The Version of Record of this manuscript has been published and is available in *Communications in Statistics–Theory and Methods* on 25 Mar 2020. https://doi.org/10.1080/03610926.2020.1742921

ON MONOTONICALLY PROCEEDING STRUCTURES AND STEPWISE INCREASING TRANSITION MATRICES OF MARKOV CHAINS

Marie-Anne Guerry and Philippe Carette

Department Business Technology and Operations
Vrije Universiteit Brussel
Pleinlaan 2, B-1050 Brussels, Belgium
marie-anne.guerry@vub.be

Department of Economics
Ghent University
Sint-Pietersplein 6, B-9000 Ghent, Belgium
philippe.carette@ugent.be

Key Words: Markov chain, transition matrix, eigenvalue, eigenvector, monotonic evolution.

ABSTRACT

In general, the transition matrix of a Markov chain is a stochastic matrix. For a system that is modeled by a Markov chain, the transition matrix must reflect the characteristics of that system. The present paper introduces a particular class of transition matrices in order to model Markov systems for which, as the length of the time interval becomes greater, a transition from one state to another is more likely. We call these transition matrices stepwise increasing. Moreover in some contexts it is less desirable that the stocks fluctuate over time. In those situations, one is interested in monotonically proceeding stock vectors. This paper examines monotonically proceeding stock vectors and stepwise increasing transition matrices. We present conditions on the transition matrix such that all stock vectors are monotonically proceeding. In particular, the set of monotonically proceeding vectors is characterized for the two-state and three-state cases.

1 Introduction

Markov chains are useful modeling tools in a number of domains such as marketing, finance, inter-generational income mobility and manpower planning. Such models allow to gain insight in how the stocks, i.e. the number of items in each of the states of the Markovian system, change over time. In a manpower planning context for example, the states are the grades of the organization or firm and the stocks are the number of employees in each of the grades.

The context and characteristics of the studied system may impose additional constraints on the Markov model. For instance, managers may not prefer to have numbers of employees that fluctuate (too much) over time (Guerry, 2013). In such a case, it is desirable to have the stocks as monotonic functions of time. Furthermore, the transition matrix of the Markov chain, the elements of which are the transition probabilities between states over one period of time, is usually restricted to a specific class of stochastic matrices. For example, monotone stochastic matrices were used to study inter-generational income mobility, see (Dardanoni, 1995). Monotone Markov models are also useful in examining the evolution of credit ratings based on a credit rating transition matrix, see (Jarrow, Lando, & Turnbull, 1997). For modeling hierarchical manpower systems, the subset of state-wise monotone transition matrices is introduced and investigated in (Guerry, 2014, 2017).

In this paper, we examine the time-evolution of the stocks. A necessary condition to have all stocks evolving monotonically over time is that the Markov transition matrix $\boldsymbol{Q}$ satisfies $(\boldsymbol{Q}^2)_{ij} \geq \boldsymbol{Q}_{ij}$ for all states $i \neq j$. We call such class of transition matrices *stepwise increasing*. Transition matrices of this kind are useful for modeling systems where the transition probability from one state to another state increases with the length of the time-interval under consideration. A number of practical applications exhibit empirical transition matrices that are stepwise increasing, such as credit risk migration (Altman & Rijken, 2004). Stepwise increasing transition matrices are also of interest to model job mobility for organizations that value seniority, as a result of which employees with greater seniority are more likely to be promoted (Mills, 1985; Ng, Sorensen, Eby, & Feldman, 2007).

The paper is organized as follows. In section 2, we introduce the concept of a monotonically proceeding structure ('mp-structure' in short), given a transition matrix $\boldsymbol{Q}$. We show a connection between t-step ($t > 2$) and 2-step mp-structures. In section 3, we study the properties of the set of 2-step mp-structures and establish a link with stepwise increasing transition matrices. In section 4, we prove some properties of stepwise increasing stochastic matrices. Section 5 is devoted to the description of the set of 2-step mp-structures for the three-state case. We also establish a result about the t-step mp-property of these structures where $t > 2$. Finally, we summarize our findings and formulate some questions for further research in section 6.

2 Monotonically proceeding structures

For a Markov chain with k states and transition matrix $\boldsymbol{Q}$, the stock vector $\boldsymbol{n}(t)$ at time t satisfies $\boldsymbol{n}(t) = \boldsymbol{n}(0)\boldsymbol{Q}^t$, where $\boldsymbol{n}(0)$ is the initial stock vector. Since $\boldsymbol{Q}$ is a stochastic matrix, the total stock $N = \sum_{i=1}^k \boldsymbol{n}_i(t)$ is constant and does not depend on t. Thus, the stock vector $\boldsymbol{n}(t)$ is uniquely determined by the probability distribution vector $\boldsymbol{v}(t) = \frac{1}{N}\boldsymbol{n}(t) \in \Pi_k$, where

$$\Pi_k = \{\boldsymbol{v} \in \mathbb{R}^{1\times k} \,|\, \forall i : \boldsymbol{v}_i \geq 0 \text{ and } \sum_{i=1}^k \boldsymbol{v}_i = 1\}.$$

We call any vector $\boldsymbol{v} \in \Pi_k$ a *structure*.

In the following, we study the evolution of the stock vectors by examining their corresponding structures.

Definition 1. *Let $\boldsymbol{v} \in \Pi_k$, where $k \geq 2$ and $t \in \mathbb{N}$, $t \geq 2$. We call $\boldsymbol{v}$ a t-*step monotonically proceeding structure (‘mp’-structure in short) with respect to the transition matrix $\boldsymbol{Q}$ *if and only if the sequence*

$$\boldsymbol{v}_i, (\boldsymbol{vQ})_i, (\boldsymbol{vQ}^2)_i, \ldots, (\boldsymbol{vQ}^t)_i$$

is monotonic for all $i \in \{1, \ldots, k\}$. The set of all t-step mp-structures with respect to $\boldsymbol{Q}$ is denoted by $\mathcal{MP}(\boldsymbol{Q}, t)$.

We observe that $\boldsymbol{v} \in \mathcal{MP}(\boldsymbol{Q}, 2)$ if and only if $(\boldsymbol{vQ})_i$ lies between $\boldsymbol{v}_i$ and $\left(\boldsymbol{vQ}^2\right)_i$ for any $i \in \{1, \ldots, k\}$ or equivalently

$$(\boldsymbol{v}(\boldsymbol{Q} - \mathbf{I}))_i \left(\boldsymbol{v}(\boldsymbol{Q}^2 - \boldsymbol{Q})\right)_i \geq 0 \quad \text{for any } i \in \{1, \ldots, k\},$$

where $\mathbf{I}$ is the identity matrix of order k. This can be rewritten as

$$(\boldsymbol{vL})_i \, (\boldsymbol{vM})_i \geq 0 \quad \text{for any } i \in \{1, \ldots, k\}, \tag{1}$$

by introducing the matrices

$$\boldsymbol{L} := \boldsymbol{Q} - \mathbf{I} \quad \text{and} \quad \boldsymbol{M} := \boldsymbol{Q}^2 - \boldsymbol{Q} = \boldsymbol{LQ} = \boldsymbol{QL}. \tag{2}$$

In this way,

$$\mathcal{MP}(\boldsymbol{Q},2) = \{\boldsymbol{v} \in \Pi_k \,|\, \forall i : (\boldsymbol{vL})_i \, (\boldsymbol{vM})_i \geq 0\}.$$

Note that $\boldsymbol{L}$ is a *Metzler*-matrix, i.e. a square matrix with non-negative off-diagonal elements. The matrix $\boldsymbol{M}$ is not necessarily a Metzler-matrix. Note also that both $\boldsymbol{L}$ and $\boldsymbol{M}$ have zero row-sums, since $\boldsymbol{Q}$ is stochastic.

The following proposition stresses the importance of the set $\mathcal{MP}(\boldsymbol{Q},2)$ in assessing the monotonic evolution of a structure.

Theorem 1. *Let $t \in \mathbb{N}$ and $t \geq 3$. If $\boldsymbol{v} \in \mathcal{MP}(\boldsymbol{Q},t)$, then $\boldsymbol{v} \in \mathcal{MP}(\boldsymbol{Q},t-1)$ and $\boldsymbol{vQ}^{t-2} \in \mathcal{MP}(\boldsymbol{Q},2)$. The converse holds if, in addition, $\boldsymbol{vQ}^{t-2} \neq \boldsymbol{vQ}^{t-1}$ componentwise.*

Proof. The first statement is an immediate consequence of definition 1.

For the converse statement, we have by definition 1 that, for each i, the sequences $\big((\boldsymbol{vQ}^\tau)_i\big)_{\tau=0}^{t-1}$ and $\big((\boldsymbol{vQ}^\tau)_i\big)_{\tau=t-2}^{t}$ are both monotonic. Also, they overlap at the elements $(\boldsymbol{vQ}^{t-2})_i$ and $(\boldsymbol{vQ}^{t-1})_i$. Since these two elements are assumed to be different, the joined sequence $\big((\boldsymbol{vQ}^\tau)_i\big)_{\tau=0}^{t}$ must also be monotonic. □

It follows from theorem 1 that a necessary condition to ensure a structure $\boldsymbol{v}$ to be t-step monotonically proceeding, is that $\boldsymbol{v}$ and all its succeeding structures $\boldsymbol{vQ}, \ldots, \boldsymbol{vQ}^{t-2}$ belong to $\mathcal{MP}(\boldsymbol{Q},2)$. In the next section we therefore examine the set of 2-step mp-structures with respect to $\boldsymbol{Q}$.

3 Properties of the set $\mathcal{MP}(\boldsymbol{Q},2)$

A 2-step mp-structure $\boldsymbol{v}$ satisfies $(\boldsymbol{vL})_i \, (\boldsymbol{vM})_i \geq 0$ for any $i \in \{1,\ldots,k\}$. Therefore in investigating properties of $\mathcal{MP}(\boldsymbol{Q},2)$, the following sets will be useful, where $i \in \{1,\ldots,k\}$:

- $\mathcal{H}_i^{\boldsymbol{L}} := \{\boldsymbol{x} \in \mathbb{R}^{1\times k} \,|\, (\boldsymbol{xL})_i = 0\}$, $\mathcal{H}_i^{\boldsymbol{M}} := \{\boldsymbol{x} \in \mathbb{R}^{1\times k} \,|\, (\boldsymbol{xM})_i = 0\}$
- $\mathcal{G}_i := \{\boldsymbol{x} \in \mathbb{R}^{1\times k} \,|\, (\boldsymbol{xL})_i(\boldsymbol{xM})_i \geq 0\}$
- $\mathcal{N}_i := \{\boldsymbol{x} \in \mathbb{R}^{1\times k} \,|\, (\boldsymbol{xL})_i(\boldsymbol{xM})_i < 0\} = \mathbb{R}^{1\times k} \setminus \mathcal{G}_i$

Note that $\mathcal{H}_i^{\boldsymbol{L}} \subset \mathcal{G}_i$ and $\mathcal{H}_i^{\boldsymbol{M}} \subset \mathcal{G}_i$. Moreover,

$$\boldsymbol{v} \in \mathcal{H}_i^{\boldsymbol{M}} \quad \Leftrightarrow \quad \boldsymbol{vQ} \in \mathcal{H}_i^{\boldsymbol{L}} \tag{3}$$

since $\boldsymbol{M} = \boldsymbol{QL}$. With the use of these notations, we can now write

$$\mathcal{MP}(\boldsymbol{Q}, 2) = \Pi_k \cap (\cap_{i=1}^k \mathcal{G}_i) = \Pi_k \setminus (\cup_{i=1}^k \mathcal{N}_i) \tag{4}$$

In what follows we denote by $\mathbf{e_i}$ the structure having the i-th element equal to 1 and all other elements equal to 0. Also, define

$$\mathcal{U}_k := \{\boldsymbol{x} \in \mathbb{R}^{1\times k} \mid \sum_{i=1}^{k} \boldsymbol{x}_i = 1\}.$$

We also need a notation for the set of left eigenvectors of $\boldsymbol{Q}$ corresponding with the eigenvalue λ:

$$\mathcal{E}_\lambda := \{\boldsymbol{x} \in \mathbb{R}^{1\times k} \,|\, \boldsymbol{x}(\boldsymbol{Q} - \lambda \mathbf{I}) = \mathbf{o}\},$$

where $\mathbf{o}$ denotes the $1 \times k$ vector consisting of zeroes. Remark that, since $\boldsymbol{Q}$ is a stochastic matrix, $\lambda = 1$ is always an eigenvalue of $\boldsymbol{Q}$. We now focus on the properties of the sets $\mathcal{G}_i$.

