

A study *c*oncerning the p*o*sitive semi-definite pr*o*perty f*o*r similarity matrices and f*o*r d*o*ubly st*o*chastic matrices with s*o*me applicati*o*ns

Rafic Nader

▶ **To cite this version:**

Rafic Nader. *A* study *c*oncerning the p*o*sitive semi-definite pr*o*perty f*o*r similarity matrices and f*o*r d*o*ubly st*o*chastic matrices with s*o*me applicati*o*ns. Computer Science [cs]. Université de Caen (France), 2019. English. NNT : . tel-03081681

HAL Id: tel-03081681

https://normandie-univ.hal.science/tel-03081681

Submitted on 18 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Normandie Université

THÈSE

Pour obtenir le diplôme de doctorat

Spécialité INFORMATIQUE

Préparée au sein de l'Université de Caen Normandie

En cotutelle internationale avec Université Libanaise , LIBAN

A study concerning the positive semi-definite property for similarity matrices and for doubly stochastic matrices with some applications

Présentée et soutenue par
Rafic NADER

Thèse soutenue publiquement le 28/06/2019 devant le jury composé de		
Mme MIRIAM PISONERO	Professeur des universités, Dpto Matematica Aplicada	Rapporteur du jury
Mme MARIA ROBBIANO	Professeur des universités, Departamento de Matematicas	Rapporteur non membre du jury
M. HASSANE ABBAS	Professeur des universités, Université Libanaise	Membre du jury
M. JACQUES BENASSENI	Professeur des universités, Université Rennes 2	Membre du jury
Mme ISABELLE BLOCH	Professeur des universités, Université Paris-Saclay	Membre du jury
M. PATRICE BOIZUMAULT	Professeur des universités, Université Caen Normandie	Membre du jury
M. FRANCOIS HENNECART	Professeur des universités, Université Saint-Etienne Jean Monnet	Membre du jury
M. BASSAM MOURAD	Professeur des universités, Université Libanaise	Membre du jury
M. ALAIN BRETTO	Professeur des universités, Université Caen Normandie	Directeur de thèse

Thèse dirigée par ALAIN BRETTO, Groupe de recherche en informatique, image, automatique et instrumentation

UNIVERSITÉ CAEN NORMANDIE

MIIS

GREYC

Acknowledgements

First, I must acknowledge my immense debt of gratitude to my supervisors Prof. Alain BRETTO and Prof. Hassan ABBAS for their excellent guidance, their valuable support and advice throughout my PhD study. It goes without saying that this would have been impossible without them. My deepest feelings of gratitude and gratitude go, of course, to my supervisor, Prof. Bassam MOURAD. I thank him for his help, direction, and support. He was always available to discuss, evaluate, correct and to enrich the work, and this even sometimes during holidays.

I also wish to thank Prof. Miriam PISONERO and Prof. Maria ROBBIANO for kindly accepting the charge to evaluate my work. I am grateful for their insightful comments and suggestions. My thanks go to, Prof. François HENNECART, Prof. Isabelle BLOCH, Prof. Patrice BOIZUMAULT and Prof. Jacques BENASSENI for the honour they gave me to be members of the examination jury.

I would also like to say a general thank you to all of the staff in the GREYC Laboratory-University of Caen and in the EDST-Lebanese University. They have been a pleasure to work with. In particular, I would like to thank all of the administration staff in Caen, who have been extremely helpful on numerous occasions. Last but by no means least, for many reasons, thanks to my family, for their endless support.

Publications

The results in this thesis have previously appeared in the following publications:

1. R. Nader, A. Bretto, B. Mourad, and H. Abbas. On the positive semi-definite property of similarity matrices. Theoretical Computer Science, 755: 13-28, 2019. https://doi.org/10.1016/j.tcs.2018.06.052.

2. R. Nader, B. Mourad, A. Bretto, and H. Abbas. A note on the real inverse spectral problem for doubly stochastic matrices. Linear Algebra and its Applications, 569:206-240, 2019. https://doi.org/10.1016/j.laa.2019.01.017.

3. R. Nader, B. Mourad, H. Abbas, and A. Bretto. On the positive semi-definite pth roots of positive semi-definite doubly stochastic matrices. Submitted to Linear and Multilinear Algebra.

Abstract

Matrix theory has shown its importance by its wide range of applications in different fields such as statistics, machine learning, economics and signal processing. This thesis concerns three main axis related to two fundamental objects of study in matrix theory and that arise naturally in many applications, that are positive semi-definite matrices and doubly stochastic matrices.

One concept which stems naturally from machine learning area and is related to the positive semi-definite property, is the one of similarity matrices. This thesis will explore the latter desirable structure for a list of similarity matrices found in the literature. The importance of characterising such matrices is also discussed. Moreover, we present new results concerning the strictly positive definite and the three positive semi-definite properties of particular similarity matrices.

