

## Matrix roots and embedding conditions for three-state discrete-time Markov chains with complex eigenvalues

Guerry, Marie-Anne

*Published in:*  
Communications in Mathematics and Statistics

*DOI:*  
[10.1007/s40304-020-00226-3](https://doi.org/10.1007/s40304-020-00226-3)

*Publication date:*  
2022

*License:*  
Other

*Document Version:*  
Accepted author manuscript

[Link to publication](#)

*Citation for published version (APA):*  
Guerry, M-A. (2022). Matrix roots and embedding conditions for three-state discrete-time Markov chains with complex eigenvalues. *Communications in Mathematics and Statistics*, 10(3), 435-450.  
<https://doi.org/10.1007/s40304-020-00226-3>

### Copyright

No part of this publication may be reproduced or transmitted in any form, without the prior written permission of the author(s) or other rights holders to whom publication rights have been transferred, unless permitted by a license attached to the publication (a Creative Commons license or other), or unless exceptions to copyright law apply.

### Take down policy

If you believe that this document infringes your copyright or other rights, please contact [openaccess@vub.be](mailto:openaccess@vub.be), with details of the nature of the infringement. We will investigate the claim and if justified, we will take the appropriate steps.

# Matrix roots and embedding conditions for three-state discrete-time Markov chains with complex eigenvalues

Marie-Anne Guerry

Received: date / Accepted: date

**Abstract** The present paper examines matrix root properties and embedding conditions for discrete-time Markov chains with three states and a transition matrix having complex eigenvalues. Necessary as well as sufficient conditions for the existence of an  $m$ -th stochastic root of the transition matrix, are investigated. Matrix roots are expressed in analytical form based on the spectral decomposition of the transition matrix and properties of these matrix roots are proved.

**Keywords** Markov chain · embedding problem · matrix root · complex eigenvalues

**Mathematics Subject Classification (2010)** 15A18 · 15B51 · 60J10

## 1 Introduction

The embedding problem for Markov chains is initially introduced in [1]. A discrete-time Markov chain with transition matrix  $\mathbf{P}$  is embeddable in the Markov chain with transition matrix  $\mathbf{A}$  in case there exists an integer  $m \in \mathbb{N} \setminus \{0, 1\}$  such that  $\mathbf{A}^m = \mathbf{P}$ . Since transition matrices are stochastic matrices, the embedding problem for discrete-time Markov chains can be reformulated in terms of  $m$ -th stochastic roots [2, 6]. An  $m$ -th stochastic root  $\mathbf{A}$  of  $\mathbf{P}$  is an  $m$ -th matrix root that is nonnegative and has all its row sums equal to 1. Within the set of discrete-time Markov chains, the embedding problem and necessary conditions for the existence of stochastic roots of the transition matrix are examined by Singer and Spilerman [12], and also in [3, 6]. More recently, nonnegative roots of matrices are studied in [9, 13].

---

Marie-Anne Guerry  
Vrije Universiteit Brussel (VUB), Department of Business Technology and Operations,  
Pleinlaan 2, 1050 Brussels, Belgium  
Tel.: +32-2-6148302  
E-mail: marie-anne.guerry@vub.be

For a Markov model, the transition matrix  $\mathbf{P}$  can be estimated in case data on stocks and flows are available for time intervals with length equal to the time unit of the Markov model. In practice however, it could happen that data is only available on yearly base while one is interested to know semiannual transition probabilities. In these situations there is a lack of data to estimate the semiannual transition probabilities. Nevertheless, for an embeddable matrix  $\mathbf{P}$ , a square root stochastic matrix could give insight in the semiannual transition probabilities.

For  $(2 \times 2)$  matrices the embedding problem is examined in detail: Necessary and sufficient conditions for the existence of stochastic roots are formulated. Furthermore, for embeddable Markov chains, the stochastic roots of the transition matrices are described in analytical form [2, 5, 7].

For  $(3 \times 3)$  stochastic matrices He and Gunn (2003) describe all possible real valued  $m$ -th root matrices that are functions of the original matrix, in the case of real as well as complex eigenvalues [5]. The expressed  $m$ -th root matrices have all row sums equal to 1 but are not necessarily nonnegative and therefore do not result automatically in  $m$ -th stochastic roots. Higham and Lin (2011) examine necessary conditions for the existence of stochastic matrix roots based on the set of all possible eigenvalues for a  $(3 \times 3)$  nonnegative matrix [6].

The case of  $(3 \times 3)$  stochastic matrices with real eigenvalues is investigated in detail and sufficient embedding conditions are presented in [4]. The present paper aims closing the gap for the  $(3 \times 3)$  case, and studies therefore  $m$ -th stochastic roots of  $(3 \times 3)$  stochastic matrices with complex eigenvalues. Necessary embedding conditions are proved in section 2. Section 3 presents row-normalized roots based on the spectrum decomposition and the projections of the original stochastic matrix. Sufficient embedding conditions are formulated in section 4. An example is presented to illustrate how the theoretical findings can be used in a beneficial way. Finally, section 5 formulates some suggestions for further research.

## 2 Necessary embedding conditions

In case  $\mathbf{P}$  is embeddable there does exist a stochastic matrix  $\mathbf{A}$  such that, for some  $m \in \mathbb{N} \setminus \{0, 1\}$ ,  $\mathbf{A}$  is an  $m$ -th root of  $\mathbf{P}$ , i.e.  $\mathbf{A}^m = \mathbf{P}$ . Higham and Lin (2011) present necessary conditions for the existence of  $m$ -th stochastic roots based on the inverse eigenvalue problem [6]. In what follows those findings are recapitulated with a focus on  $(3 \times 3)$  stochastic matrices and are further specified for the particular case of complex eigenvalues.

Let  $\mathbf{P}$  be a  $(3 \times 3)$  stochastic matrix with eigenvalues  $\lambda_1, \lambda_2$  and  $\lambda_3$  that are not all real numbers. According to the Perron-Frobenius Theorem, each stochastic matrix has an eigenvalue  $\lambda_1 = 1$ . Since the characteristic equation has real coefficients, the complex eigenvalues  $\lambda_2, \lambda_3$  are complex conjugates, i.e.  $\lambda_3 = \bar{\lambda}_2$ . In this way the spectrum of  $\mathbf{P}$  can be described as the set  $\sigma_{\mathbf{P}} = \{1, \lambda, \bar{\lambda}\}$  with  $\lambda \in \mathbb{C} \setminus \mathbb{R}$  eigenvalue of  $\mathbf{P}$ .

The eigenvalues  $1, \mu_2, \mu_3$  of a stochastic matrix  $\mathbf{A}$ , with  $\mathbf{A}^m = \mathbf{P}$ , satisfy  $\mu_2^m = \lambda$  and  $\mu_3^m = \bar{\lambda}$ . Since  $\mu_2^m = \lambda$  with  $\lambda \in \mathbb{C} \setminus \mathbb{R}$ , the eigenvalue  $\mu_2$  is neither a real number. The spectrum of an  $m$ -th stochastic root  $\mathbf{A}$  of  $\mathbf{P}$  is therefore equal to  $\sigma_{\mathbf{A}} = \{1, \mu, \bar{\mu}\}$  with  $\mu \in \mathbb{C} \setminus \mathbb{R}$  and  $\mu^m = \lambda$ .

A stochastic root  $\mathbf{A}$  is a nonnegative matrix root. The inverse eigenvalue problem is helpful in formulating necessary conditions on the spectrum  $\{1, \lambda, \bar{\lambda}\}$  of  $\mathbf{P}$  such that  $\{1, \mu, \bar{\mu}\}$ , with  $\mu^m = \lambda$ , is the set of eigenvalues of a nonnegative matrix  $\mathbf{B}$ . From Loewy and London (1978) it is known that  $\{1, \mu, \bar{\mu}\}$  is the spectrum of a nonnegative matrix if and only if the three eigenvalues  $1, \mu$  and  $\bar{\mu}$  belong to the closed triangle  $\Theta_3$  with vertices  $(1, 0)$ ,  $e^{\frac{2\pi i}{3}}$  and  $e^{-\frac{2\pi i}{3}}$  [8].

A nonnegative matrix  $\mathbf{B}$  is not necessarily a stochastic matrix. Nevertheless, in case there exists a nonnegative matrix  $\mathbf{B}$  with spectrum  $\{1, \mu, \bar{\mu}\}$ , there also exists a stochastic matrix  $\mathbf{A}$  with the same spectrum. Indeed, according to Rojo and Soto (2003), for a positive eigenvector  $\mathbf{X} = (x_1, x_2, x_3)$  corresponding with the maximal eigenvalue 1 of  $\mathbf{B}$ , the diagonal matrix  $\mathbf{D} = \text{diag}(x_1, x_2, x_3)$  gives rise to a nonnegative matrix  $\mathbf{A} = \mathbf{D}^{-1}\mathbf{B}\mathbf{D}$  with all its row sums equal to 1 [11]. In this way  $\mathbf{A} = \mathbf{D}^{-1}\mathbf{B}\mathbf{D}$  is a stochastic matrix with the same spectrum  $\{1, \mu, \bar{\mu}\}$  as the nonnegative matrix  $\mathbf{B}$ . This remark let us conclude that formulating necessary conditions for the existence of a stochastic matrix or for the existence of a nonnegative matrix are equivalent problems.