Proposition 1. *Let $\boldsymbol{\alpha} \in \Pi_k \cap \mathcal{E}_1$. Then, for any $i \in \{1, \ldots, k\}$,*

a. $\boldsymbol{\alpha} \in \mathcal{H}_i^{\boldsymbol{L}} \cap \mathcal{H}_i^{\boldsymbol{M}}$, hence $\boldsymbol{\alpha} \in \mathcal{G}_i$

b. If $\boldsymbol{M}$ is a Metzler-matrix without zero columns, $\mathbf{e_i}$ and $\mathbf{e_j}$ cannot belong to a same open half-space determined by the hyperplanes $\mathcal{H}_i^{\boldsymbol{L}}$ and $\mathcal{H}_i^{\boldsymbol{M}}$ for all $j \in \{1, \ldots, k\}$, $j \neq i$.

c. If $\boldsymbol{M}$ is a Metzler-matrix, $\mathbf{e_i} \in \mathcal{MP}(\boldsymbol{Q}, 2)$.

d. If the i-th column vector of $\boldsymbol{M}$ is not the null vector, then $\mathcal{H}_i^{\boldsymbol{L}} \neq \mathcal{H}_i^{\boldsymbol{M}}$ implies $\mathcal{N}_i \cap \mathcal{U}_k \neq \emptyset$. Moreover, if $\boldsymbol{\alpha}$ lies in the interior of Π_k, then $\mathcal{U}_k$ can be replaced by Π_k in the above statement.

e. $\mathcal{E}_\lambda + \boldsymbol{\alpha} \subset \mathcal{G}_i$ for any non-negative eigenvalue λ of $\boldsymbol{Q}$

f. If $\boldsymbol{x} \in \mathcal{G}_i$, then $\boldsymbol{u}(\theta) := \theta\boldsymbol{\alpha} + (1-\theta)\boldsymbol{x} \in \mathcal{G}_i$ for any $\theta \in \mathbb{R}$.

Proof. Recall that $\boldsymbol{L} = \boldsymbol{Q} - \mathbf{I}$ and $\boldsymbol{M} = \boldsymbol{QL} = \boldsymbol{LQ}$. Since $\boldsymbol{\alpha Q} = \boldsymbol{\alpha}$, we have $\boldsymbol{\alpha L} = \mathbf{o}$ and $\boldsymbol{\alpha M} = \mathbf{o}$.

a. Hence $\boldsymbol{\alpha} \in \mathcal{H}_i^{\boldsymbol{L}} \cap \mathcal{H}_i^{\boldsymbol{M}}$ by definition.

b. If $\boldsymbol{M}$ has no zero columns, then $\boldsymbol{L}$ has no zero columns, since $\boldsymbol{M} = \boldsymbol{QL}$. Hence $\mathcal{H}_i^{\boldsymbol{L}}$ and $\mathcal{H}_i^{\boldsymbol{M}}$ are both hyperplanes. Let $j \neq i$. Suppose $\mathbf{e_i}$ and $\mathbf{e_j}$ belong to a same open half-space determined by $\mathcal{H}_i^{\boldsymbol{L}}$. Then $(\mathbf{e_i}\boldsymbol{L})_i$ and $(\mathbf{e_j}\boldsymbol{L})_i$ are both non-zero and have the same sign. But this is impossible, since $(\mathbf{e_i}\boldsymbol{L})_i = \boldsymbol{L}_{ii} = \boldsymbol{Q}_{ii} - 1 \leq 0$ and $(\mathbf{e_j}\boldsymbol{L})_i = \boldsymbol{L}_{ji} = \boldsymbol{Q}_{ji} \geq 0$. In a similar vein, suppose $\mathbf{e_i}$ and $\mathbf{e_j}$ belong to a same open half-space determined by $\mathcal{H}_i^{\boldsymbol{M}}$. Then $(\mathbf{e_i}\boldsymbol{M})_i$ and $(\mathbf{e_j}\boldsymbol{M})_i$ are both non-zero and have the same sign. Again, this cannot happen since $(\mathbf{e_i}\boldsymbol{M})_i = \boldsymbol{M}_{ii} \leq 0$ and $(\mathbf{e_j}\boldsymbol{M})_i = \boldsymbol{M}_{ji} \geq 0$ if $\boldsymbol{M}$ is a Metzler-matrix.

c. If $\boldsymbol{M}$ is Metzler, then, for all j, $(\mathbf{e_i}\boldsymbol{L})_j = \boldsymbol{L}_{ij}$ and $(\mathbf{e_i}\boldsymbol{M})_j = \boldsymbol{M}_{ij}$ are either zero or have both the same sign. Hence $\mathbf{e_i} \in \cap_{j=1}^k \mathcal{G}_j$ and thus $\mathbf{e_i} \in \mathcal{MP}(\boldsymbol{Q}, 2)$ by equation (4).

d. If the i-th column vector of $\boldsymbol{M}$ is not the null vector, so is the i-th column vector of $\boldsymbol{L}$, since $\boldsymbol{M} = \boldsymbol{QL}$. Therefore, both $\mathcal{H}_i^{\boldsymbol{L}}$ and $\mathcal{H}_i^{\boldsymbol{M}}$ are hyperplanes in $\mathbb{R}^k$. Note that $\mathcal{H}_i^{\boldsymbol{L}} = \text{aff}\left((\mathcal{H}_i^{\boldsymbol{L}} \cap \mathcal{U}_k) \cup \{\mathbf{o}\}\right)$ and $\mathcal{H}_i^{\boldsymbol{M}} = \text{aff}\left((\mathcal{H}_i^{\boldsymbol{M}} \cap \mathcal{U}_k) \cup \{\mathbf{o}\}\right)$, where aff$(S)$ denotes the affine hull of the set S. Now, if $\mathcal{H}_i^{\boldsymbol{L}} \neq \mathcal{H}_i^{\boldsymbol{M}}$, then $\mathcal{H}_i^{\boldsymbol{L}} \cap \mathcal{U}_k \neq \mathcal{H}_i^{\boldsymbol{M}} \cap \mathcal{U}_k$ resulting in a partitioning of $\mathcal{U}_k$ into non-empty regions defined by all possible sign combinations of $(\boldsymbol{xL})_i$ and $(\boldsymbol{xM})_i$ when $\boldsymbol{x} \in \mathcal{U}_k$. Pick a $\boldsymbol{v} \in \mathcal{U}_k$ from a region where $(\boldsymbol{vL})_i$ and $(\boldsymbol{vM})_i$ are having opposite signs, then $\boldsymbol{v} \in \mathcal{N}_i$. Hence $\mathcal{N}_i \cap \mathcal{U}_k \neq \emptyset$.

Now, define $\boldsymbol{v}(t) := \boldsymbol{\alpha} + t(\boldsymbol{v} - \boldsymbol{\alpha})$, $t \in]0, 1]$. Note that $\boldsymbol{v}(t) \in \mathcal{U}_k$. Since $\boldsymbol{\alpha L} = \boldsymbol{\alpha M} = \mathbf{o}$, we have $(\boldsymbol{v}(t)\boldsymbol{L})_i = t(\boldsymbol{vL})_i$ and $(\boldsymbol{v}(t)\boldsymbol{M})_i = t(\boldsymbol{vM})_i$, so $(\boldsymbol{v}(t)\boldsymbol{L})_i(\boldsymbol{v}(t)\boldsymbol{M})_i = t^2(\boldsymbol{vL})_i(\boldsymbol{vM})_i < 0$. Hence $\boldsymbol{v}(t) \in \mathcal{N}_i \cap \mathcal{U}_k$. Now, $\boldsymbol{v}(t) \to \boldsymbol{\alpha}$ as $t \downarrow 0$. If $\boldsymbol{\alpha}$ lies in the interior of Π_k, so will $\boldsymbol{v}(t)$ for large enough t and thus $\mathcal{N}_i \cap \Pi_k \neq \emptyset$.

e. Let $\lambda \geq 0$ and $\boldsymbol{x} \in \mathcal{E}_\lambda + \boldsymbol{\alpha}$. Then $\boldsymbol{x} - \boldsymbol{\alpha} \in \mathcal{E}_\lambda$ so that $(\boldsymbol{x} - \boldsymbol{\alpha})\boldsymbol{Q} = \lambda(\boldsymbol{x} - \boldsymbol{\alpha})$ and

$$\boldsymbol{xQ} = \lambda\boldsymbol{x} + (1 - \lambda)\boldsymbol{\alpha}. \tag{5}$$

Postmultiplying both sides of (5) by $\boldsymbol{L}$, we obtain $\boldsymbol{xM} = \boldsymbol{xQL} = \lambda\boldsymbol{xL}$ since $\boldsymbol{\alpha L} = \mathbf{o}$. Consequently,

$$(\boldsymbol{xL})_i(\boldsymbol{xM})_i = \lambda(\boldsymbol{xL})_i^2$$

which is non-negative because $\lambda \geq 0$. Hence $\boldsymbol{x} \in \mathcal{G}_i$.

f. Let $x \in \mathcal{G}_i$. Since $\boldsymbol{\alpha L} = \mathbf{o} = \boldsymbol{\alpha M}$, we have for $\boldsymbol{u}(\theta) = \theta\boldsymbol{\alpha} + (1 - \theta)\boldsymbol{x}$ that $\boldsymbol{u}(\theta)\boldsymbol{L} = (1 - \theta)\boldsymbol{xL}$ and $\boldsymbol{u}(\theta)\boldsymbol{M} = (1-\theta)\boldsymbol{xM}$. Consequently, $(\boldsymbol{u}(\theta)\boldsymbol{L})_i(\boldsymbol{u}(\theta)\boldsymbol{M})_i = (1-\theta)^2(\boldsymbol{xL})_i(\boldsymbol{xM})_i$ which is non-negative since $\boldsymbol{x} \in \mathcal{G}_i$.

□

Theorem 2. *$\mathcal{MP}(\boldsymbol{Q}, 2)$ is star-convex.*

Proof. Take $\boldsymbol{\alpha} \in \Pi_k \cap \mathcal{E}_1$. By proposition 1(a), $\boldsymbol{\alpha}$ is an element of $\mathcal{MP}(\boldsymbol{Q}, 2)$. It also follows from proposition 1(f) that for any vector $\boldsymbol{v} \in \mathcal{MP}(\boldsymbol{Q}, 2)$ the segment $[\boldsymbol{\alpha}, \boldsymbol{v}]$ is contained in $\mathcal{MP}(\boldsymbol{Q}, 2)$. Therefore, $\mathcal{MP}(\boldsymbol{Q}, 2)$ is star-convex with respect to $\boldsymbol{\alpha}$. □

We now turn to characterizing the transition matrix $\boldsymbol{Q}$ such that all structures are 2-step monotonically proceeding. A sufficient condition for this to occur is that all sets $\mathcal{N}_j$ are empty. By proposition 1(d), this is connected with the coincidence of $\mathcal{H}_j^{\boldsymbol{L}}$ and $\mathcal{H}_j^{\boldsymbol{M}}$. We show that this in turn is equivalent with the j-th column of $\boldsymbol{Q}$ being quasi-constant, a concept which we define below.

