.156510	0.351305	0.156339	0.351500	0.155375	0.352606	0.105996	0.410496

$\pi_-^{(0)} = 0.15 \cdot \mathbf{1}$ and $\pi_+^{(0)} = 0.35 \cdot \mathbf{1}$ they oscillate. Finally we note that the sequence of transition matrices P_r , $r = 1, \dots, n$, that leads to a bound of some $\pi_j^{(n)}$, generally consists of different matrices. $\square$

4 Markov chains with ellipsoidal constraints

In this section we consider ellipsoidal uncertainty regions for recurrent, discrete-time Markov chains. Ellipsoidal uncertainty regions arise when confidence intervals are constructed based on observations, assuming normality. Suppose there are K_j ellipsoidal constraints on column j of the matrix A which are represented by

$$f_{j,N+k}(\mathbf{a}_j) = \sqrt{(\mathbf{a}_j - \mathbf{a}_j^0)^T F_j^{(k)} (\mathbf{a}_j - \mathbf{a}_j^0)}, \quad k = 1, \dots, K_j, \quad j \in \mathcal{S}; \quad (50)$$

here, $F_j^{(k)}$, $k = 1, \dots, K_j$, are positive semi-definite matrices, and $\mathbf{a}_j^0$ is a common vector for all constraints on $\mathbf{a}_j$, $j \in \mathcal{S}$. The constraints $f_{j,N+k}(\mathbf{a}_j) \leq 1$ are transformed by (12) into

$$f_{j,N+k}(\boldsymbol{\xi}_j + \mathbf{a}_j^0) = \sqrt{\boldsymbol{\xi}_j^T F_j^{(k)} \boldsymbol{\xi}_j} \leq x_j, \quad k = 1, \dots, K_j, \quad j \in \mathcal{S}. \quad (51)$$

Further, the constraints (6) and (7) become

$$\xi_{ij} + x_j a_{ij}^0 \leq 0, \quad i \neq j, \quad i, j \in \mathcal{S}, \quad \xi_{jj} + x_j (a_{jj}^0 - 1) \leq 0, \quad j \in \mathcal{S}. \quad (52)$$

Optimization problems with constraints of the form (51) are known as second order cone (SOCO) problems or conic quadratic programming (CQP) problems, and can be solved efficiently nowadays, cf. [1].

Example 7 Consider a three-state Markov chain with the following nominal matrix of transition probabilities and corresponding vectors of nominal values

$$P_0 = \begin{pmatrix} 0 & 0.25 & 0.75 \\ 0.75 & 0 & 0.25 \\ 0.75 & 0.25 & 0 \end{pmatrix}, \quad \mathbf{a}_1^0 = \begin{pmatrix} 1 \\ -0.25 \\ -0.75 \end{pmatrix}, \quad \mathbf{a}_2^0 = \begin{pmatrix} -0.75 \\ 1 \\ -0.25 \end{pmatrix}, \quad \mathbf{a}_3^0 = \begin{pmatrix} -0.75 \\ -0.25 \\ 1 \end{pmatrix},$$

and the following diagonal matrices for a single ellipsoidal constraint per column:

$$F_1 = \begin{pmatrix} 625 & 0 & 0 \\ 0 & 25 & 0 \\ 0 & 0 & 100 \end{pmatrix}, \quad F_2 = \begin{pmatrix} 100 & 0 & 0 \\ 0 & 625 & 0 \\ 0 & 0 & 25 \end{pmatrix}, \quad F_3 = \begin{pmatrix} 100 & 0 & 0 \\ 0 & 25 & 0 \\ 0 & 0 & 625 \end{pmatrix}.$$

The stationary distribution related to the nominal transition probabilities is $\boldsymbol{\pi}^T = (\frac{3}{7}, \frac{1}{5}, \frac{13}{35})$. Solution of the SOCO problems with $\mathbf{c} = \pm \mathbf{e}_k$, $k = 1, 2, 3$, yields the following uncertainty intervals for the stationary distribution

$$0.3977 \leq \pi_1 \leq 0.4663, \quad 0.1343 \leq \pi_2 \leq 0.2611, \quad 0.3197 \leq \pi_3 \leq 0.4250.$$