The closed triangle  $\Theta_3$  is the convex hull  $\text{Conv}\left\{(1, 0), e^{\frac{2\pi i}{3}}, e^{-\frac{2\pi i}{3}}\right\}$  and can be described in an alternative way as the following set of points:

$$\Theta_3 = \{(x, y) \in \mathbb{R}^2 \mid x \geq -0.5; x - 1 \leq \sqrt{3}y \leq 1 - x\} \quad (1)$$

The description (1) results in a practical way to verify the necessary condition for the existence of an  $m$ -th stochastic root of  $\mathbf{P}$  as formulated in Theorem 1. For  $\lambda = r(\cos \theta + i \sin \theta)$ , we introduce the notation  $\sqrt[m]{\lambda}(k)$ :

$$\sqrt[m]{\lambda}(k) = \sqrt[m]{r} e^{\frac{\theta + 2\pi k}{m}} = \sqrt[m]{r} \left( \cos \frac{\theta + 2\pi k}{m} + i \sin \frac{\theta + 2\pi k}{m} \right)$$

**Theorem 1** *If for the stochastic matrix  $\mathbf{P}$  with spectrum  $\sigma_{\mathbf{P}} = \{1, \lambda, \bar{\lambda}\}$  an  $m$ -th stochastic root does exist, then at least one  $m$ -th root of  $\lambda$  belongs to  $\Theta_3$ . This results in the following necessary embedding condition:*

$$\exists m \in \mathbb{N} \setminus \{0, 1\} \text{ and } \exists k \in \{0, 1, \dots, m-1\} : \sqrt[m]{\lambda}(k) \in \Theta_3$$

*Proof* In case the eigenvalue  $\lambda$  of  $\mathbf{P}$  has a modulus  $r$  and an argument  $\theta$ , the spectrum of  $\mathbf{P}$  equals  $\sigma_{\mathbf{P}} = \{1, r(\cos \theta + i \sin \theta), r(\cos \theta - i \sin \theta)\}$ .

An  $m$ -th stochastic root  $\mathbf{A}$  of  $\mathbf{P}$  has then the spectrum  $\{1, \mu, \bar{\mu}\}$ , with  $\mu^m = \lambda$ . According to [8] the set  $\{1, \mu, \bar{\mu}\}$  is the spectrum of a nonnegative matrix under the condition that  $\{1, \mu, \bar{\mu}\} \subset \Theta_3$ . Because of the  $X$ -axis symmetry of  $\Theta_3$ , this condition is equivalent with  $\mu \in \Theta_3$ .

Since the eigenvalue  $\mu$  of the  $m$ -th matrix root  $\mathbf{A}$  of  $\mathbf{P}$  satisfies  $\mu^m = \lambda$ , the eigenvalue  $\mu$  of  $\mathbf{A}$  is an  $m$ -th root of  $\lambda$ . In the set  $\mathbb{C}$  of complex numbers,  $\lambda$  has exactly  $m$   $m$ -th roots, being  $\sqrt[m]{\lambda}(k)$  for  $k = 0, 1, \dots, m-1$ . That proves the theorem.

The findings of Theorem 1 let formulate directly the insight that no effort has to be made to find an  $m$ -th stochastic root of the matrix  $\mathbf{P}$  in case for all  $k \in \{0, 1, \dots, m-1\}$  holds that  $\sqrt[m]{\lambda}(k) \notin \Theta_3$ .

Furthermore, it is interesting to remark that not all stochastic matrices with spectrum  $\{1, \mu, \bar{\mu}\}$ , satisfying  $\mu^m = \lambda$ , result in an  $m$ -th root matrix of  $\mathbf{P}$ . Besides the conditions formulated on the eigenvalues, also the eigenvectors of a root matrix of  $\mathbf{P}$  satisfy some properties. An  $m$ -th root  $\mathbf{A}$  of  $\mathbf{P}$  has the same eigenvectors as the matrix  $\mathbf{P}$  itself. Lemma 1 summarizes these insights.

**Lemma 1** *For an  $m$ -th root  $\mathbf{A}$  of  $\mathbf{P}$  holds:*

*$\mathbf{E}$  is eigenvector of  $\mathbf{P}$  corresponding with eigenvalue  $\rho$*

$\iff$   *$\mathbf{E}$  is eigenvector of  $\mathbf{A}$  corresponding with eigenvalue  $\tau$  satisfying  $\tau^m = \rho$*

Remark that, since  $\lambda \in \mathbb{C} \setminus \mathbb{R}$ , the matrix  $\mathbf{P}$  with spectrum  $\{1, \lambda, \bar{\lambda}\}$  has 3 eigenspaces all of dimension 1.

### 3 Row-normalized matrix roots

In what follows, for the stochastic matrix  $\mathbf{P}$  with  $\sigma_{\mathbf{P}} = \{1, \lambda, \bar{\lambda}\}$  row-normalized  $m$ -th roots are described analytically based on its spectral decomposition. Row-normalized roots have all row sums equal to 1 and are therefore candidates to result in stochastic roots.

Since  $\lambda$  is assumed to be no real number, the stochastic matrix  $\mathbf{P}$  has 3 distinct eigenvalues and is therefore diagonalizable:

$$\mathbf{P} = \mathbf{Q}\mathbf{D}_P\mathbf{Q}^{-1} \text{ with } \mathbf{D}_P = \text{diag}(1, \lambda, \bar{\lambda}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \bar{\lambda} \end{pmatrix}$$

The transformation matrix  $\mathbf{Q}$  and its inverse  $\mathbf{Q}^{-1}$  satisfy:

$\mathbf{Q} = (\mathbf{R}_1 \mathbf{R}_{\lambda} \mathbf{R}_{\bar{\lambda}})$  with  $\mathbf{R}_1, \mathbf{R}_{\lambda}, \mathbf{R}_{\bar{\lambda}}$  column vectors that are right eigenvectors of  $\mathbf{P}$  with respectively eigenvalue 1,  $\lambda$ , en  $\bar{\lambda}$ ; and

$$\mathbf{Q}^{-1} = \begin{pmatrix} \mathbf{L}_1 \\ \mathbf{L}_{\lambda} \\ \mathbf{L}_{\bar{\lambda}} \end{pmatrix} \text{ with } \mathbf{L}_1, \mathbf{L}_{\lambda}, \mathbf{L}_{\bar{\lambda}} \text{ row vectors that are left eigenvectors of } \mathbf{P}$$

with respectively eigenvalue 1,  $\lambda$ , en  $\bar{\lambda}$ .

An  $m$ -th root  $\mathbf{A}$  of  $\mathbf{P}$  has spectrum  $\sigma_{\mathbf{A}} = \{1, \mu, \bar{\mu}\}$  with  $\mu^m = \lambda$  and has the same eigenvectors as  $\mathbf{P}$  (according to Lemma 1). Therefore,  $\mathbf{P} = \mathbf{Q} \times \text{diag}(1, \lambda, \bar{\lambda}) \times \mathbf{Q}^{-1}$  implies  $\mathbf{A} = \mathbf{Q} \times \text{diag}(1, \mu, \bar{\mu}) \times \mathbf{Q}^{-1}$ . Consequently, all  $m$ -th root matrices of  $\mathbf{P}$  are of the form  $\mathbf{Q} \times \text{diag}(1, \mu, \bar{\mu}) \times \mathbf{Q}^{-1}$  with  $\mu^m = \lambda$  [10]. This insight is summarized in Lemma 2.

**Lemma 2** For a stochastic matrix  $\mathbf{P} = \mathbf{Q} \times \text{diag}(1, \lambda, \bar{\lambda}) \times \mathbf{Q}^{-1}$  and  $m \in \mathbb{N} \setminus \{0, 1\}$  holds that all  $m$ -th stochastic roots of  $\mathbf{P}$  belong to the set

$$S_m(\mathbf{P}) = \{ \mathbf{Q} \times \text{diag}(1, \mu, \bar{\mu}) \times \mathbf{Q}^{-1} \mid \mu^m = \lambda \}$$

This result points out that it is worth to examine the properties of the matrices of the form  $\mathbf{Q} \times \text{diag}(1, \mu, \bar{\mu}) \times \mathbf{Q}^{-1}$  in detail since no other possibility exists to become an  $m$ -th stochastic root of  $\mathbf{P}$ .

The matrix  $\mathbf{P} = \mathbf{Q} \mathbf{D}_P \mathbf{Q}^{-1}$  can be expressed by its spectral decomposition. Since  $\mathbf{P}$  is diagonalizable its spectral decomposition is of the form  $\mathbf{P} = \mathbf{P}_1 + \lambda \mathbf{P}_2 + \bar{\lambda} \mathbf{P}_3$ . The  $k$ -th projection is defined as  $\mathbf{P}_k = \mathbf{Q} \mathbf{I}_{kk} \mathbf{Q}^{-1}$  with  $\mathbf{I}_{kk}$  the  $(3 \times 3)$  matrix with the  $kk$ -th element equal to 1 and all the other elements equal to 0. In general, the projections satisfy by definition the following properties:

$$\mathbf{P}_i \mathbf{P}_i = \mathbf{P}_i \quad \text{for } i = 1, 2, 3 \quad \text{and} \quad \mathbf{P}_i \mathbf{P}_j = \mathbf{0} \quad \text{for } i \neq j \quad (2)$$

In the case of a matrix  $\mathbf{P}$  with spectrum  $\sigma_{\mathbf{P}} = \{1, \lambda, \bar{\lambda}\}$  Lemma 3 holds additionally.