Definition 2. *We call the j-th column of the stochastic matrix $\boldsymbol{Q}$* quasi-constant *if that column has identical off-diagonal elements and a different diagonal element, i.e.*

$$\boldsymbol{Q}\mathbf{e_j}' = \rho_j \mathbf{e}' + \theta_j \mathbf{e_j}', \quad \rho_j \in [0,1],\ \rho_j + \theta_j \in [0,1],\ \theta_j \neq 0,$$

where $\mathbf{e}$ is the $(1 \times k)$-vector having all elements equal to 1 and the prime denotes the transpose. If $\boldsymbol{Q}$ is not the identity matrix but has all columns quasi-constant, then $\boldsymbol{Q}$ is called quasi-stable.

Lemma 1. *If the j-th column of the stochastic matrix $\boldsymbol{Q}$ is quasi-constant with common off-diagonal element ρ_j and diagonal element $\rho_j + \theta_j$, then $\boldsymbol{M}\mathbf{e_j}' = \theta_j \boldsymbol{L}\mathbf{e_j}'$ and $\boldsymbol{M}_{ij} = \rho_j \theta_j$ whenever $i \neq j$.*

Proof. From definition 2 follows

$$\begin{aligned} \boldsymbol{M}\mathbf{e_j}' = \boldsymbol{L}\boldsymbol{Q}\mathbf{e_j}' &= \rho_j \boldsymbol{L}\mathbf{e}' + \theta_j \boldsymbol{L}\mathbf{e_j}' \\ &= \theta_j \boldsymbol{L}\mathbf{e_j}' \quad (\text{since } \boldsymbol{L}\mathbf{e}' = \mathbf{0}). \end{aligned}$$

Moreover, using $\boldsymbol{L} = \boldsymbol{Q} - \mathbf{I}$,

$$\begin{aligned} \boldsymbol{M}\mathbf{e_j}' &= \theta_j \boldsymbol{L}\mathbf{e_j}' \\ &= \theta_j \boldsymbol{Q}\mathbf{e_j}' - \theta_j \mathbf{e_j}' \\ &= \theta_j(\rho_j \mathbf{e}' + \theta_j \mathbf{e_j}') - \theta_j \mathbf{e_j}' \\ &= \theta_j \rho_j \mathbf{e}' + (\theta_j^2 - \theta_j)\mathbf{e_j}'. \end{aligned}$$

Hence, if $i \neq j$, $\boldsymbol{M}_{ij} = \theta_j \rho_j$. □

Lemma 2. *For all $j \in \{1, \ldots, k\}$, it holds that*

a. if the j-th column of $\boldsymbol{Q}$ is quasi-constant, then $\mathcal{H}_j^{\boldsymbol{L}} = \mathcal{H}_j^{\boldsymbol{M}}$.

b. if $\mathcal{H}_j^{\boldsymbol{L}} = \mathcal{H}_j^{\boldsymbol{M}}$ and the eigenvalue 1 of $\boldsymbol{Q}$ is simple, then the j-th column of $\boldsymbol{Q}$ is quasi-constant.

Proof. a. Suppose the j-th column of $\boldsymbol{Q}$ is quasi-constant. From lemma 1 follows $\boldsymbol{M}\mathbf{e_j}' = \theta_j \boldsymbol{L}\mathbf{e_j}'$ for some scalar $\theta_j \neq 0$. Hence, $\boldsymbol{L}\mathbf{e_j}' = \mathbf{o}'$ if and only if $\boldsymbol{M}\mathbf{e_j}' = \mathbf{o}'$ as $\theta_j \neq 0$. So, either $\mathcal{H}_j^{\boldsymbol{L}} = \mathbb{R}^{1\times k} = \mathcal{H}_j^{\boldsymbol{M}}$ or $\mathcal{H}_j^{\boldsymbol{L}}$ and $\mathcal{H}_j^{\boldsymbol{M}}$ correspond to coinciding hyperplanes of $\mathbb{R}^k$, since $\boldsymbol{L}\mathbf{e_j}'$ and $\boldsymbol{M}\mathbf{e_j}'$ are linearly dependent.

b. Suppose $\mathcal{H}_j^{\boldsymbol{L}} = \mathcal{H}_j^{\boldsymbol{M}}$. Then the j-th column vectors of $\boldsymbol{L}$ and $\boldsymbol{M}$ are both equal to the $k \times 1$ null vector or are linearly dependent. In any case, a scalar $\mu_j \neq 0$ exists such that

$$\boldsymbol{L}\mathbf{e_j}' = \mu_j \boldsymbol{M}\mathbf{e_j}' = \mu_j \boldsymbol{L}\boldsymbol{Q}\mathbf{e_j}'. \tag{6}$$

Now, consider the vector $\boldsymbol{v}_\mathbf{j}' := \mathbf{e_j}' - \mu_j \boldsymbol{Q}\mathbf{e_j}'$. By equation (6), $\boldsymbol{v}_\mathbf{j}'$ is a right eigenvector of $\boldsymbol{Q}$ with eigenvalue one. Since this eigenvalue is assumed to be simple and $\boldsymbol{Q}$ is stochastic, $\boldsymbol{v}_\mathbf{j}' = \tau_j \mathbf{e}'$ for some scalar τ_j. By the definition of $\boldsymbol{v}_\mathbf{j}'$ above, we then have $\tau_j = (\boldsymbol{v}_\mathbf{j}')_i = -\mu_j \boldsymbol{Q}_{ij}$ for all $i \neq j$ and $\tau_j = (\boldsymbol{v}_\mathbf{j}')_j = 1 - \mu_j \boldsymbol{Q}_{jj}$. It then follows immediately that the j-th column of $\boldsymbol{Q}$ is quasi-constant, since $\mu_j \neq 0$.

□

Proposition 2. *If $\mathcal{MP}(\boldsymbol{Q}, 2) = \Pi_k$, then $\boldsymbol{M}$ is a Metzler-matrix.*

Proof. Suppose (1) is true for all $\boldsymbol{v} \in \Pi_k$. Let $i \neq j$ and $\boldsymbol{v} = \mathbf{e_i}$. Then $\boldsymbol{v}_j = 0$, $(\boldsymbol{v}\boldsymbol{L})_j = \boldsymbol{Q}_{ij}$ and $(\boldsymbol{v}\boldsymbol{M})_j = \boldsymbol{M}_{ij}$. If $\boldsymbol{Q}_{ij} > 0$, (1) implies $\boldsymbol{M}_{ij} \geq 0$. If $\boldsymbol{Q}_{ij} = 0$, then $\boldsymbol{M}_{ij} = (\boldsymbol{Q}^2)_{ij} \geq 0$ because $\boldsymbol{Q}$ is non-negative. In any case $\boldsymbol{M}_{ij} \geq 0$ for all $i \neq j$, so $\boldsymbol{M}$ is a Metzler-matrix. □

The converse is not true. Consider for instance the stochastic matrix

$$\boldsymbol{Q} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \end{bmatrix}. \tag{7}$$

A routine calculation reveals that

$$\boldsymbol{L} = \frac{1}{2}\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \boldsymbol{M} = \frac{1}{4}\begin{bmatrix} -1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

so $\boldsymbol{M}$ is a Metzler-matrix. Since $(\boldsymbol{vL})_2 = \frac{1}{2}(\boldsymbol{v}_1 - \boldsymbol{v}_2)$ and $(\boldsymbol{vM})_2 = -\frac{1}{4}\boldsymbol{v}_2$, all probability vectors $\boldsymbol{v} \in \Pi_3$ with $\boldsymbol{v}_1 > \boldsymbol{v}_2 > 0$ are not belonging to $\mathcal{MP}(\boldsymbol{Q}, 2)$.

However, for $n = 2$, the converse holds true. To prove this, we first need a lemma about the relationship between the matrices $\boldsymbol{L}$ and $\boldsymbol{M}$ for the 2-state case.

Lemma 3. *Let $\boldsymbol{Q}$ be stochastic matrix of order 2 with eigenvalues 1 and λ. Then $\boldsymbol{M} = \lambda \boldsymbol{L}$.*

Proof. The Cayley-Hamilton theorem states that every matrix satisfies its characteristic equation, therefore $(\boldsymbol{Q} - \mathbf{I})(\boldsymbol{Q} - \lambda\mathbf{I}) = \mathbf{0}$. Using $\boldsymbol{L} = \boldsymbol{Q} - \mathbf{I}$ and $\boldsymbol{M} = \boldsymbol{LQ}$ we get by expansion $\boldsymbol{M} = \lambda \boldsymbol{L}$. □

Theorem 3. *For a stochastic matrix $\boldsymbol{Q}$ of order 2, $\mathcal{MP}(\boldsymbol{Q}, 2) = \Pi_2$ if and only if $\boldsymbol{M}$ is a Metzler-matrix.*

Proof. By proposition 2, only the sufficiency part remains to be shown. Assume $\boldsymbol{M}$ is a Metzler-matrix. Then, by lemma 3, $\lambda \geq 0$ or $\boldsymbol{L} = \mathbf{0}$ since $\boldsymbol{L} = \boldsymbol{Q} - \mathbf{I}$ is always a Metzler-matrix having zero row-sums. Now, let $\boldsymbol{v} \in \Pi_2$ and $j \in \{1, 2\}$. By lemma 3,

$$(\boldsymbol{vL})_j (\boldsymbol{vM})_j = \lambda (\boldsymbol{vL})_j^2$$

which is non-negative because $\lambda \geq 0$ or $\boldsymbol{L} = \mathbf{0}$. Hence $\boldsymbol{v} \in \mathcal{MP}(\boldsymbol{Q}, 2)$ by (1). □

If $k \geq 3$, necessary and sufficient conditions to have $\mathcal{MP}(\boldsymbol{Q}, 2) = \Pi_k$ are given in the following theorem for the case when $\boldsymbol{Q}$ is primitive :

Theorem 4. *Let $\boldsymbol{Q}$ be a primitive stochastic matrix of order $k \geq 3$. Then, $\mathcal{MP}(\boldsymbol{Q}, 2) = \Pi_k$ if and only if $\boldsymbol{M}$ is a Metzler-matrix and $\boldsymbol{Q}$ is either quasi-stable or $\boldsymbol{Q} = \boldsymbol{e}'\boldsymbol{\alpha}$ where $\boldsymbol{\alpha}$ is the unique stationary distribution of $\boldsymbol{Q}$.*

Proof. We first prove the sufficiency. Let $\boldsymbol{v} \in \Pi_k$ and $j \in \{1, \ldots, k\}$. We show that $(\boldsymbol{vL})_j (\boldsymbol{vM})_j \geq 0$. If $\boldsymbol{Q} = \boldsymbol{e}'\boldsymbol{\alpha}$, then $\boldsymbol{Q}$ is idempotent so $\boldsymbol{M} = \mathbf{0}$ and we are done. If $\boldsymbol{Q}$ is quasi-stable, then by lemma 1, $(\boldsymbol{vM})_j = \theta_j (\boldsymbol{vL})_j$ where $\theta_j = \boldsymbol{Q}_{jj} - \rho_j$ and $\rho_j \in [0, 1]$ is the common off-diagonal value in the jth-column of $\boldsymbol{Q}$. Also by lemma 1, $\boldsymbol{M}_{ij} = \rho_j \theta_j$ for all $i \neq j$, hence $\rho_j \theta_j \geq 0$ since $\boldsymbol{M}$ is a Metzler-matrix. If $\rho_j > 0$, then $\theta_j \geq 0$. If $\rho_j = 0$, then $\theta_j = \boldsymbol{Q}_{jj} \geq 0$. In both cases $\theta_j \geq 0$, hence $(\boldsymbol{vL})_j (\boldsymbol{vM})_j = \theta_j (\boldsymbol{vL})_j^2 \geq 0$.