Matrices of transition probabilities corresponding to the lower and the upper bound of e.g. π_1 are, respectively,

$$P_{\text{low}} = \begin{pmatrix} 0.0000 & 0.2364 & 0.7636 \\ 0.6602 & 0.0035 & 0.3363 \\ 0.6602 & 0.3363 & 0.0035 \end{pmatrix}, \quad P_{\text{up}} = \begin{pmatrix} 0.0394 & 0.2185 & 0.7421 \\ 0.8394 & 0.0000 & 0.1606 \\ 0.8394 & 0.1606 & 0.0000 \end{pmatrix}.$$

The matrix P_{low} contains relatively large probabilities of cycling between states 2 and 3, while the matrix P_{up} contains relatively large probabilities of transitions to state 1. Note that the solution P_{low} is not unique: it can be shown that any values of p_{12}, p_{13} with $|p_{12} - 0.25| \leq \frac{1}{25}\sqrt{5}$ and $p_{13} = 1 - p_{12}$ yield the lower bound on π_1 due to the symmetry in the remaining part of P_{low} . The first row of the matrix P_{up} exhibits a typical aspect of a quadratic optimization problem: to maximize p_{11} it is optimal to have $p_{13} - 0.75 < 0$ although its coefficient in the constraint of 100 is larger than that of 25 of $p_{12} - 0.25$. $\square$

Remark 10 As indicated in Remark 3, the shift vectors $\mathbf{a}_j^0$ in (9) do not need to represent nominal values for the vectors $\mathbf{a}_j$, $j \in \mathcal{S}$. Consider the following vectors with corresponding matrix $P_0 = I - A_0^T$:

$$\mathbf{a}_1^0 = \begin{pmatrix} 0.05 \\ -0.05 \\ \alpha \end{pmatrix}, \quad \mathbf{a}_2^0 = \begin{pmatrix} -0.5 \\ 1 - \beta \\ -0.5 \end{pmatrix}, \quad \mathbf{a}_3^0 = \begin{pmatrix} \alpha \\ -0.05 \\ 0.05 \end{pmatrix}, \quad P_0 = \begin{pmatrix} 0.95 & 0.05 & \alpha \\ 0.50 & \beta & 0.50 \\ \alpha & 0.05 & 0.95 \end{pmatrix},$$

and the following diagonal matrices for a single ellipsoidal constraint per column: $F_j = 100I$, $j = 1, 2, 3$. Only if $\alpha = \beta = 0$, the matrix P_0 represents a matrix of transition probabilities. In this case, matrices with $p_{11} = 1$ and/or $p_{33} = 1$, which correspond to Markov chains with one or two absorbing states, are feasible. Still, the solutions to the corresponding SOCO problems yield sensible results in the sense of limiting distributions: $0 \leq \pi_1 \leq 1$, $0 \leq \pi_2 \leq 0.1162$, $0 \leq \pi_3 \leq 1$. For instance, if $p_{11} = 1$ and $p_{33} < 1$, state 1 is absorbing and states 2 and 3 are transient, so that $\pi_1 = 1$ and $\pi_2 = \pi_3 = 0$. Similarly, $p_{11} < 1$ and $p_{33} = 1$ leads to $\pi_1 = \pi_2 = 0$ and $\pi_3 = 1$. The feasible case that $p_{11} = 1$ and $p_{33} = 1$, in which (2) does not have a unique solution, does not play a role in the optimization problems. The upper bound $\pi_2 = 0.1162$ corresponds to a stationary Markov chain in which p_{22} and $p_{12} = p_{32}$ are maximal. Returning to the shift vectors $\mathbf{a}_j^0$, also for $\alpha \neq 0$ and/or $\beta \neq 0$, the corresponding SOCO problems have feasible solutions. For instance, taking $\beta = -0.1$ forces the vector $\mathbf{a}_2$ to be $(-0.5 \ 1 \ -0.5)^T$ (corresponding to $p_{22} = 0$, $p_{21} = p_{23} = 0.5$). And for $\alpha = \beta = -0.05$, matrices with $p_{11} = 1$ and/or $p_{33} = 1$ are still feasible, and the corresponding SOCO problems lead to the following bounds on the limiting distributions: $0 \leq \pi_1 \leq 1$, $0 \leq \pi_2 \leq 0.1043$, $0 \leq \pi_3 \leq 1$. Only for $|\alpha| > \frac{1}{20}\sqrt{3} \approx 0.0866$, matrices with $p_{11} = 1$ and/or $p_{33} = 1$ are no longer feasible. For $\alpha = -0.09$, $\beta = 0$, the following bounds on the limiting *and* stationary distributions are found: $0.1471 \leq \pi_1, \pi_3 \leq 0.8250$, $0.0188 \leq \pi_2 \leq 0.0809$. $\square$