On the other hand, an interesting research field in matrix analysis involves the study of roots of stochastic matrices which is important in Markov chain models in finance and healthcare. We extend the analysis of this problem to positive semi-definite doubly stochastic matrices. Our contributions include some geometrical properties of the set of all $n \times n$ positive semi-definite doubly stochastic matrices with nonnegative pth roots for a given integer p, denoted by $K_n^{1/p}$. We also present methods for finding elements in $K_n^{1/p}$ by making use of the theory of M-Matrices and the symmetric doubly stochastic inverse eigenvalue problem (SDIEP), which is also of independent interest.

In the context of the SDIEP, which is the problem of characterising those lists of real numbers which are realisable as the spectrum of some symmetric doubly stochastic matrix, we present some new results along this line. In particular, we propose to use a recursive method on constructing doubly stochastic matrices from smaller size matrices with known spectra to obtain new independent sufficient conditions for SDIEP. Finally, we focus our attention on the realizability by a symmetric doubly stochastic matrix of normalised Suleimanova spectra which is a normalized variant of the spectra introduced by Suleimanova.

Table of contents

List of tables **xi**

1 Introduction **1**
- 1.1 Notations . . . 2
- 1.2 Positive semi-definite matrices . . . 2
- 1.3 Nonnegative and doubly stochastic matrices . . . 5
- 1.4 Inverse eigenvalue problem . . . 7
- 1.5 Distance and similarity . . . 8
- 1.6 About this thesis . . . 10

2 Positive semi-definite similarity matrices and some applications **13**
- 2.1 Introduction . . . 13
- 2.2 General definitions . . . 15
- 2.3 Similarity measures for binary data . . . 17
  - 2.3.1 Definition and Notation . . . 17
  - 2.3.2 Similarity measures with nonnegative entries . . . 18
  - 2.3.3 Coefficients of correlation and association . . . 20
- 2.4 Main results concerning PSD matrices . . . 20
- 2.5 A list of PSD similarity matrices . . . 28
  - 2.5.1 Notations . . . 28
  - 2.5.2 Family of PSD similarity matrices . . . 28
  - 2.5.3 Other similarity matrices . . . 30
- 2.6 Positive definiteness of similarity matrices . . . 32
  - 2.6.1 Sufficient condition for the positive definiteness of similarity matrices 33
  - 2.6.2 The complete Tversky similarity matrix is positive definite . . . 34
- 2.7 Three-positive semi-definite similarity matrices . . . 38
- 2.8 Non-PSD similarity matrices . . . 43
- 2.9 Similarity measures for numerical data . . . 44

2.10 Applications 46

3 Positive semi-definite pth roots of positive semi-definite doubly stochastic matrices 51

3.1 Introduction of the problem of interest 51
3.2 Square root of a PSD doubly stochastic matrix 53
3.3 Some geometrical properties of the set $K_n^{1/p}$ 57
3.4 Generating elements in $K_n^{1/p}$ using M-matrices 64
3.5 Constructing elements of $K_n^{1/p}$ via the use of eigenvalues. 68
3.6 A family of $K_n^{1/2}$ via the use of eigenvalues. 75
3.6.1 A method to define $K_3^{1/2}$ 75
3.6.2 Generalization to order n 82

4 The symmetric doubly stochastic inverse eigenvalue problem 87

4.1 Introduction 87
4.2 Preliminairies and overview of some results 88
4.3 Recursive sufficient conditions for SDIEP 92
4.3.1 Inaccurate refinement of Soules 94
4.3.2 An alternative sufficient condition when n is odd 95
4.3.3 The recursive approach to SDIEP 102

5 On the realizability of normalized Suleimanova spectra 117

5.1 Introduction 117
5.2 NS-SDIEP for n odd 118
5.3 Sufficient conditions 120
5.4 Sequences of sufficient conditions for NS-SDIEP 129

6 Conclusions and Futur works 137

6.1 Summary 137
6.2 Futur work 139

Appendix A 141

Appendix B: Synthèse en français 145

References 165

List of tables

2.1 Bivariate counts table for binary variables. 18
2.2 Similarity measures ignoring d 19
2.3 Similarity measures including d 20
2.4 Similarity measures with negative values 21

Chapter 1

Introduction

Matrix theory has developed rapidly in the last few decades because of its wide range of applications and many links with different fields of mathematics, economics, machine learning and signal processing. One of the most desirable structures a matrix can possess is the positive semi-definite property. The class of positive semi-definite matrices is fundamental in matrix analysis and arises naturally in many applications such as data analysis, complex analysis, harmonic analysis, mechanics, and engineering. Moreover, positive semi-definite matrices are used as covariance matrices in statistics, as kernels in machine learning and tensors in medical imaging. In addition, the concept of positive semi-definiteness is of a certain importance when considering similarity matrices [35, 36, 131], which are matrices used to quantify the resemblance of elements of a data space. The mentioned property allows us to construct distance metrics between elements of a data set.