We now turn to the necessity part. Let $\mathcal{MP}(\boldsymbol{Q}, 2) = \Pi_k$. Then, by proposition 2, $\boldsymbol{M}$ is a Metzler-matrix. If $\boldsymbol{M} = \mathbf{0}$, then $\boldsymbol{Q}$ is idempotent, so that each row of $\boldsymbol{Q}$ is a left eigenvector of $\boldsymbol{Q}$ having unit eigenvalue. By primitivity of $\boldsymbol{Q}$, each row of $\boldsymbol{Q}$ must then be equal to the unique stationary distribution $\boldsymbol{\alpha}$ of $\boldsymbol{Q}$, hence $\boldsymbol{Q} = \boldsymbol{e}'\boldsymbol{\alpha}$.

Now suppose $\boldsymbol{M} \neq \mathbf{0}$. We first show that every non-null column of $\boldsymbol{M}$ coincides with a quasi-constant column of $\boldsymbol{Q}$. Let the index j be such that $\boldsymbol{M}\boldsymbol{e_j}' \neq \mathbf{o}'$. Then $\boldsymbol{L}\boldsymbol{e_j}' \neq \mathbf{o}'$ because $\boldsymbol{M} = \boldsymbol{Q}\boldsymbol{L}$. In that case, $\mathcal{H}_j^{\boldsymbol{L}}$ and $\mathcal{H}_j^{\boldsymbol{M}}$ are both hyperplanes. Suppose they are not coincident. Using the fact that $\boldsymbol{\alpha}$ lies in the interior of Π_k (a consequence of the Perron-Frobenius theorem for primitive matrices), proposition 1 (d) then yields $\mathcal{N}_j \cap \Pi_k \neq \emptyset$, contradicting $\mathcal{MP}(\boldsymbol{Q}, 2) = \Pi_k$. Hence $\mathcal{H}_j^{\boldsymbol{L}} = \mathcal{H}_j^{\boldsymbol{M}}$. Furthermore, the Perron-Frobenius theorem states that the unit eigenvalue of $\boldsymbol{Q}$ is simple. By lemma 2, the j-th column of $\boldsymbol{Q}$ is then quasi-constant.

Finally, we prove that $\boldsymbol{M}$ cannot have a null column given $\boldsymbol{M} \neq \mathbf{0}$. Suppose the j-th column of $\boldsymbol{M}$ is null. Let the l-th column of $\boldsymbol{M}$ be non-null ($l \neq j$). We have proven above that the l-th column of $\boldsymbol{Q}$ is quasi-constant. Let ρ_l be the common off-diagonal element of $\boldsymbol{Q}$ and $\rho_l + \theta_l = \boldsymbol{Q}_{ll}$ with $\theta_l \neq 0$. Then, by lemma 1, $\boldsymbol{M}_{jl} = \rho_l \theta_l$. But $\boldsymbol{M}_{jl} = 0$ since $\boldsymbol{M}_{jj} = 0$ and $\boldsymbol{M}$ is a Metzler-matrix having zero-sum rows. Therefore $\rho_l = 0$, as $\theta_l \neq 0$, and thus $\boldsymbol{Q}\boldsymbol{e_l}' = \theta_l \boldsymbol{e_l}$. But then, $\boldsymbol{Q}$ cannot be primitive because all natural powers of $\boldsymbol{Q}$ contain zero elements at the off-diagonal positions of their l-th column. □

In the next section, we consider the transition matrices $\boldsymbol{Q}$ for which $\boldsymbol{M}$ is a Metzler-matrix in more detail.

4 Stepwise increasing transition matrices

If $\boldsymbol{M} = \boldsymbol{Q}^2 - \boldsymbol{Q}$ is a Metzler-matrix, then $(\boldsymbol{Q}^2)_{ij} \geq \boldsymbol{Q}_{ij}$ for all $i \neq j$, i.e. a transition from one state to another state is more likely to occur over time intervals of double the length of the unit time interval. This is a sensible assumption in a number of applications. For instance, in manpower systems where the states are job-levels and where employees face promotion prospects which depend on seniority (Mills, 1985), a promotion to a higher job-level becomes more likely if the time horizon increases. Also, in credit migration studies, where the states are credit ratings of financial assets, transitions from one credit rating to another rating seem more likely to occur over longer time intervals, see the examples in (Altman & Rijken, 2004).

This motivates the introduction of the notion of "stepwise increasingness".

Definition 3. *A stochastic matrix* $\boldsymbol{Q}$ *is called* stepwise increasing *if and only if for all* $i \neq j$,

$$(\boldsymbol{Q}^2)_{ij} \geq \boldsymbol{Q}_{ij},$$

which is to say that $\boldsymbol{M} = \boldsymbol{Q}^2 - \boldsymbol{Q}$ *is a Metzler-matrix.*

The following results relate stepwise increasingness to the eigenvalues.

Proposition 3. *Let* $\boldsymbol{Q}$ *be a stochastic matrix of order* 2. *Then* $\boldsymbol{Q}$ *is stepwise increasing if and only if* $\boldsymbol{Q}$ *has no negative eigenvalues.*

Proof. By lemma 3, $\boldsymbol{M} = \lambda\boldsymbol{L}$, where the eigenvalues of $\boldsymbol{Q}$ are 1 and λ.

We first prove the sufficiency part. If $\lambda \geq 0$, then $\boldsymbol{M} = \lambda\boldsymbol{L}$ is Metzler because $\boldsymbol{L}$ is. So $\boldsymbol{Q}$ is stepwise increasing by definition.

Conversely, let $\boldsymbol{Q}$ be stepwise increasing and assume $\lambda < 0$. Then $\boldsymbol{L} = \boldsymbol{0}$, because $\boldsymbol{L}$ and $\boldsymbol{M} = \lambda\boldsymbol{L}$ are Metzler-matrices with zero row sums. From $\boldsymbol{L} = \boldsymbol{0}$ follows $\boldsymbol{Q} = \mathbf{I}$. But then $\lambda = 1$, contradicting the assumption $\lambda < 0$. Hence $\lambda \geq 0$. □

When the order of $\boldsymbol{Q}$ exceeds 2, not having a negative eigenvalue does not guarantee stepwise increasingness. For instance, consider the stochastic matrix $\boldsymbol{Q}$ of the form

$$\boldsymbol{Q} = \begin{bmatrix} a & c & 1-a-c \\ 0 & b & 1-b \\ 0 & 0 & 1 \end{bmatrix}, \quad a+b<1,\ 0<c<1-a. \tag{8}$$

Then $\boldsymbol{Q}$ has eigenvalues 1, a and b. In this case, $(\boldsymbol{Q}^2)_{12} = (a+b)c < c = \boldsymbol{Q}_{12}$, so $\boldsymbol{Q}$ is not stepwise increasing.

Not having negative eigenvalues is nevertheless a necessary condition to stepwise increasingness when $\boldsymbol{Q}$ is primitive.

Theorem 5. *Let the stochastic matrix* $\boldsymbol{Q}$ *of order* $k \geq 3$ *be stepwise increasing and primitive. Then every real eigenvalue of* $\boldsymbol{Q}$ *is non-negative.*

Proof. Since $\boldsymbol{Q}$ is stepwise increasing, $\boldsymbol{M} = \boldsymbol{Q}^2 - \boldsymbol{Q}$ is a Metzler-matrix. By a Perron-Frobenius-type theorem for Metzler-matrices (Ngoc, 2006), $\boldsymbol{M}$ has a real eigenvalue μ^* which dominates $\mathrm{Re}(\mu)$ for all $\mu \in \sigma(\boldsymbol{M})$, the set of eigenvalues of the matrix $\boldsymbol{M}$. Now, suppose an eigenvalue $\lambda < 0$ of $\boldsymbol{Q}$ exists. Then, since $\lambda^2 - \lambda \in \sigma(\boldsymbol{M})$, $\mu^* \geq \lambda^2 - \lambda > 0$.

Now there exists a scalar $s > 0$ such that $\boldsymbol{Z} = \boldsymbol{M} + s\mathbf{I}$ is non-negative, $\boldsymbol{M}$ being a Metzler-matrix. Furthermore, by primitivity of $\boldsymbol{Q}$, $\boldsymbol{Q}$ as well as $\boldsymbol{Q}^2$ are irreducible and so $\boldsymbol{Z}$ is irreducible. Hence the Perron-Frobenius theorem for non-negative irreducible matrices applies to $\boldsymbol{Z}$. Let r^* be the Perron-root of $\boldsymbol{Z}$. Then, $\min_i \sum_j z_{ij} \leq r^* \leq \max_i \sum_j z_{ij}$ (see Corollary 1 of theorem 1.1 in (Seneta, 2006)), meaning $r^* = s$ since all row sums of $\boldsymbol{Z}$ are equal to s. But, since $\sigma(\boldsymbol{Z}) = \sigma(\boldsymbol{M}) + \{s\}$, we have $r^* = \mu^* + s$. Consequently $\mu^* = 0$ because $r^* = s$. We now have a contradiction with $\mu^* > 0$. □

The following results characterize stepwise increasingness in terms of the matrix elements.

Theorem 6. *Let $\boldsymbol{Q}$ be a stepwise increasing stochastic matrix of order $k \geq 3$. Then, $\boldsymbol{Q}_{ij} \leq \boldsymbol{Q}_{jj}$ for all indices i and j with $i \neq j$.*

Proof. Let $r \neq j$ be an index such that $\boldsymbol{Q}_{rj} = \max_{s \neq j} \boldsymbol{Q}_{sj}$. We prove that $\boldsymbol{Q}_{rj} \leq \boldsymbol{Q}_{jj}$. Suppose $\boldsymbol{Q}_{rj} > \boldsymbol{Q}_{jj}$. Then,

$$\begin{aligned} 0 \leq (\boldsymbol{Q}^2)_{rj} - \boldsymbol{Q}_{rj} &= \sum_l \boldsymbol{Q}_{rl}\boldsymbol{Q}_{lj} - \boldsymbol{Q}_{rj} \sum_l \boldsymbol{Q}_{rl} \\ &= \sum_{l;\, l \neq r} \boldsymbol{Q}_{rl}(\boldsymbol{Q}_{lj} - \boldsymbol{Q}_{rj}) \\ &= \boldsymbol{Q}_{rj}(\boldsymbol{Q}_{jj} - \boldsymbol{Q}_{rj}) + \sum_{l; l \neq r, l \neq j} \boldsymbol{Q}_{rl}(\boldsymbol{Q}_{lj} - \boldsymbol{Q}_{rj}) \\ &< \sum_{l;\, l \neq r, l \neq j} \boldsymbol{Q}_{rl}(\boldsymbol{Q}_{lj} - \boldsymbol{Q}_{rj}). \end{aligned}$$

Consequently, there exists some $s \neq r$ and $s \neq j$ with $\boldsymbol{Q}_{rs}(\boldsymbol{Q}_{sj} - \boldsymbol{Q}_{rj}) > 0$ and thus $\boldsymbol{Q}_{sj} > \boldsymbol{Q}_{rj}$. This is contradictory to the fact that $\boldsymbol{Q}_{rj} = \max_{s \neq j} \boldsymbol{Q}_{sj}$. □

Theorem 6 gives a necessary condition for a stochastic matrix to be stepwise increasing: all off-diagonal elements must be dominated columnwise by the corresponding diagonal elements. This, however, is not a sufficient condition, as the example (8) with the additional constraint $c \leq b$ demonstrates.

Proposition 2 and theorem 3 connect the notion of a stepwise increasing transition matrix to the concept of a mp-structure. In the next section, we examine the set of mp-structures in more detail for $k = 3$.