5 Markovian queueing models

Markov process models for queueing systems often concern the description of the evolution over time of the number of customers in a queue. The generator of the Markov process is usually built up of a small number of parameters. For instance, the structure of the generator for the M/M/1/K system, with a Poisson arrival process with rate λ , exponentially distributed service times with rate μ , a single server and a buffer capacity for K customers, the one in service included, yields the following matrix A , cf. (5):

$$A = -G^T = \begin{pmatrix} \lambda & -\mu & 0 & \cdots & 0 & 0 \\ -\lambda & \lambda + \mu & -\mu & \cdots & 0 & 0 \\ 0 & -\lambda & \lambda + \mu & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda + \mu & -\mu \\ 0 & 0 & 0 & \cdots & -\lambda & \mu \end{pmatrix}. \quad (53)$$

The problem of determining uncertainty intervals for the stationary queue length probabilities ν , cf. (4), or the mean queue length $E\{N\}$, given some uncertainty region for the arrival rate λ and the service rate μ , does not fall into the framework of the optimization problem (11), because the parameters λ and μ occur in multiple columns of the matrix A . Figure 1 contains the graphs of the stationary queue length probabilities ν_j , $j = 0, \dots, 4$, for the M/M/1/4 system as functions of the load $\rho \doteq \lambda/\mu$. It illustrates that in general the probabilities ν_j , $j = 1, \dots, K-1$, have a maximum for some load ρ since they vanish both as $\rho \downarrow 0$ and as $\rho \rightarrow \infty$. This indicates that it will not be possible to find uncertainty intervals for these probabilities via an LP formulation.

Remark 11 If we nevertheless solve LP problems of the type (22) for the stationary queue length probabilities ν_j , $j = 0, \dots, 4$, of the M/M/1/4 system constructed with matrices A_- and A_+ in (20) based on a fixed service rate $\mu = 1$ and the arrival rate on the interval $0.5 \leq \lambda \leq 0.7$ in (53), we correctly find the intervals $0.361 \leq \nu_0 \leq 0.516$, $0.032 \leq \nu_4 \leq 0.087$, where the bounds are related to generators with either $\lambda = 0.5$ or $\lambda = 0.7$. But these LP solutions yield too wide intervals for not only the probability ν_1 , which has a maximum of 0.261 at $\lambda = 0.568$, cf. Figure 1, but also ν_2 and ν_3 , of which the boundaries are related to generators with different values of λ in its rows. For the mean queue length, which is an increasing function of λ , the correct interval $0.839 \leq E\{N\} \leq 1.323$ is obtained. $\square$

A model with more independent parameters is, for instance, the M/M/1/K system with balking customers. Balking is the phenomenon that customers may decide not to join a queue upon arrival. Let ψ_m denote the probability that a potential customer from a Poisson arrival stream joins the queue when the server is busy and m customers are already waiting for service, $m = 0, 1, \dots, K-1$. A logical constraint is that customers are more reluctant to join the queue when the queue is longer:

$$0 = \psi_{K-1} \leq \psi_{K-2} \leq \cdots \leq \psi_1 \leq \psi_0 \leq 1. \quad (54)$$

Figure 1: The probabilities ν_j , $j = 0, \dots, 4$, for the M/M/1/4 system as functions of ρ .