**Lemma 3** For a matrix  $\mathbf{P}$  with spectrum  $\sigma_{\mathbf{P}} = \{1, \lambda, \bar{\lambda}\}$ , the projections  $\mathbf{P}_2$  and  $\mathbf{P}_3$  are related and satisfy:

$$\mathbf{P}_3 = \overline{\mathbf{P}_2}$$

In searching for an  $m$ -th root of  $\mathbf{P}$ , it is interesting to mention that the complex eigenvalues  $\lambda = r(\cos \theta + i \sin \theta)$  and  $\bar{\lambda} = r(\cos(-\theta) + i \sin(-\theta))$  have both  $m$   $m$ -th roots in  $\mathbb{C}$ , namely

$$\begin{aligned} \sqrt[m]{\lambda}(k_2) &= \sqrt[m]{r} \left( \cos \frac{\theta + 2\pi k_2}{m} + i \sin \frac{\theta + 2\pi k_2}{m} \right) \\ \sqrt[m]{\bar{\lambda}}(k_3) &= \sqrt[m]{r} \left( \cos \frac{-\theta + 2\pi k_3}{m} + i \sin \frac{-\theta + 2\pi k_3}{m} \right) \end{aligned}$$

with  $k_2, k_3 \in \{0, 1, \dots, m-1\}$

An  $m$ -th stochastic root of  $\mathbf{P}$  has eigenvalues  $\mu$  and  $\bar{\mu}$  that are complex conjugates. The  $m$ -th roots  $\sqrt[m]{\lambda}(k_2)$  and  $\sqrt[m]{\bar{\lambda}}(k_3)$  are complex conjugates for  $k_3 = m - k_2$  since in that situation the argument of  $\sqrt[m]{\bar{\lambda}}(k_3)$  equals

$$\frac{-\theta + 2\pi k_3}{m} = 2\pi - \frac{\theta + 2\pi k_2}{m}$$

For  $\mathbf{P} = \mathbf{P}_1 + \lambda \mathbf{P}_2 + \bar{\lambda} \mathbf{P}_3$ ,  $k_2 = k \in \{0, 1, \dots, m-1\}$  and  $k_3 = m - k$ , the matrix  $\mathbf{A}(m, k)$  is introduced:

$$\mathbf{A}(m, k) = \mathbf{P}_1 + \sqrt[m]{\lambda}(k) \mathbf{P}_2 + \sqrt[m]{\bar{\lambda}}(m - k) \mathbf{P}_3$$

Eq.(2) results in:

$$(\mathbf{A}(m, k))^m = \mathbf{P}_1 + \left( \sqrt[m]{\lambda}(k) \right)^m \mathbf{P}_2 + \left( \sqrt[m]{\bar{\lambda}}(m - k) \right)^m \mathbf{P}_3 = \mathbf{P}_1 + \lambda \mathbf{P}_2 + \bar{\lambda} \mathbf{P}_3 = \mathbf{P}$$

In this way, for  $k \in \{0, 1, \dots, m-1\}$ , the matrices  $\mathbf{A}(m, k)$  are  $m$ -th roots of  $\mathbf{P}$ . Moreover, in accordance to Lemma 2, an  $m$ -th root of  $\mathbf{P}$  is necessarily one of the matrices  $\mathbf{A}(m, k) = \mathbf{Q} \times \text{diag}(1, \mu, \bar{\mu}) \times \mathbf{Q}^{-1}$  with  $\mu = \sqrt[m]{\lambda}(k)$  and  $\bar{\mu} = \sqrt[m]{\lambda}(m-k)$ .

For a diagonalizable stochastic matrix  $\mathbf{P} = \mathbf{P}_1 + \lambda \mathbf{P}_2 + \bar{\lambda} \mathbf{P}_3$  all row sums of the projections  $\mathbf{P}_2$  and  $\mathbf{P}_3$  are equal to zero ([4], lemma 1) and  $\mathbf{P}_1$  is a stochastic matrix. Consequently, all row sums of  $\mathbf{A}(m, k) = \mathbf{P}_1 + \sqrt[m]{\lambda}(k) \mathbf{P}_2 + \sqrt[m]{\lambda}(m-k) \mathbf{P}_3$  are equal to one. All matrices  $\mathbf{A}(m, k)$  are therefore row-normalized  $m$ -th roots of  $\mathbf{P}$ . Theorem 2 summarizes these findings.

**Theorem 2** *For the stochastic matrix  $\mathbf{P} = \mathbf{P}_1 + \lambda \mathbf{P}_2 + \bar{\lambda} \mathbf{P}_3$  holds:*

$\mathbf{A}$  is a row-normalized  $m$ -th root of  $\mathbf{P} \iff \mathbf{A} = \mathbf{A}(m, k)$  for some  $k \in \{0, 1, \dots, m-1\}$

Remark that a matrix  $\mathbf{A}(m, k)$  is not necessarily an  $m$ -th stochastic root of  $\mathbf{P}$  since  $\mathbf{A}(m, k)$  can have negative elements.

#### 4 Sufficient embedding conditions

Since all  $m$ -th roots of the stochastic matrix  $\mathbf{P}$  are of the form  $\mathbf{A}(m, k)$ , a sufficient embedding condition concerns a condition under which at least one of the matrices  $\mathbf{A}(m, k)$  is a stochastic matrix. According to Theorem 2, all  $\mathbf{A}(m, k)$  are row-normalized and therefore a nonnegative matrix  $\mathbf{A}(m, k)$  is automatically a stochastic matrix. Besides, in studying embedding conditions, it is useful to point out that in case a stochastic matrix  $\mathbf{P}$  has an  $m$ -th stochastic root  $\mathbf{A}$  for some even number  $m$ , it has automatically also a stochastic square root, namely  $\mathbf{A}^{\frac{m}{2}}$ . Theorem 3 formulates these insights.

**Theorem 3** *The stochastic matrix  $\mathbf{P}$  is embeddable if and only if*

- *there exist  $m \in \mathbb{N} \setminus \{0, 1\}$  and  $k \in \{0, 1, \dots, m-1\}$  with  $\mathbf{A}(m, k)$  nonnegative or*
- *$\mathbf{P}$  has a stochastic square root or an  $m$ -th stochastic root for some odd  $m$ .*

In the case a stochastic matrix  $\mathbf{P}$  has no stochastic square root, one can conclude according to Theorem 3 that  $\mathbf{P}$  has neither an  $m$ -th stochastic root for whatever even number  $m$ . In what follows further properties of the matrices  $\mathbf{A}(m, k)$  are investigated.

**Theorem 4** *For the stochastic matrix  $\mathbf{P} = \mathbf{P}_1 + \lambda \mathbf{P}_2 + \bar{\lambda} \mathbf{P}_3$ ,  $m \in \mathbb{N} \setminus \{0, 1\}$  and  $k \in \{0, 1, \dots, m-1\}$ , the matrix*

$$\mathbf{A}(m, k) = \mathbf{P}_1 + \sqrt[m]{\lambda}(k) \mathbf{P}_2 + \sqrt[m]{\lambda}(m-k) \mathbf{P}_3$$

*is a real valued matrix with for  $h, l \in \{1, 2, 3\}$  the  $(h, l)$ -th element equal to*

$$(\mathbf{A}(m, k))_{hl} = (\mathbf{P}_1)_{hl} + 2r_{hl} \sqrt[m]{r} \cos\left(\theta_{hl} + \frac{\theta + 2\pi k}{m}\right) \quad (3)$$

*$r$  and  $\theta$  are respectively the modulus and the argument of  $\lambda$ ;  $r_{hl}$  and  $\theta_{hl}$  are respectively the modulus and the argument of  $(\mathbf{P}_2)_{hl}$ .*

*Proof* For  $h, l \in \{1, 2, 3\}$  and  $(\mathbf{P}_2)_{hl} = r_{hl}(\cos \theta_{hl} + i \sin \theta_{hl})$  holds:

$$\begin{aligned} (\mathbf{A}(m, k))_{hl} &= (\mathbf{P}_1)_{hl} + 2\Re[\sqrt[m]{\lambda}(k)(\mathbf{P}_2)_{hl}] \\ &= (\mathbf{P}_1)_{hl} + 2r_{hl} \sqrt[m]{r} \left[ \cos(\theta_{hl}) \cos\left(\frac{\theta + 2\pi k}{m}\right) - \sin(\theta_{hl}) \sin\left(\frac{\theta + 2\pi k}{m}\right) \right] \\ &= (\mathbf{P}_1)_{hl} + 2r_{hl} \sqrt[m]{r} \cos\left(\theta_{hl} + \frac{\theta + 2\pi k}{m}\right) \end{aligned}$$

since  $\sqrt[m]{\lambda}(k)$  and  $\sqrt[m]{\lambda}(m-k)$  are complex conjugates and  $(\mathbf{P}_3)_{hl} = \overline{(\mathbf{P}_2)_{hl}}$  according to Lemma 3. That proves the theorem.