5 Monotonically proceeding structures: the three-state case

Let $\boldsymbol{Q}$ be a stepwise increasing stochastic matrix. If $\boldsymbol{Q}$ is of order 2, this property is both necessary and sufficient in order to have all the structures 2-step monotonically proceeding (theorem 3). When the order of $\boldsymbol{Q}$ exceeds two, the stepwise increasing property does not guarantee that all structures are 2-step monotonically proceeding, see proposition 2 and example (7). This section aims to describe the set $\mathcal{MP}(\boldsymbol{Q},2)$ of 2-step mp-structures for the case of order three.

Let us assume that $\boldsymbol{Q}$ is a stepwise increasing matrix of order $k=3$. We also assume that $\boldsymbol{Q}$ is primitive, so that $\lambda=1$ is a simple eigenvalue and there is a unique steady-state probability vector $\boldsymbol{\alpha}\in\mathcal{E}_1$ having no zero components. We denote $\boldsymbol{A}=\lim_{n\to\infty}\boldsymbol{Q}^n=\boldsymbol{e}'\boldsymbol{\alpha}$, where $\boldsymbol{e}$ is the row vector consisting of ones.

We note that the primitivity of $\boldsymbol{Q}$ ensures that no column of $\boldsymbol{L}$ is identically zero. In addition, if $\boldsymbol{Q}$ has no zero eigenvalue, $\boldsymbol{M}$ does not have any zero columns as well, since $\boldsymbol{M}=\boldsymbol{Q}\boldsymbol{L}$.

Suppose $\boldsymbol{Q}$ has a complex conjugate pair of eigenvalues $\{\lambda,\bar{\lambda}\}$ where $\lambda=r\mathrm{e}^{\mathrm{i}\theta}$, $0<r<1$ and $0<\theta<\pi$. By Sylvester's formula, $\boldsymbol{Q}^t=\boldsymbol{A}+\lambda^t\boldsymbol{Q}_1+\bar{\lambda}^t\boldsymbol{Q}_2$ for all $t\in\mathbb{N}$, where $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$ are the Frobenius covariants corresponding to λ and $\bar{\lambda}$, which satisfy $\mathbf{I}=\boldsymbol{A}+\boldsymbol{Q}_1+\boldsymbol{Q}_2$ and $\overline{\boldsymbol{Q}_2}=\boldsymbol{Q}_1$. Since $\lambda^t=r^t\mathrm{e}^{\mathrm{i}t\theta}$, $\bar{\lambda}^t=r^t\mathrm{e}^{-\mathrm{i}t\theta}$ and $\mathrm{e}^{\mathrm{i}t\theta}=\cos t\theta+\mathrm{i}\sin t\theta$, we obtain $\boldsymbol{Q}^t=\boldsymbol{A}+r^t\cos t\theta\boldsymbol{A}_1+r^t\sin t\theta\boldsymbol{A}_2$, where $\boldsymbol{A}_1=\mathbf{I}-\boldsymbol{A}$ and $\boldsymbol{A}_2=\mathrm{i}(\boldsymbol{Q}_1-\boldsymbol{Q}_2)$ are both real matrices. Thus, for any $\boldsymbol{v}\neq\boldsymbol{\alpha}$, each component of $(\boldsymbol{v}\boldsymbol{Q}^t)$ is an oscillating function of t. As we are studying monotonically proceeding structures, we henceforth assume that all eigenvalues are real.

According to theorem 5, all real eigenvalues must be non-negative, so we denote the set $\sigma(\boldsymbol{Q})$ of eigenvalues of $\boldsymbol{Q}$ as $\sigma(\boldsymbol{Q})=\{1,\Lambda,\lambda\}$ with $0\leq\lambda\leq\Lambda<1$.

5.1 Description of the set $\mathcal{MP}(\boldsymbol{Q},2)$

We discuss all possible cases regarding the eigenvalues of $\boldsymbol{Q}$.

In case $0=\lambda=\Lambda$, then $\operatorname{trace}(\boldsymbol{Q})=1=\operatorname{trace}(\boldsymbol{Q}^2)$, consequently $\operatorname{trace}(\boldsymbol{M})=0$. Since $\boldsymbol{M}$ has zero row sums and is assumed to be a Metzler-matrix ($\boldsymbol{Q}$ is stepwise increasing), we necessarily have $\boldsymbol{M}=\mathbf{0}$. So (1) is automatically satisfied for any structure $\boldsymbol{v}\in\Pi_3$. Hence $\mathcal{MP}(\boldsymbol{Q},2)=\Pi_3$.

In case $0=\lambda<\Lambda$, then, by Sylvester's formula (see (Gantmacher, 1960), p. 101), for each vector $\boldsymbol{v}\in\Pi_3$ there exists a vector $\boldsymbol{u}$ such that $\boldsymbol{v}\boldsymbol{Q}^n=\boldsymbol{\alpha}+\Lambda^n\boldsymbol{u}$ for all $n\in\mathbb{N}_0$. Hence, the components of $\boldsymbol{v}\boldsymbol{Q}^n$ are monotonic functions of n and therefore $\mathcal{MP}(\boldsymbol{Q},t)=\Pi_3$ for all $t\geq 2$. In particular, $\mathcal{MP}(\boldsymbol{Q},2)=\Pi_3$.

In case $0<\lambda=\Lambda<1$ and $\boldsymbol{Q}$ is diagonalizable, then $(\boldsymbol{Q}-\mathbf{I})(\boldsymbol{Q}-\lambda\mathbf{I})=\mathbf{0}$. Using (2), this equation can be rewritten as $\boldsymbol{M}=\lambda\boldsymbol{L}$. Hence $(\boldsymbol{v}\boldsymbol{L})_i(\boldsymbol{v}\boldsymbol{M})_i=\lambda(\boldsymbol{v}\boldsymbol{L})_i^2\geq 0$ for all $\boldsymbol{v}\in\Pi_3$

and $i \in \{1, 2, 3\}$. So, by (1), $\mathcal{MP}(\boldsymbol{Q}, 2) = \Pi_3$.

All the previous cases result in $\mathcal{MP}(\boldsymbol{Q}, 2) = \Pi_3$. In all these situations, $\boldsymbol{Q}$ is quasi-stable (theorem 4).

The case $0 < \lambda = \Lambda < 1$ where $\boldsymbol{Q}$ is not diagonalizable, turns out to a limiting situation, where $\lambda \to \Lambda$, of the case $0 < \lambda < \Lambda < 1$ (described below) and will be discussed at the end of this section.

We now turn to the case $0 < \lambda < \Lambda < 1$. According to (4), the set $\mathcal{MP}(\boldsymbol{Q}, 2)$ is the complement in Π_3 of $\cup_{j=1}^3 \mathcal{N}_j$, where

$$\mathcal{N}_j = \{\boldsymbol{x} \in \mathbb{R}^{1\times 3} \,|\, (\boldsymbol{xL})_j(\boldsymbol{xM})_j < 0\}.$$

Using proposition 1 (b and c) and theorem 2, the location of the sets $\mathcal{N}_j \cap \Pi_3$ within the simplex with vertices $\mathbf{e_1}$, $\mathbf{e_2}$, $\mathbf{e_3}$ (which is a representation of Π_3) is such that $\mathcal{MP}(\boldsymbol{Q}, 2) = \cup_{j=1}^3 \mathcal{C}_j$, the union of three “bowtie-shaped” subsets $\mathcal{C}_j$ which are star-convex with respect to $\boldsymbol{\alpha}$. The labeling is defined in such a way that $\mathcal{C}_j$ contains $\mathbf{e_j}$ and is bordered by $\mathcal{N}_i \cup \mathcal{N}_l$ where $\{i, j, l\} = \{1, 2, 3\}$, see figure 1. We remark that $\mathrm{Int}(\mathcal{C}_j)$, the interior of $\mathcal{C}_j$ within Π_3, is precisely that subset of elements $\boldsymbol{v} \in \mathcal{MP}(\boldsymbol{Q}, 2)$ for which $(\boldsymbol{vL})_i$ and $(\boldsymbol{vL})_l$ have the same sign, i.e.

$$\mathrm{Int}(\mathcal{C}_j) = \{\boldsymbol{v} \in \mathcal{MP}(\boldsymbol{Q}, 2) \,|\, (\boldsymbol{vL})_i(\boldsymbol{vL})_l > 0\}. \tag{9}$$

Note that $\mathcal{C}_j$ is the union of two closed cones with vertex $\boldsymbol{\alpha}$ that are point-symmetric with respect to $\boldsymbol{\alpha}$. Moreover,

$$\mathrm{Int}(\mathcal{C}_i) \cap \mathrm{Int}(\mathcal{C}_j) = \emptyset, \quad \text{whenever } i \neq j. \tag{10}$$

For each $\mu \in \{\lambda, \Lambda\}$, we also consider the following sets of structures

$$\mathcal{S}_\mu = \{\boldsymbol{v} \in \Pi_3 \,|\, \boldsymbol{v} - \boldsymbol{\alpha} \in \mathcal{E}_\mu\} = \Pi_3 \cap (\mathcal{E}_\mu + \boldsymbol{\alpha}).$$

All vectors $\boldsymbol{u} \in \mathcal{E}_\mu$ satisfy $\boldsymbol{u}(\boldsymbol{Q} - \mu\mathbf{I}) = \mathbf{o}$. Upon postmultiplying both sides of this equation by $\boldsymbol{e}'$, we obtain $(1 - \mu)\boldsymbol{ue}' = 0$ since $\boldsymbol{Q}$ has unit row sums. Hence,

$$\boldsymbol{ue}' = \sum_{i=1}^{3} \boldsymbol{u}_i = 0, \tag{11}$$

since $\mu \neq 1$. Therefore, $\mathcal{S}_\mu$ are line segments in Π_3 containing $\boldsymbol{\alpha}$. Furthermore, $\mathcal{S}_\mu \subset \mathcal{MP}(\boldsymbol{Q}, 2)$ by proposition 1(e). Hence $\mathcal{S}_\mu$ must be contained in one of the subsets $\mathcal{C}_j$. The following lemma excludes the boundary of such a subset as a possible location of $\mathcal{S}_\mu$ when $\boldsymbol{Q}$ has no quasi-constant columns.

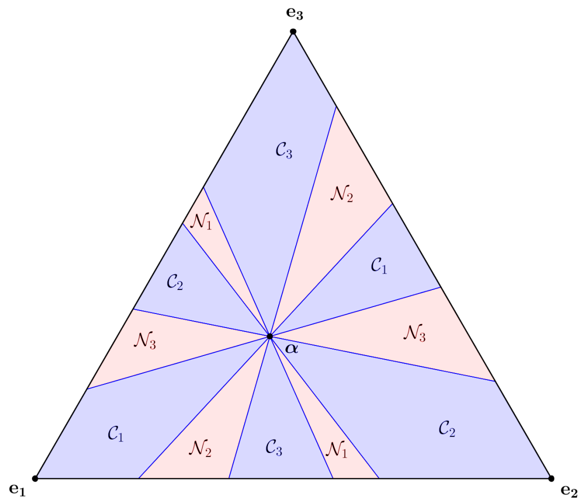

Figure 1: Graphical representation of $\mathcal{MP}(\boldsymbol{Q}, 2) = \cup_{j=1}^{3} \mathcal{C}_j$

Lemma 4. *Let $\mu \in \{\lambda, \Lambda\}$. If the j-th column of $\boldsymbol{Q}$ is not quasi-constant, then $\mathcal{E}_\mu + \boldsymbol{\alpha} \not\subset \mathcal{H}_j^{\boldsymbol{L}}$ and $\mathcal{E}_\mu + \boldsymbol{\alpha} \not\subset \mathcal{H}_j^{\boldsymbol{M}}$.*

Proof. Let $\boldsymbol{v} \in \mathcal{E}_\mu + \boldsymbol{\alpha}$. Then, $(\boldsymbol{v} - \boldsymbol{\alpha})\boldsymbol{Q} = \mu(\boldsymbol{v} - \boldsymbol{\alpha})$. Postmultiplying by $\boldsymbol{L}$ and using $\boldsymbol{M} = \boldsymbol{Q}\boldsymbol{L}$ and $\boldsymbol{\alpha}\boldsymbol{L} = \mathbf{o}$, we have $\boldsymbol{v}\boldsymbol{M} = \mu\boldsymbol{v}\boldsymbol{L}$. As $\mu \neq 0$, it then follows that $\boldsymbol{v} \in \mathcal{H}_j^{\boldsymbol{M}}$ if and only if $\boldsymbol{v} \in \mathcal{H}_j^{\boldsymbol{L}}$ and hence, $\mathcal{E}_\mu + \boldsymbol{\alpha} \subset \mathcal{H}_j^{\boldsymbol{L}}$ if and only if $\mathcal{E}_\mu + \boldsymbol{\alpha} \subset \mathcal{H}_j^{\boldsymbol{M}}$.