Further, the probabilities must be estimated in practice. The structure of the generator for this system yields the following matrix A :

$$A = -G^T = \begin{pmatrix} \lambda & -\mu & 0 & \cdots & 0 & 0 & 0 \\ -\lambda & \psi_0\lambda + \mu & -\mu & \cdots & 0 & 0 & 0 \\ 0 & -\psi_0\lambda & \psi_1\lambda + \mu & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \psi_{K-3}\lambda + \mu & -\mu & 0 \\ 0 & 0 & 0 & \cdots & -\psi_{K-3}\lambda & \psi_{K-2}\lambda + \mu & -\mu \\ 0 & 0 & 0 & \cdots & 0 & -\psi_{K-2}\lambda & \mu \end{pmatrix}. \quad (55)$$

The problem of determining uncertainty intervals for the stationary queue length probabilities $\boldsymbol{\nu}$, cf. (4), or the mean queue length $E\{N\}$, given some uncertainty region for the probabilities ψ_m , $m = 0, 1, \dots, K-1$, but with fixed arrival rate λ and service rate μ , falls into the column-wise framework of the optimization problem (11), provided that the constraints (54) are ignored or not explicitly required because the uncertainty intervals for the probabilities ψ_m , $m = 0, 1, \dots, K-1$, are so small that they do not overlap.

Example 8 Consider the following case: $\lambda = \mu = 1$, $K = 3$, $\frac{1}{2} \leq \psi_0 \leq 1$, $0 \leq \psi_1 \leq \frac{1}{2}$. Since the uncertainty intervals for ψ_0 and ψ_1 do not overlap, the uncertainty intervals for the stationary probabilities and the mean queue length can be determined by solving LP problems of the type (22):

$$\frac{2}{7} \leq \nu_0 \leq \frac{2}{5}, \quad \frac{2}{7} \leq \nu_1 \leq \frac{2}{5}, \quad \frac{2}{11} \leq \nu_2 \leq \frac{1}{3}, \quad 0 \leq \nu_3 \leq \frac{1}{7}, \quad \frac{4}{5} \leq E\{N\} \leq \frac{9}{7}.$$

In this case, the reduced LP problems (26) yield the same uncertainty intervals. Two inequalities of these reduced LP problems imply that $\nu_0 = \nu_1$ so that we can eliminate these probabilities from the remaining inequalities together with the normalization: $\nu_0 = \nu_1 = \frac{1}{2}(1 - \nu_2 - \nu_3)$. The remaining six inequalities,

$$5\nu_2 + \nu_3 \geq 1, \quad 3\nu_2 - \nu_3 \leq 1, \quad 2\nu_3 \leq \nu_2, \quad 3\nu_2 + \nu_3 \leq 1, \quad 7\nu_2 - 3\nu_3 \geq 1, \quad \nu_3 \geq 0,$$

determine a quadrangle in the (ν_2, ν_3) plane as shown in Figure 2 (right). The corner points are $(\frac{1}{5}, 0)$, $(\frac{1}{3}, 0)$, $(\frac{2}{7}, \frac{1}{7})$, and $(\frac{2}{11}, \frac{1}{11})$. The boundaries of the uncertainty interval for the mean queue length $E\{N\}$ correspond to the corner points $\psi_0 = \frac{1}{2}$, $\psi_1 = 0$, or $\nu_0 = \nu_1 = \frac{2}{5}$, $\nu_2 = \frac{1}{5}$, $\nu_3 = 0$, and $\psi_0 = 1$, $\psi_1 = \frac{1}{2}$, or $\nu_0 = \nu_1 = \nu_2 = \frac{2}{7}$, $\nu_3 = \frac{1}{7}$, respectively.

Next, consider the following related case with partially overlapping uncertainty intervals: $\lambda = \mu = 1$, $K = 3$, $\frac{2}{5} \leq \psi_0 \leq \frac{4}{5}$, $\frac{1}{5} \leq \psi_1 \leq \frac{3}{5}$. Here, we would like to add $\psi_1 \leq \psi_0$, cf. (54), to the uncertainty region. But first

Figure 2: Uncertainty regions for M/M/1/3 with balking in (ψ_0, ψ_1) (left) and (ν_2, ν_3) (right) planes.

we determine the uncertainty intervals for the stationary probabilities and the mean queue length by solving LP problems of the type (22):

$$\frac{25}{82} \leq \nu_0 \leq \frac{25}{62}, \quad \frac{25}{82} \leq \nu_1 \leq \frac{25}{62}, \quad \frac{5}{33} \leq \nu_2 \leq \frac{10}{37}, \quad \frac{1}{31} \leq \nu_3 \leq \frac{6}{41}, \quad \frac{51}{62} \leq E\{N\} \leq \frac{101}{82}.$$