Since  $\mathbf{P} = \mathbf{P}_1 + \lambda \mathbf{P}_2 + \overline{\lambda} \mathbf{P}_3$  and since the projections sum up to the identity matrix  $\mathbf{I}$ , the following equations hold:

$$\mathbf{P}_1 + 2\Re(\lambda \mathbf{P}_2) = \mathbf{P} \quad \text{and} \quad \mathbf{P}_1 + 2\Re(\mathbf{P}_2) = \mathbf{I} \quad (4)$$

Hereby is  $\Re(\mathbf{P}_2)$  the matrix with elements  $\Re((\mathbf{P}_2)_{hl})$ . The real and imaginary part of the elements of the projection  $\mathbf{P}_2$  can then be expressed as follows:

$$\Re(\mathbf{P}_2) = \frac{1}{2}(\mathbf{I} - \mathbf{P}_1) \quad \text{and} \quad \Im(\mathbf{P}_2) = \frac{1}{2\Im(\lambda)}[\mathbf{P}_1 - \mathbf{P} + \Re(\lambda)(\mathbf{I} - \mathbf{P}_1)] \quad (5)$$

In this way the projection  $\mathbf{P}_2$  is written in terms of  $\mathbf{P}$ ,  $\mathbf{P}_1$  and  $\lambda$ . The modulus  $r_{hl}$  can then be expressed as:

$$r_{hl} = \frac{1}{2} \sqrt{\frac{r^2(\delta_{hl} - \alpha_l)^2 + (\alpha_l - \mathbf{P}_{hl})^2 + 2\Re(\lambda)(\alpha_l - \mathbf{P}_{hl})(\delta_{hl} - \alpha_l)}{(\Im(\lambda))^2}} \quad (6)$$

and the argument  $\theta_{hl}$  satisfies  $\tan \theta_{hl} = \frac{\Im((\mathbf{P}_2)_{hl})}{\Re((\mathbf{P}_2)_{hl})}$ .

An alternative for Eq. (3) to express the matrix  $\mathbf{A}(m, k)$  is then:

$$\mathbf{A}(m, k) = \mathbf{P}_1 + 2 \sqrt[m]{r} \left[ \cos\left(\frac{\theta + 2\pi k}{m}\right) \Re(\mathbf{P}_2) - \sin\left(\frac{\theta + 2\pi k}{m}\right) \Im(\mathbf{P}_2) \right] \quad (7)$$

with  $\Re(\mathbf{P}_2)$  and  $\Im(\mathbf{P}_2)$  as in Eq. (5).

For a root  $\sqrt[m]{\lambda}(k)$  satisfying the necessary conditions formulated in Theorem 1 holds  $x \geq -0.5$ ;  $x - 1 \leq \sqrt{3}y \leq 1 - x$  with  $x = \Re(\sqrt[m]{\lambda}(k)) = \sqrt[m]{r} \cos \frac{\theta + 2\pi k}{m}$  and  $y = \Im(\sqrt[m]{\lambda}(k)) = \sqrt[m]{r} \sin \frac{\theta + 2\pi k}{m}$ . Theorem 5 formulates further linear conditions on  $\sqrt[m]{\lambda}(k)$  to guarantee that  $\mathbf{A}(m, k)$  is a stochastic matrix root of  $\mathbf{P}$ .

It is useful to remark that the rows of the first projection  $\mathbf{P}_1$  are identical stochastic vectors. Denoting this row vector as  $(\alpha_1, \alpha_2, \alpha_3)$  results in  $0 < (\mathbf{P}_1)_{hl} = \alpha_l < 1$ .

**Theorem 5** *For the stochastic matrix  $\mathbf{P}$ , the  $m$ -th root  $\mathbf{A}(m, k)$  is a stochastic matrix if and only if  $\sqrt[m]{\lambda}(k)$  satisfies*

$$x - 1 \leq c_{hl}y \quad \forall h \neq l \quad \text{and} \quad c_{ll}y \leq x + \frac{\alpha_l}{1 - \alpha_l} \quad \forall l$$

with  $c_{hl} = \frac{(\mathbf{P}_1 - \mathbf{P})_{hl}}{\Im(\lambda)(\mathbf{I} - \mathbf{P}_1)_{hl}} + \frac{\Re(\lambda)}{\Im(\lambda)} \quad \forall h, l$ .



*Proof* The matrix  $\mathbf{A}(m, k)$  is row-normalized and therefore it is a stochastic matrix root of  $\mathbf{P}$  if and only if it is nonnegative.

On the one hand for all  $h \neq l$  holds that  $(\mathbf{P}_1)_{hl} + 2\Re((\mathbf{P}_2)_{hl}) = 0$  according to Eq. (4). Consequently,  $(\mathbf{A}(m, k))_{hl} \geq 0$  is equivalent with  $(\mathbf{A}(m, k))_{hl} \geq (\mathbf{P}_1)_{hl} + 2\Re((\mathbf{P}_2)_{hl})$ . Furthermore, Eq. (7) results in  $(\mathbf{A}(m, k))_{hl} = (\mathbf{P}_1)_{hl} + 2\Re(\sqrt[m]{\lambda}(k)) \cdot \Re(\mathbf{P}_2)_{hl} - 2\Im(\sqrt[m]{\lambda}(k)) \cdot \Im(\mathbf{P}_2)_{hl}$  and according to Eq. (5) holds that  $\Re(\mathbf{P}_2)_{hl} = -\frac{1}{2}(\mathbf{P}_1)_{hl} < 0$ . Therefore the following equivalence holds:

$$(\mathbf{A}(m, k))_{hl} \geq 0 \iff \Re(\sqrt[m]{\lambda}(k)) - 1 \leq \frac{\Im((\mathbf{P}_2)_{hl})}{\Re((\mathbf{P}_2)_{hl})} \Im(\sqrt[m]{\lambda}(k)) \quad \forall h \neq l$$

On the other hand for all  $l = 1, 2, 3$  is  $(\mathbf{A}(m, k))_{ll} = (\mathbf{P}_1)_{ll} + 2\Re(\sqrt[m]{\lambda}(k)) \cdot \Re(\mathbf{P}_2)_{ll} - 2\Im(\sqrt[m]{\lambda}(k)) \cdot \Im(\mathbf{P}_2)_{ll}$  and  $\Re(\mathbf{P}_2)_{ll} = \frac{1}{2}(1 - (\mathbf{P}_1)_{ll}) > 0$  (according to Eq. (5)), resulting in:

$$(\mathbf{A}(m, k))_{ll} \geq 0 \iff \frac{\Im((\mathbf{P}_2)_{ll})}{\Re((\mathbf{P}_2)_{ll})} \Im(\sqrt[m]{\lambda}(k)) \leq \Re(\sqrt[m]{\lambda}(k)) + \frac{(\mathbf{P}_1)_{ll}}{2\Re((\mathbf{P}_2)_{ll})} \quad \forall l = 1, 2, 3$$

with  $\frac{(\mathbf{P}_1)_{ll}}{2\Re((\mathbf{P}_2)_{ll})} = \frac{\alpha_l}{1 - \alpha_l}$ .

Consequently, the sufficient condition to have  $(\mathbf{A}(m, k))_{hl}$  nonnegative is that the root  $\sqrt[m]{\lambda}(k)$  satisfies  $x - 1 \leq c_{hl}y \quad \forall h \neq l$  and  $c_{ll}y \leq x + \frac{\alpha_l}{1 - \alpha_l} \quad \forall l$  with  $c_{hl} = \frac{\Im((\mathbf{P}_2)_{hl})}{\Re((\mathbf{P}_2)_{hl})} \quad \forall h, l$ . These conditions in combination with the expressions for  $\Re((\mathbf{P}_2)_{hl})$  and  $\Im((\mathbf{P}_2)_{hl})$  as in Eq. (5) prove the theorem.

Theorem 6 formulates a sufficient embedding condition relating the modulus  $r$  of  $\lambda$ , the moduli  $r_{hl}$  of the elements of  $\mathbf{P}_2$  and the elements  $\alpha_l$  of the first projection  $\mathbf{P}_1$ . The formulated condition is easy to implement in practice: The values of  $r_{hl}$  can be computed by Eq. (6) and the vector  $(\alpha_1, \alpha_2, \alpha_3)$  can be determined as the unique stochastic fixpoint of  $\mathbf{P}$ .