Suppose $\mathcal{E}_\mu + \boldsymbol{\alpha} \subset \mathcal{H}_j^{\boldsymbol{L}}$ or $\mathcal{E}_\mu + \boldsymbol{\alpha} \subset \mathcal{H}_j^{\boldsymbol{M}}$. By the above argument, the hyperplanes $\mathcal{H}_j^{\boldsymbol{L}}$ and $\mathcal{H}_j^{\boldsymbol{M}}$ are then coincident since they both contain the line $\mathcal{E}_\mu + \boldsymbol{\alpha}$ and the vector $\mathbf{o}$ which lies outside $\mathcal{E}_\mu + \boldsymbol{\alpha}$ as $\mu \neq 1$. But by lemma 2 (b), the j-th column of $\boldsymbol{Q}$ would then be quasi-constant. □

To summarize, if no column of $\boldsymbol{Q}$ is quasi-constant, $\mathcal{S}_\mu \setminus \{\boldsymbol{\alpha}\} \subset \mathrm{Int}(\mathcal{C}_j)$ for exactly one $j \in \{1, 2, 3\}$. We denote this set $\mathcal{C}_j$ as $\mathcal{C}(\mu)$. In the next theorem, we show that the index j is given by the column index for which μ lies between the smallest en largest absolute differences between the off-diagonal elements and the diagonal element. Let us therefore denote

$$\underline{\boldsymbol{Q}}_j := \min\{\boldsymbol{Q}_{ij}, \boldsymbol{Q}_{lj}\} \quad \text{and} \quad \overline{\boldsymbol{Q}}_j := \max\{\boldsymbol{Q}_{ij}, \boldsymbol{Q}_{lj}\} \quad \text{where } \{j, i, l\} = \{1, 2, 3\}.$$

Theorem 7. *Let $\boldsymbol{Q}$ be a stepwise increasing and primitive stochastic matrix of order 3 having spectrum $\sigma(\boldsymbol{Q}) = \{1, \Lambda, \lambda\}$ where $0 < \lambda < \Lambda < 1$. Suppose $\boldsymbol{Q}$ has no quasi-constant columns.*

Then, for all $\mu \in \{\lambda, \Lambda\}$ and $j \in \{1,2,3\}$,

$$\mathcal{C}(\mu) = \mathcal{C}_j \quad \Leftrightarrow \quad \boldsymbol{Q}_{jj} - \overline{\boldsymbol{Q}}_j < \mu < \boldsymbol{Q}_{jj} - \underline{\boldsymbol{Q}}_j. \tag{12}$$

Proof. Let $\{j, i, l\} = \{1,2,3\}$, $\mu \in \{\lambda, \Lambda\}$ and $\boldsymbol{u} \in \mathcal{E}_\mu \setminus \{\mathbf{o}\}$. Let $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{\alpha}$. Then, using (2), $\boldsymbol{vL} = (\mu - 1)\boldsymbol{u}$ and, because $\mu \neq 1$,

$$(\boldsymbol{vL})_i(\boldsymbol{vL})_l > 0 \;\Leftrightarrow\; \boldsymbol{u}_i\boldsymbol{u}_l > 0. \tag{13}$$

Moreover, $\boldsymbol{u} \in \mathcal{E}_\mu$ implies

$$\boldsymbol{u}_j(\boldsymbol{Q}_{jj} - \mu) + \boldsymbol{u}_i\boldsymbol{Q}_{ij} + \boldsymbol{u}_l\boldsymbol{Q}_{lj} = 0,$$

which can be rewritten as

$$-(\boldsymbol{Q}_{jj} - \mu - \boldsymbol{Q}_{ij})\boldsymbol{u}_i = (\boldsymbol{Q}_{jj} - \mu - \boldsymbol{Q}_{lj})\boldsymbol{u}_l \tag{14}$$

using (11).

We are now ready to prove the necessity part. Suppose $\mathcal{C}(\mu) = \mathcal{C}_j$, then, by definition, $\mathcal{S}_\mu \setminus \{\boldsymbol{\alpha}\} \subset \mathrm{Int}(\mathcal{C}_j)$. Let $\boldsymbol{v} \in \mathcal{S}_\mu \setminus \{\boldsymbol{\alpha}\}$, i.e. $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{\alpha}$ where $\boldsymbol{u} \in \mathcal{E}_\mu \setminus \{\mathbf{o}\}$. Then, $\boldsymbol{v} \in \mathrm{Int}(\mathcal{C}_j)$ yielding $(\boldsymbol{vL})_i(\boldsymbol{vL})_l > 0$. Using (13) and (14), we then have

$$(\boldsymbol{Q}_{jj} - \mu - \boldsymbol{Q}_{ij})(\boldsymbol{Q}_{jj} - \mu - \boldsymbol{Q}_{lj}) < 0,$$

because none of these factors can be zero assuming the j-th column of $\boldsymbol{Q}$ is not quasi-constant. The conclusion is now follows from the fact that $\{\boldsymbol{Q}_{ij}, \boldsymbol{Q}_{lj}\} = \{\underline{\boldsymbol{Q}}_j, \overline{\boldsymbol{Q}}_j\}$.

We now turn to the sufficiency part. Let $\boldsymbol{v} \in \mathcal{S}_\mu \setminus \{\boldsymbol{\alpha}\}$, i.e. $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{\alpha}$ where $\boldsymbol{u} \in \mathcal{E}_\mu \setminus \{\mathbf{o}\}$. We must prove that $\boldsymbol{v} \in \mathrm{Int}(\mathcal{C}_j)$, which is equivalent to $\boldsymbol{u}_i\boldsymbol{u}_l > 0$, by (13). Now, the two bracketed expressions in (14) are both non-zero and have opposite signs by assumption. For (14) to hold, one must necessarily have $\boldsymbol{u}_i\boldsymbol{u}_l > 0$, since $\boldsymbol{u} \neq \mathbf{o}$ and hence, by (11), $\boldsymbol{u}_i$ and $\boldsymbol{u}_l$ cannot both be equal to zero. □

Theorem 7 does not exclude the possibility that $\mathcal{S}_\lambda$ and $\mathcal{S}_\Lambda$ lie in the same set $\mathcal{C}_j$, i.e. $\mathcal{C}(\lambda) = \mathcal{C}(\Lambda)$. Conditions under which $\mathcal{C}(\lambda) \neq \mathcal{C}(\Lambda)$ are described in the following corollary.

Corollary 1. *Let $\boldsymbol{Q}$ be a stepwise increasing and primitive stochastic matrix of order 3 having spectrum $\sigma(\boldsymbol{Q}) = \{1, \Lambda, \lambda\}$ where $0 < \lambda < \Lambda < 1$. Suppose $\boldsymbol{Q}$ has no quasi-constant columns. Then,*

$$\begin{aligned} \mathcal{C}(\lambda) \neq \mathcal{C}(\Lambda) \quad &\Leftrightarrow \quad \boldsymbol{Q}_{jj} - \overline{\boldsymbol{Q}}_j < \lambda < \boldsymbol{Q}_{jj} - \underline{\boldsymbol{Q}}_j < \Lambda \quad \textit{for some } j \in \{1,2,3\} \\ &\Leftrightarrow \quad \lambda < \boldsymbol{Q}_{ll} - \overline{\boldsymbol{Q}}_l < \Lambda < \boldsymbol{Q}_{ll} - \underline{\boldsymbol{Q}}_l \quad \textit{for some } l \in \{1,2,3\} \end{aligned}$$

Proof. By virtue of theorem 7,

$$\mathcal{C}(\lambda) = \mathcal{C}_j \neq \mathcal{C}(\Lambda) \quad \Leftrightarrow \quad \boldsymbol{Q}_{jj} - \overline{\boldsymbol{Q}}_j < \lambda < \boldsymbol{Q}_{jj} - \underline{\boldsymbol{Q}}_j < \Lambda$$

and

$$\mathcal{C}(\Lambda) = \mathcal{C}_l \neq \mathcal{C}(\lambda) \quad \Leftrightarrow \quad \lambda < \boldsymbol{Q}_{ll} - \overline{\boldsymbol{Q}}_l < \Lambda < \boldsymbol{Q}_{ll} - \underline{\boldsymbol{Q}}_l \ ,$$

since $\lambda < \Lambda$. □

We can now consider the remaining case $0 < \lambda = \Lambda < 1$ where $\boldsymbol{Q}$ is not diagonalizable. As the eigenspace $\mathcal{E}_\lambda = \mathcal{E}_\Lambda$ is one-dimensional, the line segments $\mathcal{S}_\lambda$ and $\mathcal{S}_\Lambda$ are coincident. In fact, this case arises as a limiting case of $0 < \lambda < \Lambda < 1$ where $\lambda \to \Lambda$. The description of the set $\mathcal{MP}(\boldsymbol{Q}, 2) = \mathcal{C}_1 \cup \mathcal{C}_2 \cup \mathcal{C}_3$ using (9) and theorem 7 continues to hold with the additional constraint $\lambda = \Lambda$.

The following section is concerned with the question of which 2-step mp-structures $\boldsymbol{v}$ will have a successor $\boldsymbol{vQ}$ that is also 2-step monotonically proceeding.

5.2 Location of the succeeding structure with respect to $\mathcal{MP}(\boldsymbol{Q}, 2)$

Suppose $\boldsymbol{v}$ is a t-step mp-structure ($t \geq 3$). By theorem 1, it is then necessary that all intermediary structures $\boldsymbol{u} = \boldsymbol{vQ}^\tau$ ($\tau \in \mathbb{N}$, $\tau \leq t-3$), have the property that both $\boldsymbol{u}$ and $\boldsymbol{uQ}$ are 2-step monotonically proceeding. It is therefore important to examine the set

$$\{\boldsymbol{u} \in \mathcal{MP}(\boldsymbol{Q}, 2) \,|\, \boldsymbol{uQ} \in \mathcal{MP}(\boldsymbol{Q}, 2)\}.$$

In section 5.1, we have seen that the above set is equal to Π_3, except for the cases where $\boldsymbol{Q}$ has distinct eigenvalues or $\boldsymbol{Q}$ has a pair of coinciding eigenvalues and is non-diagonalizable. As the latter case can be viewed as a limiting case of the former, we will restrict our attention to stepwise increasing and primitive stochastic matrices $\boldsymbol{Q}$ having $\sigma(\boldsymbol{Q}) = \{1, \lambda, \Lambda\}$ with $0 < \lambda < \Lambda < 1$.