Again, the reduced LP problems (26) yield the same uncertainty intervals. As before we have $\nu_0 = \nu_1 = \frac{1}{2}(1 - \nu_2 - \nu_3)$. The remaining six inequalities,

$$6\nu_2 + \nu_3 \geq 1, \quad 8\nu_2 - 3\nu_3 \leq 2, \quad 5\nu_3 \leq 3\nu_2, \quad 7\nu_2 + 2\nu_3 \leq 2, \quad 9\nu_2 - 4\nu_3 \geq 1, \quad 5\nu_3 \geq \nu_2,$$

determine a quadrangle in the (ν_2, ν_3) plane as shown in Figure 3. The corner points are $(\frac{5}{31}, \frac{1}{31})$, $(\frac{10}{37}, \frac{2}{37})$, $(\frac{10}{41}, \frac{6}{41})$, and $(\frac{5}{33}, \frac{1}{11})$. The last point, however, is infeasible since it corresponds to the values $\psi_0 = \frac{2}{5} < \psi_1 = \frac{3}{5}$. This only affects the above lower bound on ν_2 . The boundaries of the uncertainty interval for the mean queue length $E\{N\}$ correspond to the corner points $\psi_0 = \frac{2}{5}$, $\psi_1 = \frac{1}{5}$, or $\nu_0 = \nu_1 = \frac{25}{62}$, $\nu_2 = \frac{5}{31}$, $\nu_3 = \frac{1}{31}$, and $\psi_0 = \frac{4}{5}$, $\psi_1 = \frac{3}{5}$, or $\nu_0 = \nu_1 = \frac{25}{82}$, $\nu_2 = \frac{10}{41}$, $\nu_3 = \frac{6}{41}$, respectively.

Figure 3: Uncertainty regions for M/M/1/3 with balking in (ψ_0, ψ_1) (left) and (ν_2, ν_3) (right) planes.

The condition $\psi_1 = \psi_0$ leads by substitution into the actual expressions for the stationary queue length probabilities,

$$\nu_0 = \nu_1 = \frac{1}{2 + \psi_0 + \psi_0\psi_1}, \quad \nu_2 = \frac{\psi_0}{2 + \psi_0 + \psi_0\psi_1}, \quad \nu_3 = \frac{\psi_0\psi_1}{2 + \psi_0 + \psi_0\psi_1},$$

to a nonlinear relation between ν_3 and ν_2 :

$$\nu_3 = \frac{1}{2} \left[1 - \nu_2 - \sqrt{(1 - \nu_2)^2 - 8\nu_2^2} \right].$$

This leads to a curved boundary between the points $(\frac{5}{32}, \frac{1}{16})$ and $(\frac{15}{74}, \frac{9}{74})$. Figure 3 (right) illustrates that the uncertainty region in the (ν_2, ν_3) plane is not convex. The actual lower bound on ν_2 , taking into account the condition $\psi_0 \geq \psi_1$, follows as $\frac{5}{32}$. We finally note that the above corner points $(\frac{5}{32}, \frac{1}{16})$ and $(\frac{15}{74}, \frac{9}{74})$ (but of course not the non convex boundary) are generated if we replace the inequality $9\nu_2 - 4\nu_3 \geq 1$ which is the only inequality that corresponds to an inequality involving both $\psi_0^- = \frac{2}{5}$ and $\psi_1^+ = \frac{3}{5}$, namely $-\psi_0^- \nu_1 + (1 + \psi_1^+) \nu_2 - \nu_3 \geq 0$, by two inequalities $8\nu_2 - 4\nu_3 \geq 1$ and $19\nu_2 - 7\nu_3 \geq 3$, corresponding to $-\psi_0^- \nu_1 + (1 + \psi_0^-) \nu_2 - \nu_3 \geq 0$ and $-\psi_1^+ \nu_1 + (1 + \psi_1^+) \nu_2 - \nu_3 \geq 0$, respectively. $\square$

6 Final remarks

In this last section we would like to discuss other possible applications of our analysis. First of all note that our analysis can also be applied to general linear programming problems. In that sense our paper also extends the results in [13] and [4] to more general cases.