**Lemma 4** *In case  $\sqrt[m]{r} \leq \min_{h,l} \left\{ \frac{\alpha_l}{2r_{hl}}, \frac{1 - \alpha_l}{2r_{hl}} \right\}$  then  $\mathbf{A}(m, k)$  is a stochastic matrix for all  $k \in \{0, 1, \dots, m - 1\}$ .*

*Proof* Since the matrix  $\mathbf{A}(m, k)$  is row-normalized, Eq. (3) results in:

$$\mathbf{A}(m, k) \text{ is a stochastic matrix} \iff \sqrt[m]{r} \cos\left(\theta_{hl} + \frac{\theta + 2\pi k}{m}\right) \in \left[\frac{-\alpha_l}{2r_{hl}}, \frac{1 - \alpha_l}{2r_{hl}}\right] \quad \forall h, l.$$

Therefore  $\mathbf{A}(m, k)$  is a stochastic matrix in case  $\sqrt[m]{r} \leq \min_{h,l} \left\{ \frac{\alpha_l}{2r_{hl}}, \frac{1 - \alpha_l}{2r_{hl}} \right\}$ .

**Theorem 6** *The stochastic matrix  $\mathbf{P} = \mathbf{P}_1 + \lambda\mathbf{P}_2 + \overline{\lambda}\overline{\mathbf{P}_2}$  is embeddable in case*

$$\sqrt{r} \leq \min_{h,l} \left\{ \frac{\alpha_l}{2r_{hl}}, \frac{1 - \alpha_l}{2r_{hl}} \right\}$$

*Proof* According to Theorem 3,  $\mathbf{P}$  is embeddable if and only if the matrix  $\mathbf{A}(m, k)$  is stochastic for at least one pair  $(m, k)$  with  $m \in \mathbb{N} \setminus \{0, 1\}$  and  $k \in \{0, 1, \dots, m - 1\}$ . Therefore lemma 4 let conclude that  $\mathbf{P}$  is embeddable in case there exists a value for  $m$  satisfying  $\sqrt[m]{r} \leq \min_{h,l} \left\{ \frac{\alpha_l}{2r_{hl}}, \frac{1 - \alpha_l}{2r_{hl}} \right\}$ . This condition

is less restrictive for  $m = 2$  since  $\sqrt[m]{r}$  is increasing with  $m$ . Thus in case  $\sqrt[m]{r} \leq \min_{h,l} \left\{ \frac{\alpha_l}{2r_{hl}}, \frac{1-\alpha_l}{2r_{hl}} \right\}$  for some  $m > 2$  then also  $\sqrt{r} \leq \min_{h,l} \left\{ \frac{\alpha_l}{2r_{hl}}, \frac{1-\alpha_l}{2r_{hl}} \right\}$ . Furthermore, the condition  $\sqrt{r} \leq \min_{h,l} \left\{ \frac{\alpha_l}{2r_{hl}}, \frac{1-\alpha_l}{2r_{hl}} \right\}$  guarantees that  $\mathbf{A}(2, k)$  is a stochastic square root of  $\mathbf{P}$ .

**Corollary 1** *In case  $\sqrt[m]{r} \leq \min_{h,l} \left\{ \frac{\alpha_l}{2r_{hl}}, \frac{1-\alpha_l}{2r_{hl}} \right\}$  then  $\mathbf{P}$  has at least  $\frac{m(m+1)}{2} - 1$  stochastic roots.*

This corollary results directly from Lemma 4 and Theorem 6 since  $\sqrt[m]{r} \leq \min_{h,l} \left\{ \frac{\alpha_l}{2r_{hl}}, \frac{1-\alpha_l}{2r_{hl}} \right\}$  implies that  $\mathbf{A}(\tilde{m}, k)$  is a stochastic matrix for all  $2 \leq \tilde{m} \leq m$  and for all  $k \in \{0, 1, \dots, \tilde{m} - 1\}$ .

In searching for further sufficient embedding conditions, let us introduce for  $h, l \in \{1, 2, 3\}$  and  $k \in \mathbb{N}$ , the function

$$\begin{aligned} f_{hl}^k(t) &= \alpha_l + 2r_{hl}r^t \cos(\theta_{hl} + (\theta + 2\pi k)t) \\ &= \alpha_l + 2r^t [\cos((\theta + 2\pi k)t) \Re((\mathbf{P}_2)_{hl}) - \sin((\theta + 2\pi k)t) \Im((\mathbf{P}_2)_{hl})] \end{aligned}$$

This function satisfies according to Eq. (3):

$$f_{hl}^k\left(\frac{1}{m}\right) = (\mathbf{A}(m, k))_{hl} \quad (8)$$

Furthermore holds:

$$f_{hl}^k(0) = \alpha_l + 2r_{hl} \cos \theta_{hl} = (\mathbf{P}_1)_{hl} + (\mathbf{P}_2)_{hl} + (\overline{\mathbf{P}_2})_{hl} = \mathbf{I}_{hl} = \delta_{hl} \quad (9)$$

$$f_{hl}^k(1) = \alpha_l + 2r_{hl}r \cos(\theta_{hl} + \theta + 2\pi k) = (\mathbf{P}_1)_{hl} + \lambda (\mathbf{P}_2)_{hl} + \bar{\lambda} (\overline{\mathbf{P}_2})_{hl} = \mathbf{P}_{hl} \quad (10)$$

In this way, the value of the function  $f_{hl}^k$  is nonnegative in  $t = 0$  as well as in  $t = 1$ .

In what follows further insights are gained on the  $m$ -th roots  $\mathbf{A}(m, k)$  of the stochastic matrix  $\mathbf{P}$  by examining properties of  $f_{hl}^k(t)$  for  $t \in [0, 1]$ . A pertinent question is under what condition at least one of the roots  $\mathbf{A}(m, k)$  results in a stochastic matrix. Lemma 5 describes the critical values of the function  $f_{hl}^k(t)$ . The notation  $\arctan$  refers to the arctangent function.

**Lemma 5** *The critical values of  $f_{hl}^k(t) = \alpha_l + 2r_{hl}r^t \cos(\theta_{hl} + (\theta + 2\pi k)t)$  are elements of the set*

$$S_{hl} = \left\{ \frac{1}{\theta + 2\pi k} \left[ \arctan\left(\frac{\ln r}{\theta + 2\pi k}\right) - \theta_{hl} + \pi q \right] \mid q \in \mathbb{Z} \right\}$$

*Proof* The derivative of  $f_{hl}^k(t)$  equals:

$\frac{d}{dt} f_{hl}^k(t) = 2r_{hl}r^t [\ln r \cos(\theta_{hl} + (\theta + 2\pi k)t) - (\theta + 2\pi k) \sin(\theta_{hl} + (\theta + 2\pi k)t)]$ . In solving the equation  $\frac{d}{dt} f_{hl}^k(t) = 0$  we can take into account that  $\theta + 2\pi k \neq 0$  since  $\theta$  is the argument of the eigenvalue  $\lambda \in \mathbb{C} \setminus \mathbb{R}$ . Furthermore a value for

$t$  satisfying  $\cos(\theta_{hl} + (\theta + 2\pi k)t) = 0$  cannot be a solution of the equation  $\frac{d}{dt}f_{hl}^k(t) = 0$  since this would imply that also  $\sin(\theta_{hl} + (\theta + 2\pi k)t) = 0$ , which is impossible. Therefore:

$$\frac{d}{dt}f_{hl}^k(t) = 0 \Leftrightarrow \tan(\theta_{hl} + (\theta + 2\pi k)t) = \frac{\ln r}{\theta + 2\pi k}$$

with solutions  $t_{hl}^k(q) = \frac{1}{\theta + 2\pi k} \left[ \arctan\left(\frac{\ln r}{\theta + 2\pi k}\right) - \theta_{hl} + \pi q \right] \quad (q \in \mathbb{Z})$ .

The  $m$   $m$ -th roots  $\sqrt[m]{\lambda}(k)$  of  $\lambda$  correspond with  $k \in \{0, 1, \dots, m-1\}$ . The fact that  $k \leq m-1$  implies  $\frac{1}{m} \leq \frac{1}{k+1}$ . Since  $(\mathbf{A}(m, k))_{hl} = f_{hl}^k\left(\frac{1}{m}\right)$  and  $\frac{1}{m} \in ]0, \frac{1}{k+1}]$ , in examining for a particular value for  $k$  conditions that guarantee that the  $m$ -th root  $\mathbf{A}(m, k)$  is a stochastic matrix, only the critical values  $t_{hl}^k(q)$  that belong to  $]0, \frac{1}{k+1}]$  are of importance. The following lemma gives information on the number of critical values  $t_{hl}^k(q) \in S_{hl} \cap ]0, \frac{1}{k+1}]$ .

**Lemma 6** *For  $k$  fix, the number of critical values  $t_{hl}^k(q) \in S_{hl} \cap ]0, \frac{1}{k+1}]$  of  $f_{hl}^k$  is at most equal to 2. In particular for  $k = 0$ , the number of critical values is at most equal to 1.*

*Proof* For two subsequent critical values  $t_{hl}^k(q)$  and  $t_{hl}^k(q+1)$  holds that  $|t_{hl}^k(q+1) - t_{hl}^k(q)| = \frac{\pi}{|\theta + 2\pi k|}$ . Therefore expressing the length of the interval  $]0, \frac{1}{k+1}]$  proportionally to the length of the interval  $[t_{hl}^k(q), t_{hl}^k(q+1)]$  results in

$$\frac{\frac{1}{k+1}}{\frac{\pi}{|\theta + 2\pi k|}} = \left| \frac{\theta + 2\pi k}{(k+1)\pi} \right|$$

The argument  $\theta$  of the eigenvalue  $\lambda \in \mathbb{C} \setminus \mathbb{R}$  satisfies  $\theta \in ]-\pi, \pi[$ . Therefore  $\frac{\theta + 2\pi k}{(k+1)\pi} \in ]\frac{2k-1}{k+1}, \frac{2k+1}{k+1}[$  with  $\frac{2k+1}{k+1} < 2$ . Consequently the number of critical values  $t_{hl}^k(q) \in S_{hl}$  that belong to  $]0, \frac{1}{k+1}]$  is at most equal to 2.