Our main result in this section is presented in theorem 8. It states that, in case $\mathcal{C}(\lambda) \neq \mathcal{C}(\Lambda)$, all structures $\boldsymbol{v}$ of $\mathcal{C}(\Lambda)$ have their successors $\boldsymbol{vQ}$ belonging to the same set $\mathcal{C}(\Lambda)$ and thus to $\mathcal{MP}(\boldsymbol{Q}, 2)$. To prove this, we first need the following lemma and hereby define for any $\boldsymbol{v} \in \Pi_3$ and $\theta \in \mathbb{R}$, $\theta \neq 1$, the vector $\boldsymbol{v}(\theta)$ as

$$\boldsymbol{v}(\theta) := \frac{1}{1-\theta}\boldsymbol{v}(\boldsymbol{Q} - \theta\mathbf{I}).$$

Note that $\boldsymbol{v}(\theta) \neq \boldsymbol{v}$ if $\boldsymbol{vQ} \neq \boldsymbol{v}$, i.e. if $\boldsymbol{v} \neq \boldsymbol{\alpha}$ ($\boldsymbol{Q}$ is primitive).

Lemma 5. *For all $\boldsymbol{v} \in \Pi_3$, $\boldsymbol{v} \neq \boldsymbol{\alpha}$, it holds that*

a. $\boldsymbol{vQ} \in]\boldsymbol{v}, \boldsymbol{v}(\lambda)[$ *and* $\boldsymbol{vQ} \in]\boldsymbol{v}, \boldsymbol{v}(\Lambda)[$

b. $\boldsymbol{v}(\lambda) \in]\boldsymbol{v}, \boldsymbol{v}(\Lambda)[$

c. $\boldsymbol{v}(\lambda) \in \boldsymbol{\alpha} + \mathcal{E}_\Lambda$ *and* $\boldsymbol{v}(\Lambda) \in \boldsymbol{\alpha} + \mathcal{E}_\lambda$

Proof. a. For $\boldsymbol{v} \in \Pi_3$ and $0 < \theta < 1$ holds

$$\boldsymbol{v}(\theta) = \frac{1}{1-\theta}\boldsymbol{v}(\boldsymbol{Q} - \theta\mathbf{I}) \quad \Leftrightarrow \quad \boldsymbol{vQ} = \theta\boldsymbol{v} + (1-\theta)\boldsymbol{v}(\theta)$$

Consequently $\boldsymbol{vQ} \in]\boldsymbol{v}, \boldsymbol{v}(\theta)[$, the open line segment with endpoints $\boldsymbol{v}$ and $\boldsymbol{v}(\theta)$.

b. The vector $\boldsymbol{v}(\lambda)$ can be rewritten as follows

$$\begin{aligned}\boldsymbol{v}(\lambda) &= \frac{1}{1-\lambda}\boldsymbol{v}(\boldsymbol{Q} - \lambda\mathbf{I}) \\ &= \frac{1}{1-\lambda}\boldsymbol{v}(\boldsymbol{Q} - \Lambda\mathbf{I} + \Lambda\mathbf{I} - \lambda\mathbf{I}) \\ &= \frac{1}{1-\lambda}\boldsymbol{v}(\boldsymbol{Q} - \Lambda\mathbf{I}) + \frac{\Lambda-\lambda}{1-\lambda}\boldsymbol{v} \\ &= \frac{1-\Lambda}{1-\lambda}\boldsymbol{v}(\Lambda) + \frac{\Lambda-\lambda}{1-\lambda}\boldsymbol{v}\end{aligned}$$

with coefficients $\frac{1-\Lambda}{1-\lambda}$ and $\frac{\Lambda-\lambda}{1-\lambda}$ that are both non-negative and sum to unity. Consequently $\boldsymbol{v}(\lambda)$ belongs to the open line segment with endpoints $\boldsymbol{v}$ and $\boldsymbol{v}(\Lambda)$.

c. First, note that $\boldsymbol{\alpha}(\lambda) = \boldsymbol{\alpha}$ since $\boldsymbol{\alpha Q} = \boldsymbol{\alpha}$. Let $\boldsymbol{A} = \lim_{n\to\infty} \boldsymbol{Q}^n$, then $\boldsymbol{A} = \frac{1}{(1-\lambda)(1-\Lambda)}(\boldsymbol{Q} - \lambda\mathbf{I})(\boldsymbol{Q} - \Lambda\mathbf{I})$. Hence,

$$\begin{aligned}(\boldsymbol{v}(\lambda) - \boldsymbol{\alpha})(\boldsymbol{Q} - \Lambda\mathbf{I}) &= (\boldsymbol{v}(\lambda) - \boldsymbol{\alpha}(\lambda))(\boldsymbol{Q} - \Lambda\mathbf{I}) \\ &= \frac{1}{1-\lambda}(\boldsymbol{v} - \boldsymbol{\alpha})(\boldsymbol{Q} - \lambda\mathbf{I})(\boldsymbol{Q} - \Lambda\mathbf{I}) \\ &= (1-\Lambda)(\boldsymbol{v} - \boldsymbol{\alpha})\boldsymbol{A} \\ &= (1-\Lambda)(\boldsymbol{\alpha} - \boldsymbol{\alpha}) = \mathbf{o},\end{aligned}$$

i.e. $\boldsymbol{v}(\lambda) - \boldsymbol{\alpha} \in \mathcal{E}_\Lambda$. The result $\boldsymbol{v}(\Lambda) - \boldsymbol{\alpha} \in \mathcal{E}_\lambda$ can be proven analogously. □

Theorem 8. *Let $\boldsymbol{Q}$ be a stepwise increasing and primitive stochastic matrix of order* 3 *having spectrum $\sigma(\boldsymbol{Q}) = \{1, \Lambda, \lambda\}$ where $0 < \lambda < \Lambda < 1$. Suppose $\boldsymbol{Q}$ has no quasi-constant columns and that $\mathcal{C}(\lambda) \neq \mathcal{C}(\Lambda)$. Then for all $\boldsymbol{v} \in \mathcal{C}(\Lambda)$, $\boldsymbol{v} \neq \boldsymbol{\alpha}$, we have $\boldsymbol{vQ} \in \mathrm{Int}(\mathcal{C}(\Lambda))$.*

Proof. Let $\boldsymbol{v} \in \mathcal{C}(\Lambda)$ and $\boldsymbol{v} \neq \boldsymbol{\alpha}$. According to lemma 5, the vectors $\boldsymbol{v}$, $\boldsymbol{vQ}$, $\boldsymbol{v}(\lambda)$ and $\boldsymbol{v}(\Lambda)$ are collinear in such a way that

$$\boldsymbol{vQ} \in]\boldsymbol{v}, \boldsymbol{v}(\lambda)[\subset]\boldsymbol{v}, \boldsymbol{v}(\Lambda)[\tag{15}$$

By definition, $\mathcal{C}(\Lambda) = \Pi_3 \cap (\mathcal{C}' \cup \mathcal{C}'')$, where $\mathcal{C}'$ and $\mathcal{C}''$ are convex cones that are point-symmetric with respect to their common vertex $\boldsymbol{\alpha}$ (see also figure 1). Since $\boldsymbol{v} \neq \boldsymbol{\alpha}$, $\boldsymbol{v}$ belongs to exactly one of these cones, say $\boldsymbol{v} \in \mathcal{C}'$. Applying lemma 5 (c), $\boldsymbol{v}(\lambda) \in \mathcal{C}' \cup \mathcal{C}''$. We now prove that $\boldsymbol{v}(\lambda) \in \mathcal{C}'$. Suppose, on the contrary, that $\boldsymbol{v}(\lambda) \in \mathcal{C}''$. Then, as $\mathcal{C}(\lambda) \neq \mathcal{C}(\Lambda)$, the open line segment $]\boldsymbol{v}, \boldsymbol{v}(\lambda)[$ must cross $\mathcal{C}(\lambda)$ and therefore it intersects the line $\boldsymbol{\alpha} + \mathcal{E}_\lambda$ containing $\boldsymbol{v}(\Lambda)$, see lemma 5 (c). In this way $\boldsymbol{v}(\Lambda) \in]\boldsymbol{v}, \boldsymbol{v}(\lambda)[$, which contradicts (15).

Now, as $\boldsymbol{v}$ and $\boldsymbol{v}(\lambda)$ are elements of the same convex cone $\mathcal{C}'$, (15) entails $\boldsymbol{vQ} \in \mathcal{C}'$. In fact, $\boldsymbol{v}(\lambda) \in \mathrm{Int}(\mathcal{C}')$ since by lemma 4 the line $\mathcal{E}_\Lambda + \boldsymbol{\alpha}$ containing $\boldsymbol{v}(\lambda)$ does not coincide with the boundary of $\mathcal{C}(\Lambda)$. Hence $\boldsymbol{vQ} \in \mathrm{Int}(\mathcal{C}')$ and thus $\boldsymbol{vQ} \in \mathrm{Int}(\mathcal{C}(\Lambda))$. □

5.3 Long-run evolution of a structure

It follows from theorem 1 that structures $\boldsymbol{v}$ having the entire path $(\boldsymbol{vQ}^t)_{t\in\mathbb{N}}$ in $\mathcal{MP}(\boldsymbol{Q}, 2)$ and satisfying $\boldsymbol{vQ}^{t+1} \neq \boldsymbol{vQ}^{t+2}$ componentwise for all $t \in \mathbb{N}$ will belong to $\mathcal{MP}(\boldsymbol{Q}, \infty) := \cup_{\tau\geq 2}\mathcal{MP}(\boldsymbol{Q}, \tau)$. In other words, structures $\boldsymbol{v}$ will be monotonically proceeding over any number of steps if all structures in their path $(\boldsymbol{vQ}^t)_{t\in\mathbb{N}}$ are monotonically proceeding over two steps and all pairs of consecutive structures in that path differ componentwise. Note that the condition $(\boldsymbol{vQ}^{t+1})_j \neq (\boldsymbol{vQ}^{t+2})_j$ is equivalent to $\boldsymbol{vQ}^t \notin \mathcal{H}_j^{\boldsymbol{M}}$.

As in the previous section, we will only consider stepwise increasing and primitive stochastic matrices $\boldsymbol{Q}$ having $\sigma(\boldsymbol{Q}) = \{1, \lambda, \Lambda\}$ and $0 < \lambda < \Lambda < 1$. The main result of this section is that, in case $\mathcal{C}(\lambda) \neq \mathcal{C}(\Lambda)$, all structures in the interior of $\mathcal{C}(\Lambda)$ or in $\mathcal{S}_\lambda$ are monotonically proceeding over any number of steps. Other structures $\boldsymbol{v}$ will have a $\boldsymbol{vQ}^t$ that belongs to $\mathcal{MP}(\boldsymbol{Q}, \infty)$ for large enough t.

The proof of this result relies on the following lemma, which describes the path $(\boldsymbol{vQ}^t)_t$ towards the stationary structure $\boldsymbol{\alpha}$.