Moreover, the analysis given in this paper can also be applied for solving general linear systems of equations with inexact data. Note that the well-known total least-squares method ([6]) tries to find the solution which fits best to all possible realizations of the data. However, we try to analyze the effect of the uncertain data on the solution of the system.

Another application of our analysis is fuzzy Markov chains. The basic idea used in literature on fuzzy Markov chains, see e.g. [2], is to transform a fuzzy chain with finite capacity to a family of conventional crisp chains by applying the α -cut approach. For each fixed value of α the problem reduces exactly to a Markov chain with uncertain transition matrix, for which the methods developed in this paper can be used.

It is also useful to investigate whether our techniques can be used in Markov decision problems with uncertain probabilities.

Finally we would like to describe the possible application in the fascinating field of the so-called “small world phenomenon”. See e.g. [5]. This phenomenon can be modeled as a Markov chain with uncertainties. Using our analysis we may be able to calculate the so-called maximum or average mean hitting time exactly for a specifically chosen network topology and uncertainty structure.

Acknowledgements

The authors thank Dr. Anja de Waegenaere for some useful comments.

References

- [1] Boyd, S., L. Vandenberghe (2004), *Convex Optimization*, Cambridge University Press.
- [2] Chen, S.-P. (2004) Parametric nonlinear programming for analyzing fuzzy queues with finite capacity, *European Journal of Operational Research* **157**, 429–438.
- [3] Dantzig, G.B. (1963) *Linear Programming and Extensions*, 11th printing (1998), Princeton University Press.
- [4] Fiedler, M., J. Nedoma, J. Ramik, J. Rohn, and K. Zimmermann (2006) *Linear Optimization Problems with Inexact Data*, Springer.
- [5] Higham, D.J. (2007) A matrix perturbation view of the small world phenomenon, *SIAM Review* **49**, 1, 91–108.

- [6] Huffel, S. V., P. Lemmerling (2002) *Total Least Squares and Errors-in-Variables Modeling: Analysis, Algorithms and Applications*, Kluwer Academic Publishers.
- [7] Ke, J.-C., H.-I. Huang, C.-H. Lin (2007) On retrial queueing model with fuzzy parameters, *Physica A* **374**, 272–280.
- [8] Kruse, R., R. Buck-Emden, R. Cordes (1987) Processor power considerations - an application of fuzzy Markov chains, *Fuzzy Sets and Systems* **21**, 289–299.
- [9] Kurano, M., M. Yasuda, J. Nakagami, Y. Yoshida (2006) A fuzzy approach to Markov decision processes with uncertain transition probabilities, *Fuzzy Sets and Systems* **157**, 2674–2682.
- [10] Li, R.J., E.S. Lee (1989) Analysis of fuzzy queues, *Computers and Mathematical Applications* **17**, 1143–1147.
- [11] Satia, J.K., R.E. Lave (1973) Markovian decision processes with uncertain transition probabilities, *Operations Research* **21** (3), 728–740.
- [12] Schweitzer, P.J. (1968) Perturbation theory and finite Markov chains, *Journal of Applied Probability* **5**, 401–413.
- [13] Serafini, P. (2005) Linear programming with variable matrix entries, *Operations Research Letters* **33** (2), 165–170.
- [14] Silver, E. (1963) Markovian decision processes with uncertain transition probabilities or rewards, Interim Techn. Rep. No. 1, Operations Research Center, MIT.
- [15] Smith, J.L. (1971) Approximate stationary probability vectors of a finite Markov chain, *SIAM Journal of Applied Mathematics* **20**, 612–618.
- [16] Stanford, R.E. (1982) The set of limiting distributions for a Markov chain with fuzzy transition probabilities, *Fuzzy Sets and Systems* **7**, 71–78.
- [17] Walton, G.H., J.H. Poore (2000) Measuring complexity and coverage of software specifications, *Information and Software Technology* **42**, 859–872.
- [18] Wolff, R.W. (1989) *Stochastic Modeling and the Theory of Queues*, Prentice-Hall.
- [19] Xue, J. (2001) Blockwise perturbation theory for nearly uncoupled Markov chains and its application, *Linear Algebra and its Applications* **326**, 173–191.