In particular for  $k = 0$  holds  $\left| \frac{\theta + 2\pi k}{(k+1)\pi} \right| = \left| \frac{\theta}{\pi} \right| < 1$ .

For the root matrices  $\mathbf{A}(m, k)$  under study is  $k \leq m-1$  and thus  $m \geq k+1$ . The following theorem presents a sufficient condition that guarantees that the  $(h, l)$ -th element of  $\mathbf{A}(m, k)$  is nonnegative for all  $m \geq k+1$ . The condition is formulated on the value of  $b_{hl}(k)$ :

$$b_{hl}(k) = \frac{\ln r \cdot \Re((\mathbf{P}_2)_{hl}) - (\theta + 2\pi k) \cdot \Im((\mathbf{P}_2)_{hl})}{(\theta + 2\pi k) \cdot \Re((\mathbf{P}_2)_{hl}) + \ln r \cdot \Im((\mathbf{P}_2)_{hl})} \quad (11)$$

with  $\Re((\mathbf{P}_2)_{hl})$  and  $\Im((\mathbf{P}_2)_{hl})$  as in Eq. (5).

**Theorem 7** *In case for  $h, l \in \{1, 2, 3\}$  and  $k \in \mathbb{N}$  holds that  $b_{hl}(k) \notin \text{Conv}\{0, \frac{\Im(\lambda)}{\Re(\lambda)}\}$  then  $(\mathbf{A}(m, k))_{hl} \geq 0$  for all  $m \geq k+1$ .*

*Proof* Critical values  $t^*$  of  $f_{hl}^k(t)$  satisfy  $\tan(\theta t^*) = b_{hl}(k)$  with  $b_{hl}(k)$  defined as in Eq.(11). Since  $\tan(\theta) = \frac{\Im(\lambda)}{\Re(\lambda)}$ , the condition  $b_{hl}(k) \notin \text{Conv}\{0, \frac{\Im(\lambda)}{\Re(\lambda)}\}$  implies that  $f_{hl}^k(t)$  has no critical value in  $[0, 1]$ . As a result,  $f_{hl}^k(t)$  is monotonously evolving on  $[0, 1]$ . Moreover, according to Eq.(9) and Eq.(10), both  $f_{hl}^k(0)$  and  $f_{hl}^k(1)$  are nonnegative. Consequently  $(\mathbf{A}(m, k))_{hl} = f_{hl}^k(\frac{1}{m}) \geq 0 \quad \forall m \geq k + 1$ .

The following theorem formulates a sufficient embedding condition for  $\mathbf{P}$  by ensuring that the row-normalized root  $\mathbf{A}(m, k)$  is a stochastic matrix.

**Theorem 8** *In case  $f_{hl}^k(\frac{1}{k+1}) \in [0, 1]$  as well as  $f_{hl}^k(t_{hl}^k(q)) \in [0, 1]$ , for all  $t_{hl}^k(q) \in S_{hl} \cap ]0, \frac{1}{k+1}]$  and for all  $(h, l) \in \{1, 2, 3\} \times \{1, 2, 3\}$ , then  $\mathbf{A}(m, k)$  is an  $m$ -th stochastic root of  $\mathbf{P}$ , for all  $m \geq k + 1$ .*

*Proof* According to Eq.(9) holds that  $f_{hl}^k(0) \in [0, 1]$ . Therefore in case  $f_{hl}^k(\frac{1}{k+1}) \in [0, 1]$  as well as  $f_{hl}^k(t_{hl}^k(q)) \in [0, 1]$  for all  $t_{hl}^k(q) \in S_{hl} \cap ]0, \frac{1}{k+1}]$ , the infimum and the supremum of  $f_{hl}^k$  on  $]0, \frac{1}{k+1}]$  are elements of  $[0, 1]$ . Consequently  $f_{hl}^k(t)$  is an element of  $[0, 1]$  for all  $t \in ]0, \frac{1}{k+1}]$ . Furthermore for the  $m$ -th root  $\mathbf{A}(m, k)$  holds  $(\mathbf{A}(m, k))_{hl} = f_{hl}^k(\frac{1}{m})$ . In this way, under the stated conditions, the  $m$ -th root  $\mathbf{A}(m, k)$  is a stochastic matrix.

In what follows the study regarding the stochastic property of the row-normalized matrix  $\mathbf{A}(m, k)$  is approached in a geometrical way. Let us therefore denote the  $i$ th row of the matrix  $\mathbf{A}(m, k)$  by  $\mathbf{a}_i(m, k)$  for  $i = 1, 2, 3$ . Then by definition:

$$\mathbf{a}_i(m, k) = \left( f_{i1}^k\left(\frac{1}{m}\right) \quad f_{i2}^k\left(\frac{1}{m}\right) \quad f_{i3}^k\left(\frac{1}{m}\right) \right)$$

and

$$\mathbf{A}(m, k) \text{ is a stochastic matrix} \iff \mathbf{a}_i(m, k) \in \text{Conv}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} \text{ for } i = 1, 2, 3$$

with  $\text{Conv}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  the convex hull of  $\mathbf{e}_1 = (100)$ ,  $\mathbf{e}_2 = (010)$  and  $\mathbf{e}_3 = (001)$ .

For  $i \in \{1, 2, 3\}$  and  $k \in \mathbb{N}$ , the path with parametric equation  $(f_{i1}^k(t) \quad f_{i2}^k(t) \quad f_{i3}^k(t))$  and  $t \in [0, 1]$  is denoted by  $\mathcal{P}_i^k$ . Then, for  $m \in \mathbb{N}$  with  $m \geq k + 1$ , the row vectors  $\mathbf{a}_i(m, k)$  correspond with discrete points on the path  $\mathcal{P}_i^k$ . Furthermore, according to Eq.(9) and Eq.(10) the row vector  $\mathbf{e}_i$  and the  $i$ th row  $\mathbf{p}_i$  of  $\mathbf{P}$  satisfy

$$\mathbf{e}_i = (f_{i1}^k(0) \quad f_{i2}^k(0) \quad f_{i3}^k(0)) \quad \text{and} \quad \mathbf{p}_i = (f_{i1}^k(1) \quad f_{i2}^k(1) \quad f_{i3}^k(1))$$

Consequently, the path  $\mathcal{P}_i^k$  has starting point  $\mathbf{e}_i$  and ending point  $\mathbf{p}_i$ . Since the boundary of  $\text{Conv}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  is  $\cup_{r \neq s} \text{Conv}\{\mathbf{e}_r, \mathbf{e}_s\}$  and the points of  $\text{Conv}\{\mathbf{e}_r, \mathbf{e}_s\}$  have equal (zero)  $l$ th component for  $l \notin \{r, s\}$ , lemma 6 guarantees that the path  $\mathcal{P}_i^k$  intersects the segment  $\text{Conv}\{\mathbf{e}_r, \mathbf{e}_s\}$  at most twice. Indeed, a part of the path  $\mathcal{P}_i^k$  where the  $l$ th component is evolving monotonously,

intersects  $\text{Conv}\{\mathbf{e}_r, \mathbf{e}_s\}$  at most once. Those intersection points are important in determining stochastic row vectors  $\mathbf{a}_i(m, k)$  that are laying on parts of the path  $\mathcal{P}_i^k$  within  $\text{Conv}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ . This path-approach is useful in examining geometrically the stochasticity of a root  $\mathbf{A}(m, k)$ : in case for all the 3 paths  $\mathcal{P}_1^k$ ,  $\mathcal{P}_2^k$  and  $\mathcal{P}_3^k$  the point with  $t = \frac{1}{m}$  is on a part of the paths within  $\text{Conv}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ , the matrix  $\mathbf{A}(m, k)$  is a stochastic root of  $\mathbf{P}$ .