Lemma 6. *Let $\boldsymbol{v} - \boldsymbol{\alpha} \notin \mathcal{E}_\lambda$ and define the vector $\boldsymbol{u}(t) = \boldsymbol{vQ}^{t+1} - \boldsymbol{vQ}^t$, $t \in \mathbb{N}$. Then, for any matrix norm $||\cdot||$, $\lim_{t\to\infty} \frac{\boldsymbol{u}(t)}{||\boldsymbol{u}(t)||}$ exists and is an element of $\mathcal{E}_\Lambda$.*

Proof. Using Sylvester's formula, $\boldsymbol{Q}^t = \boldsymbol{A} + \Lambda^t \boldsymbol{Q}_\Lambda + \lambda^t \boldsymbol{Q}_\lambda$ for all $t \in \mathbb{N}$, where $\boldsymbol{Q}^0 := \mathbf{I}$ and $\boldsymbol{Q}_\Lambda := \frac{1}{(\Lambda-1)(\Lambda-\lambda)}(\boldsymbol{Q} - \mathbf{I})(\boldsymbol{Q} - \lambda\mathbf{I})$ and $\boldsymbol{Q}_\lambda := \frac{1}{(\lambda-1)(\lambda-\Lambda)}(\boldsymbol{Q} - \mathbf{I})(\boldsymbol{Q} - \Lambda\mathbf{I})$ are the Frobenius covariants. Hence

$$\begin{aligned}\boldsymbol{u}(t) &= (\Lambda^{t+1} - \Lambda^t)\boldsymbol{v}\boldsymbol{Q}_\Lambda + (\lambda^{t+1} - \lambda^t)\boldsymbol{v}\boldsymbol{Q}_\lambda \\ &= \Lambda^t(\Lambda - 1)\left(\boldsymbol{v}\boldsymbol{Q}_\Lambda + \left(\frac{\lambda}{\Lambda}\right)^t \frac{\lambda - 1}{\Lambda - 1}\,\boldsymbol{v}\boldsymbol{Q}_\lambda\right)\end{aligned} \tag{16}$$

Now, if $\boldsymbol{v}-\boldsymbol{\alpha} \notin \mathcal{E}_\lambda$, then $\boldsymbol{v}\boldsymbol{Q}_\Lambda \neq \mathbf{o}$. Indeed, should $\boldsymbol{v}\boldsymbol{Q}_\Lambda = \mathbf{o}$, then $\boldsymbol{v}-\boldsymbol{\alpha} = \boldsymbol{v}\boldsymbol{Q}_\lambda$ (because $\mathbf{I} = \boldsymbol{A} + \boldsymbol{Q}_\Lambda + \boldsymbol{Q}_\lambda$) and $\boldsymbol{v}\boldsymbol{Q} - \boldsymbol{\alpha} = \lambda\boldsymbol{v}\boldsymbol{Q}_\lambda$ (because $\boldsymbol{Q} = \boldsymbol{A} + \Lambda\boldsymbol{Q}_\Lambda + \lambda\boldsymbol{Q}_\lambda$). Combining these equations and using $\boldsymbol{\alpha} = \boldsymbol{\alpha}\boldsymbol{Q}$, we have then $(\boldsymbol{v}-\boldsymbol{\alpha})\boldsymbol{Q} = \lambda(\boldsymbol{v}-\boldsymbol{\alpha})$ and thus $\boldsymbol{v}-\boldsymbol{\alpha} \in \mathcal{E}_\lambda$.

Let $||\cdot||$ be any matrix norm. Note that $||\boldsymbol{v}\boldsymbol{Q}_\Lambda|| \neq 0$, since $\boldsymbol{v}\boldsymbol{Q}_\Lambda \neq \mathbf{o}$. Consequently, by (16) and the fact that $(\frac{\lambda}{\Lambda})^t \to 0$ as $t \to \infty$,

$$\frac{\boldsymbol{u}(t)}{||\boldsymbol{u}(t)||} \to -\frac{\boldsymbol{v}\boldsymbol{Q}_\Lambda}{||\boldsymbol{v}\boldsymbol{Q}_\Lambda||} \quad \text{as } t \to \infty.$$

The conclusion now follows from the the Cayley-Hamilton theorem yielding $\boldsymbol{Q}_\Lambda(\boldsymbol{Q}-\lambda\mathbf{I}) = \mathbf{0}$ and hence $\boldsymbol{v}\boldsymbol{Q}_\Lambda(\boldsymbol{Q}-\lambda\mathbf{I}) = \mathbf{o}$.

□

Theorem 9. *Let $\boldsymbol{Q}$ be a stepwise increasing and primitive stochastic matrix of order 3 having spectrum $\sigma(\boldsymbol{Q}) = \{1, \Lambda, \lambda\}$ where $0 < \lambda < \Lambda < 1$. Suppose $\boldsymbol{Q}$ has no quasi-constant columns and $\mathcal{C}(\Lambda) \neq \mathcal{C}(\lambda)$. Then, $\boldsymbol{v} \in \mathcal{MP}(\boldsymbol{Q}, \infty)$ for any structure $\boldsymbol{v} \in \mathrm{Int}(\mathcal{C}(\Lambda)) \cup \mathcal{S}_\lambda$. Moreover, if $\boldsymbol{v} \notin \mathrm{Int}(\mathcal{C}(\Lambda)) \cup \mathcal{S}_\lambda$, there exists $T(\boldsymbol{v}) \in \mathbb{N}_0$ such that $\boldsymbol{v}\boldsymbol{Q}^{T(\boldsymbol{v})} \in \mathcal{MP}(\boldsymbol{Q}, \infty)$.*

Proof. If $\boldsymbol{v} \in \mathrm{Int}(\mathcal{C}(\Lambda))$ or $\boldsymbol{v} \in \mathcal{S}_\lambda$, we prove that, for all $t \in \mathbb{N}$ and $j \in \{1,2,3\}$, $\boldsymbol{v}\boldsymbol{Q}^t \in \mathcal{MP}(\boldsymbol{Q}, 2)$ and $\boldsymbol{v}\boldsymbol{Q}^t \notin \mathcal{H}_j^{\boldsymbol{M}}$. By theorem 1, we have then $\boldsymbol{v} \in \mathcal{MP}(\boldsymbol{Q}, \infty)$.

If $\boldsymbol{v} \in \mathrm{Int}(\mathcal{C}(\Lambda))$, then $\boldsymbol{v}\boldsymbol{Q}^t \in \mathrm{Int}(\mathcal{C}(\Lambda))$ and hence $\boldsymbol{v}\boldsymbol{Q}^t \in \mathcal{MP}(\boldsymbol{Q}, 2)$ for all t, by theorem 8. Furthermore, as $\boldsymbol{v}\boldsymbol{Q}^t$ lies in the interior of $\mathcal{C}(\Lambda)$, it cannot be part of any of the bounding hyperplanes $\mathcal{H}_j^{\boldsymbol{M}}$, $j \in \{1,2,3\}$.

If $\boldsymbol{v} \in \mathcal{S}_\lambda$, then $\boldsymbol{v}-\boldsymbol{\alpha} \in \mathcal{E}_\lambda$. Since $\mathcal{E}_\lambda$ is a vector space, $\lambda(\boldsymbol{v}-\boldsymbol{\alpha}) \in \mathcal{E}_\lambda$. Furthermore $\boldsymbol{v}\boldsymbol{Q} - \boldsymbol{\alpha} = (\boldsymbol{v}-\boldsymbol{\alpha})\boldsymbol{Q} = \lambda(\boldsymbol{v}-\boldsymbol{\alpha}) \in \mathcal{E}_\lambda$. Hence, $\boldsymbol{v}\boldsymbol{Q} \in \mathcal{S}_\lambda$. Repeating this argument, $\boldsymbol{v}\boldsymbol{Q}^t \in \mathcal{S}_\lambda \subset \mathcal{MP}(\boldsymbol{Q}, 2)$ for all t. Using lemma 4, $\boldsymbol{v}\boldsymbol{Q}^t \notin \mathcal{H}_j^{\boldsymbol{M}}$ for all j.

If $\boldsymbol{v} \notin \mathrm{Int}(\mathcal{C}(\Lambda)) \cup \mathcal{S}_\lambda$, lemma 6 states that the structures $\boldsymbol{v}\boldsymbol{Q}^t$ approach the stationary structure $\boldsymbol{\alpha}$ along a path that is tangent to the line segment $\mathcal{S}_\Lambda$. As that line segment is contained in the interior of $\mathcal{C}(\Lambda)$, the trajectory of $\boldsymbol{v}$ enters $\mathrm{Int}(\mathcal{C}(\Lambda))$ at a certain time, say $T(\boldsymbol{v}) \in \mathbb{N}_0$. The first statement of this theorem then holds for the structure $\boldsymbol{v}\boldsymbol{Q}^{T(\boldsymbol{v})}$, and the second statement follows. □

6 Conclusions and further research questions

In manpower planning it is important for managers to know which structures of the manpower system evolve in a monotonous way as time passes. Therefore, we introduce the notion of a t-step monotonically proceeding structure in a Markov chain model ($t \in \mathbb{N}$, $t \geq 2$). The basic questions we consider in this paper are: (1) what transition matrices yield all structures to be monotonically proceeding and (2) what structures are monotonically proceeding given a transition matrix?

For two states, we prove that all structures are 2-step monotonically proceeding if and only if the transition matrix $\boldsymbol{Q}$ has the property that all off-diagonal elements of $\boldsymbol{Q}^2 - \boldsymbol{Q}$ are nonnegative (theorem 3). We call such transition matrices stepwise increasing. The equivalence does not hold for a higher number of states. For three states and primitive stochastic transition matrices $\boldsymbol{Q}$, we establish that all structures are 2-step monotonically proceeding if and only if $\boldsymbol{Q}$ is stepwise increasing and quasi-stable (theorem 4). We define a quasi-stable matrix as a square matrix where the off-diagonal elements are identical in each column. If $\boldsymbol{Q}$ is not quasi-stable, structures exist that are not 2-step monotonically proceeding. In general, the set of all 2-step monotonically proceeding structures is star-convex with respect to any stationary distribution vector (theorem 2).

Concerning stepwise increasing stochastic matrices of order three and above, we find the following necessary conditions: (1) all real eigenvalues are non-negative (theorem 5) and (2) the off-diagonal elements do not dominate the diagonal element in each column (theorem 6).

We also describe the set of all 2-step monotonically proceeding structures for the case of primitive stepwise increasing transition matrices $\boldsymbol{Q}$ of order three. If the off-diagonal elements in each column of $\boldsymbol{Q}$ are unequal, this set is a union of three point-symmetric cones with respect to the unique stationary probability distribution $\boldsymbol{\alpha}$ of $\boldsymbol{Q}$ having $\boldsymbol{\alpha}$ as their sole common point (figure 1). In addition, we show that all structures in the interior of one of these cones are monotononically proceeding over any number of steps (theorem 9). This cone is the unique one containing the structures $\boldsymbol{v}$ for which $\boldsymbol{v} - \boldsymbol{\alpha}$ is an eigenvector of $\boldsymbol{Q}$ associated with the second largest eigenvalue of $\boldsymbol{Q}$.

If a structure $\boldsymbol{v}$ is not monotonically proceeding over a certain number of steps, it would be interesting to measure the extent to which the monotonic progression fails. Further research could identify the control variables for adjusting the model parameters (transition probabilities) in order to minimize the departure from a monotonic progression.

References

Altman, E. I., & Rijken, H. A. (2004). How rating agencies achieve rating stability. *Journal of Banking & Finance*, *28*(11), 2679–2714.

Dardanoni, V. (1995). Income distribution dynamics: monotone markov chains make light work. *Social Choice and Welfare*, *12*(2), 181–192.

Gantmacher, F. R. (1960). *The theory of matrices, volume one.* Chelsea Publishing Company.

Guerry, M.-A. (2013). On the embedding problem for discrete-time markov chains. *Journal of Applied Probability*, *50*(4), 918–930.

Guerry, M.-A. (2014). Some results on the embeddable problem for discrete-time markov models in manpower planning. *Communications in Statistics-Theory and Methods*, *43*(7), 1575–1584.

Guerry, M.-A. (2017). Necessary embedding conditions for state-wise monotone markov chains. *Linear and Multilinear Algebra*, *65*(8), 1529–1539.

Jarrow, R. A., Lando, D., & Turnbull, S. M. (1997). A markov model for the term structure of credit risk spreads. *Review of Financial studies*, *10*(2), 481–523.

Mills, Q. D. (1985). Seniority versus ability in promotion decisions. *Industrial and Labor Relations Review*, *38*(3), 421–425.

Ng, T. W., Sorensen, K. L., Eby, L. T., & Feldman, D. C. (2007). Determinants of job mobility: A theoretical integration and extension. *Journal of Occupational and Organizational Psychology*, *80*(3), 363–386.

Ngoc, P. H. A. (2006). A perron–frobenius theorem for a class of positive quasi-polynomial matrices. *Applied Mathematics Letters*, *19*(8), 747–751.

Seneta, E. (2006). *Non-negative matrices and markov chains.* Springer Science & Business Media.