### Example

The following discussion and example illustrate how the theoretical findings can be useful in the context of stochastic matrix roots and for the embedding problem. The roots  $\mathbf{A}(m, k)$  are examined for the stochastic matrix

$$\mathbf{P} = \begin{pmatrix} 0.31 & 0.49 & 0.2 \\ 0.29 & 0.61 & 0.1 \\ 0.2 & 0.59 & 0.21 \end{pmatrix} \text{ that has eigenvalues } \lambda_1 = 1, \lambda_2 = 0.065 + \frac{\sqrt{0.0279}}{2}i,$$

with modulus  $r = 0.10583$  and argument  $\theta = \arctan \frac{\sqrt{0.0279}}{0.13} \approx 0.9095$ , and  $\lambda_3$  the complex conjugate of  $\lambda_2$ .

$$\text{The first projection } \mathbf{P}_1 \text{ equals } \mathbf{P}_1 = \begin{pmatrix} 0.2827 & 0.5732 & 0.1441 \\ 0.2827 & 0.5732 & 0.1441 \\ 0.2827 & 0.5732 & 0.1441 \end{pmatrix}$$

For the projection  $\mathbf{P}_2$  the moduli  $r_{hl}$  and the arguments  $\theta_{hl}$  of the elements  $(\mathbf{P}_2)_{hl}$  are as follows

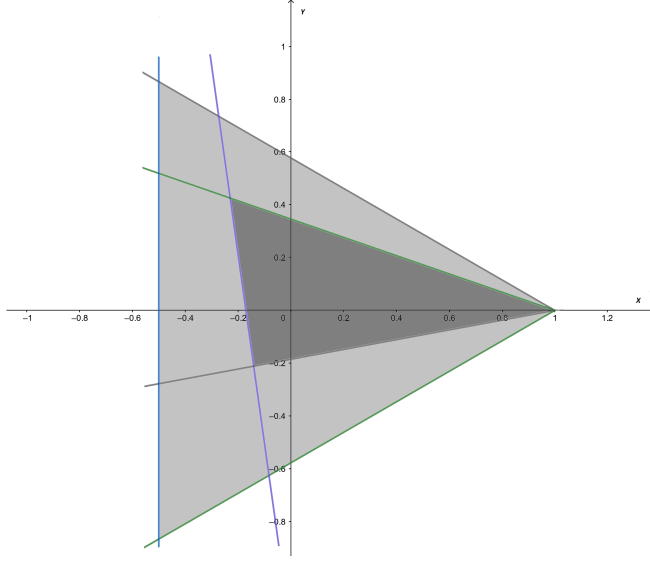
$$(r_{hl}) = \begin{pmatrix} 0.3768 & 0.3972 & 0.3972 \\ 0.2089 & 0.2202 & 0.2202 \\ 0.4101 & 0.4323 & 0.4323 \end{pmatrix} \text{ and}$$

$$(\theta_{hl}) = \begin{pmatrix} 0.3118 & 2.3768 & -1.7532 \\ -2.3140 & -0.2480 & 1.9042 \\ 1.9226 & -2.2955 & -0.1424 \end{pmatrix}$$

The projection  $\mathbf{P}_3$  satisfies  $\mathbf{P}_3 = \overline{\mathbf{P}_2}$ , according to Lemma 3.

For some particular values of  $m$  and  $k$ , and for  $\Theta_3$  as specified in (1), the necessary condition  $\sqrt[m]{\lambda}(k) \in \Theta_3$  (as formulated in Theorem 1) let conclude that  $\mathbf{A}(m, k)$  cannot be a nonnegative matrix and therefore cannot be an  $m$ -th stochastic root of  $\mathbf{P}$ . For example for  $m = 6$  and  $k = 1$ , the corresponding  $m$ -th root of  $\lambda$  results in  $\sqrt[6]{\lambda}(1) = \sqrt[6]{r} \cos \frac{\theta+2\pi}{6} + i \sqrt[6]{r} \sin \frac{\theta+2\pi}{6}$  that does not satisfy the condition  $x - 1 \leq \sqrt{3}y \leq 1 - x$ . Since  $\sqrt[6]{\lambda}(1) \notin \Theta_3$ ,  $\mathbf{A}(6, 1)$  is a row-normalized 6-th root of  $\mathbf{P}$  that has some negative elements and therefore does not result in a stochastic root of  $\mathbf{P}$ .

The sufficient embedding condition formulated in Theorem 6 is not satisfied since  $0.3253 \approx \sqrt{r} > \min_{h,l} \left\{ \frac{\alpha_l}{2r_{hl}}, \frac{1-\alpha_l}{2r_{hl}} \right\} \approx 0.1667$  and gives therefore, for the example under study, no information on whether or not the matrix  $\mathbf{P}$  is embeddable.



**Fig. 1** Graphical presentation of  $\Theta_3$  (in light grey) and of the region where  $\sqrt[m]{\lambda}(k)$  satisfies the sufficient condition (according to Theorem 5) to have for  $\mathbf{P}$  the stochastic root  $\mathbf{A}(m, k)$  (in dark grey)

Theorem 5 formulates a sufficient condition on  $\sqrt[m]{\lambda}(k)$  to have the matrix root  $\mathbf{A}(m, k)$  nonnegative. This condition is expressed in terms of  $c_{hl} = \frac{(\mathbf{P}_1 - \mathbf{P})_{hl}}{\Im(\lambda)(\mathbf{I} - \mathbf{P}_1)_{hl}} + \frac{\Re(\lambda)}{\Im(\lambda)} \forall h, l$ . For the example under consideration these  $c_{hl}$  are as follows:

$$(c_{hl}) = \begin{pmatrix} 0.3226 & -0.9597 & 5.4232 \\ 1.0875 & -0.2541 & -2.8861 \\ -2.7245 & 1.1292 & -0.1436 \end{pmatrix}$$

The region determined by the sufficient condition (according to Theorem 5) is graphically presented in Figure 1. For each  $\sqrt[m]{\lambda}(k)$  belonging to this region, the matrix  $\mathbf{A}(m, k)$  is a stochastic root of  $\mathbf{P}$ . The square root  $\sqrt{\lambda}(0)$ , for example, has  $x = 0.2923$  and  $y = 0.1429$  that satisfy the formulated sufficient condition. Consequently, we can conclude that  $\mathbf{A}(2, 0)$  is a stochastic root of  $\mathbf{P}$ .

In examining whether the row-normalized roots  $\mathbf{A}(m, k)$  of  $\mathbf{P}$  result in stochastic roots, and for what values of  $m$  this would be the case, the behavior of the function  $f_{hl}^k$  and the sufficient conditions formulated in Theorem 8 and Theorem 7 can provide useful insights. For example in analyzing the root matrices  $\mathbf{A}(m, k)$  for  $k = 0$ , the values of  $b_{hl}(0)$  as in Eq. (11) and the condition formulated in Theorem 7, let us conclude that  $(\mathbf{A}(m, 0))_{hl} \geq 0$  for all  $m \geq 2$  and  $(h, l) \notin \{(1, 3), (2, 3), (3, 1)\}$ . The reason for this is that for  $(h, l) \notin \{(1, 3), (2, 3), (3, 1)\}$  the function  $f_{hl}^k$  has no critical value in  $[0, 1]$  and is therefore monotone in  $[0, 1]$ . Moreover, according to Eq.(9) and Eq.(10), both  $f_{hl}^0(0)$  and  $f_{hl}^0(1)$  are elements of  $[0, 1]$ . Consequently, for all  $m \in \{2, 3, 4, \dots\}$

the  $m$ -th root  $\mathbf{A}(m, 0)$  of  $\mathbf{P}$  has for all  $(h, l) \notin \{(1, 3), (2, 3), (3, 1)\}$  elements that belong to  $[0, 1]$ .

For  $(h, l) \in \{(1, 3), (2, 3), (3, 1)\}$  the properties of  $f_{hl}^0$  and  $(\mathbf{A}(m, 0))_{hl}$  are examined separately in what follows. Regarding the critical values  $t_{hl}^0(q) = \frac{1}{\theta} [\arctan(\frac{\ln r}{\theta}) - \theta_{hl} + \pi q]$  ( $q \in \mathbb{Z}$ ) of the functions  $f_{hl}^0(t)$ , there can be remarked that the only critical values that belong to  $[0, 1]$  are  $t_{13}^0(0)$ ,  $t_{23}^0(1)$  and  $t_{31}^0(1)$ .

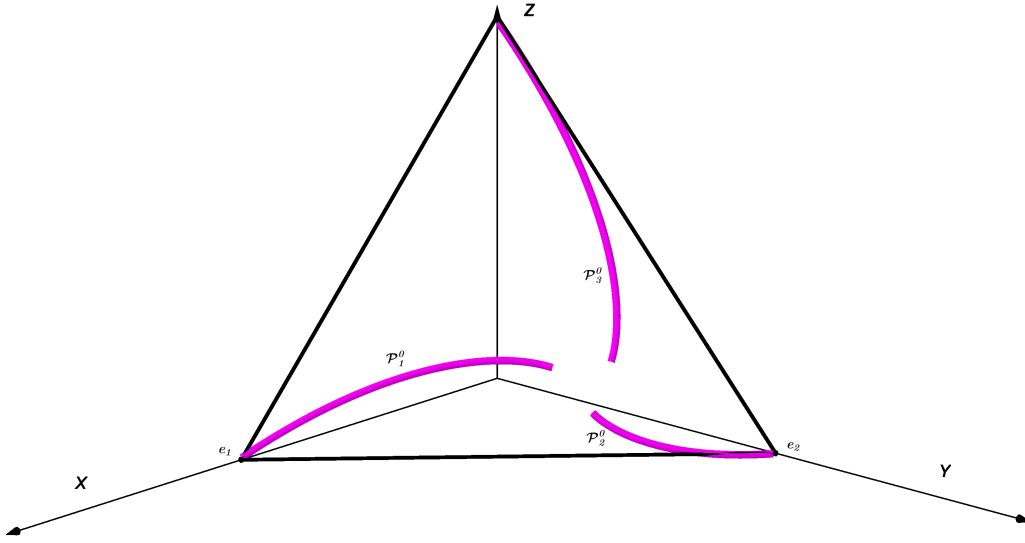
For  $(h, l) = (1, 3)$ , the critical value  $t_{13}^0(0) = 0.6237 > \frac{1}{2}$  and therefore  $f_{13}^0$  is monotonous in  $[0, \frac{1}{2}]$ . Furthermore  $f_{13}^0(0) = 0$  and  $f_{13}^0(\frac{1}{2}) = 0.0425$  so that for all  $t = \frac{1}{m}$  the value of  $f_{13}^0(t)$  is in between 0 and 1. Consequently,  $(\mathbf{A}(m, 0))_{13} \in [0, 1]$  for all  $m \geq 2$ .

For  $(h, l) = (2, 3)$ , the critical value  $t_{23}^0(1) = 0.0565$  with  $f_{23}^0(t_{23}^0(1)) = -0.0015$ . The function  $f_{23}^0$  is monotonous in  $[t_{23}^0(1), 1]$ . Furthermore  $f_{23}^0(1) = \mathbf{P}_{23}$ ,  $f_{23}^0(\frac{1}{8}) > 0$  and  $f_{23}^0(\frac{1}{9}) < 0$ . Since  $f_{23}^0(1)$  and  $f_{23}^0(\frac{1}{8})$  are both elements of  $[0, 1]$ , also  $(\mathbf{A}(m, 0))_{23} \in [0, 1]$  for all  $m$ -th roots  $\mathbf{A}(m, 0)$  of  $\mathbf{P}$  with  $2 \leq m \leq 8$ . A similar reasoning for  $(h, l) = (3, 1)$  results in the insight that  $(\mathbf{A}(m, 0))_{31} \in [0, 1]$  for all  $m$ -th roots  $\mathbf{A}(m, 0)$  of  $\mathbf{P}$  with  $2 \leq m \leq 9$ .

In accordance with Theorem 8 these insights regarding the functions  $f_{hl}^0$  let conclude that all  $\mathbf{A}(m, 0)$  with  $m = 2, \dots, 8$  are stochastic roots of  $\mathbf{P}$ . For  $m = 2, \dots, 8$  the  $m$ -th stochastic roots  $\mathbf{A}(m, 0)$  of  $\mathbf{P}$  are as follows:

$$\begin{aligned} \mathbf{A}(2, 0) &= \begin{pmatrix} 0.4593 & 0.3271 & 0.2136 \\ 0.2440 & 0.7134 & 0.0426 \\ 0.0900 & 0.4982 & 0.4118 \end{pmatrix} & \mathbf{A}(3, 0) &= \begin{pmatrix} 0.5739 & 0.2367 & 0.1894 \\ 0.1985 & 0.7812 & 0.0203 \\ 0.0463 & 0.4059 & 0.5478 \end{pmatrix} \\ \mathbf{A}(4, 0) &= \begin{pmatrix} 0.6516 & 0.1839 & 0.1645 \\ 0.1652 & 0.8243 & 0.0105 \\ 0.0266 & 0.3379 & 0.6355 \end{pmatrix} & \mathbf{A}(5, 0) &= \begin{pmatrix} 0.7062 & 0.1499 & 0.1439 \\ 0.1408 & 0.8536 & 0.0056 \\ 0.0164 & 0.2882 & 0.6954 \end{pmatrix} \\ \mathbf{A}(6, 0) &= \begin{pmatrix} 0.7463 & 0.1264 & 0.1273 \\ 0.1225 & 0.8746 & 0.0029 \\ 0.0105 & 0.2507 & 0.7388 \end{pmatrix} & \mathbf{A}(7, 0) &= \begin{pmatrix} 0.7770 & 0.1091 & 0.1139 \\ 0.1082 & 0.8905 & 0.0013 \\ 0.0069 & 0.2217 & 0.7714 \end{pmatrix} \\ \mathbf{A}(8, 0) &= \begin{pmatrix} 0.8011 & 0.0960 & 0.1029 \\ 0.0969 & 0.9028 & 0.0003 \\ 0.0046 & 0.1986 & 0.7968 \end{pmatrix} \end{aligned}$$

The path-approach provides additionally a graphical presentation of the evolution of the row vectors  $\mathbf{a}_i(m, 0)$ : The  $i$ -th row of  $\mathbf{A}(m, 0)$  corresponds with the point on the path  $\mathcal{P}_i^0$  with parametric equation  $(f_{i1}^0(t), f_{i2}^0(t), f_{i3}^0(t))$  for  $t = \frac{1}{m} \in [0, 1]$ . For the example under study, the paths  $\mathcal{P}_1^0$ ,  $\mathcal{P}_2^0$  and  $\mathcal{P}_3^0$  are graphically presented in Figure 2. In accordance to the above computed results, the path  $\mathcal{P}_1^0$  lie within the triangle  $\text{Conv}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  for the whole range of  $t$ -values in  $[0, 1]$ , while for the path  $\mathcal{P}_2^0$  the points corresponding with  $t \leq \frac{1}{9}$  (i.e.  $m \geq 9$ ) are lying outside  $\text{Conv}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ .



**Fig. 2** Graphical presentation of the paths  $\mathcal{P}_1^0$ ,  $\mathcal{P}_2^0$  and  $\mathcal{P}_3^0$

## 5 Further research questions

For Markov chains with two states and that satisfy the necessary and sufficient embedding conditions, stochastic roots of the  $(2 \times 2)$  transition matrices are known in analytical form from previous work [2, 5, 7]. For  $(3 \times 3)$  transition matrices with real eigenvalues the embedding problem is discussed in [4–6]. The present paper examines matrix roots and the embedding problem for Markov chains with three states in the case the transition matrix  $\mathbf{P}$  has complex eigenvalues. Necessary embedding conditions are formulated, row-normalized roots  $\mathbf{A}(m, k)$  are constructed and sufficient conditions for the existence of an  $m$ -th stochastic root are presented. In this way, the  $(2 \times 2)$  and  $(3 \times 3)$  case are examined in depth.

For an  $(n \times n)$  transition matrix that is a block diagonal matrix  $P = \text{diag}(P_1, \dots, P_k)$ , the matrix  $A = \text{diag}(A_1, \dots, A_k)$  is an  $m$ -th root of  $P$  if and only if for all  $i \in \{1, \dots, k\}$  holds that  $A_i$  is an  $m$ -th root of  $P_i$ . Therefore, for  $P = \text{diag}(P_1, \dots, P_k)$  with all blocks  $P_i$  of order  $(2 \times 2)$  or  $(3 \times 3)$ , the known insights for the  $(2 \times 2)$  and  $(3 \times 3)$  case are useful in examining matrix roots and the embedding problem. For future research it would be interesting to find out whether the dissemination of the presented approaches could result in insights for matrix roots and for the embedding problem in the case of general  $(n \times n)$  stochastic matrices.



## References

1. Elfving, G.: Zur theorie der markoffschen ketten. *Acta Soc. Sci. Fennicae n. Ser. A* 2. **8**, 1–17 (1937)
2. Guerry, M.: On the embedding problem for discrete-time markov chains. *J. of Appl. Probab.* **50**(4), 918–930 (2013)
3. Guerry, M.: Some results on the embeddable problem for discrete-time markov models in manpower planning. *Commun. in Stat. - Theory and Methods.* **43**(7), 1575–1584 (2014)
4. Guerry, M.: Sufficient embedding conditions for three-state discrete-time markov chains with real eigenvalues. *Linear and Multilinear Algebra.* **67**(1), 106–120 (2019)
5. He, Q.M., Gunn, E.: A note on the stochastic roots of stochastic matrices. *J. of Syst. Sci. and Syst. Eng.* **12**(2), 210–223 (2003)
6. Higham, N., Lin, L.: On  $p$ th roots of stochastic matrices. *Linear Algebra and its Applications.* **435**, 448–463 (2011)
7. Lin, L.: Roots of stochastic matrices and fractional matrix powers [dissertation]. Manchester, UK: The University of Manchester, Manchester Institute for Mathematical Sciences. **MIMS EPrint 2011.9** (2010)
8. Loewy, R., London, D.: A note on an inverse problem for nonnegative matrices. *Linear and Multilinear Algebra.* **6**(1), 83–90 (1978)
9. McDonald, J., Paparella, P.: Matrix roots of imprimitive irreducible nonnegative matrices. *Linear Algebra and its Applications.* **498**, 244–461 (2016)
10. Meyer, C.: Matrix analysis and applied linear algebra. Society for Industrial and Applied Mathematics (2000)
11. Rojo, O., Soto, R.: Existence and construction of nonnegative matrices with complex spectrum. *Linear Algebra and its Applications.* **368**, 53–69 (2003)
12. Singer, B., Spilerman, S.: Social mobility models for heterogeneous populations. in: H.L. Costner (Ed.) *Soc. Methodol.* **5**, 356–401 (1974)
13. Tam, B.S., Huang, P.R.: Nonnegative square roots of matrices. *Linear Algebra and its Applications.* **498**, 404–440 (2016)