On the other hand, nonnegative matrices form an integral part of matrix theory, started by Oskar Perron [98] and and George Frobenius [33] at the start of the twentieth century. Of special importance of nonnegative matrices is the class of doubly stochastic matrices that is particularly endowed with a rich collection of applications in other areas of mathematics such as regular graph theory, matrix inequalities and majorization, combinatorics, numerical analysis, and also in other areas such as economics, statistics, clustering and the theory of communications [14, 75, 124, 130, 134]. A problem that have been of a special interest in matrix analysis is the study of functions mapping on matrices (i.e trigonometric functions, pth root functions and others [41]). More particularly, finding roots that preserves nonnegativity of nonnegative matrices [73, 74, 89, 125] and recently of stochastic matrices [37, 38] play an important role in both theory and applications. By combining the doubly stochastic property and the positive semi-definite property in some matrix A, it is natural to seek conditions under which the pth root of A, with $p \geq 2$ positive integer, is positive semi-definite doubly stochastic matrix.

In addition to studying the power functions $t \longrightarrow t^{1/p}$ with a given positive integer p, an other intriguing object of study in the theory of doubly stochastic matrices is that of the inverse eigenvalue problem which explores the necessary and sufficient conditions on a complex (or real) prescribed spectrum for the existence of a doubly stochastic matrix with the prescribed spectral data.

In this thesis, we make advances in some related problems and applications in the area of positive semi-definite and doubly stochastic matrices. Before explaining the main three axis of the thesis (Section 1.6), we will give a general introduction to some of the basic facts needed from the theory of positive semi-definite matrices, nonnegative matrices, and doubly stochastic matrices. This introduction is given to provide a convenient repository for all readers. We discuss briefly the material we shall require from these theories and in each section we point the reader to the suitable reference(s).

1.1 Notations

First let us give our notation used in the whole thesis. We denote by $\mathbb{C}^{n\times n}$ (respectively $\mathbb{R}^{n\times n}$) the set of all $n\times n$ complex matrices (respectively $n\times n$ real matrices). A matrix A of dimension n with real entries is said to be nonnegative (respectively positive) if all of its entries are nonnegative (respectively positive). In this case, we write $A \geqslant 0$ (respectively $A > 0$).

We say that C is a principal submatrix of $A \in \mathbb{C}^{n\times n}$, if C is a square submatrix obtained from A by deleting some rows and columns indexed by the same set of indices. The determinant of an $r\times r$ submatrix of A is called a minor of size r.

For any matrix (or vector) A, we denote by A^* its conjugate transpose and by A^T its transpose. $A\oplus B$ will denote the direct sum of A and B, for any $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$.

Finally, let $\mathrm{diag}(1,\lambda_2,\lambda_3,...,\lambda_n)$ be the $n\times n$ diagonal matrix with diagonal entries $1,\lambda_2,\lambda_3,...,\lambda_n$. Let I_n and J_n be the $n\times n$ identity matrix and the $n\times n$ matrix whose all entries are $\frac{1}{n}$ respectively. Also, let e_n be the vector of dimension n with every entry equal to $\frac{1}{\sqrt{n}}$, i.e. $e_n = \frac{1}{\sqrt{n}}(1,...,1)^T$.

1.2 Positive semi-definite matrices

In this section, we will give the briefest of summaries on the theory of positive semi-definite matrices.

Definition 1.2.1. A matrix $A \in \mathbb{C}^{n\times n}$ ($A = A^*$) is said to be positive semi-definite (PSD), denoted by $A \succeq 0$, if

$$x^*Ax \geq 0$$

for all $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{C}^n$. Furthermore, A is said to be positive definite (PD) denoted by $A \succ 0$, if

$$x^*Ax > 0,$$

for all $x \neq 0 \in \mathbb{C}^n$.

Remark 1. Clearly, a symmetric matrix $A \in \mathbb{R}^{n\times n}$ is called positive semi-definite if $x^TAx \geq 0$ for all $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$, and is called positive definite, if $x^TAx > 0$ for all nonzero $x \in \mathbb{R}^n$.

From Definition 1.2.1 , we conclude the following simple observation, for which the proof can be found in [43, page 430]

Observation 1.2.2. *Let $A \in \mathbb{C}^{n\times n}$ be positive semi-definite matrix, then all of its principal submatrices are positive semi-definite. Moreover, if A is positive definite, then all of its principal submatrices are positive definite.*

Moreover, positive semi-definite matrices can be characterised in many different ways. We shall make frequent use in Chapter 2 of the following characterisations.

Theorem 1.2.3. *[43] A Hermitian matrix is positive semi-definite if and only if all of its eigenvalues are nonnegative. It is positive definite if and only if all of its eigenvalues are positive.*

This leads us to the following corollary.

Corollary 1.2.4. *[43] If $A \in \mathbb{C}^{n\times n}$ is positive semi-definite, then so is each A^k, $k = 1, 2, ...$*

The following theorem gives a characterization using the determinant of principal submatrices of a Hermitian matrix.

Theorem 1.2.5. *(Sylvester's criterion)[43]. Let $A \in \mathbb{C}^{n\times n}$ be Hermitian. If every principal minor of A (including det A) is nonnegative, then A is positive semi-definite.*

Next, we give the the Schur decomposition [139, Theorem 3.3] and the spectral decomposition [139, Theorem 3.4] of matrices that play a critical role in matrix theory.

Theorem 1.2.6. *Schur Decomposition [139]*
Let $\lambda_1, \lambda_2, ..., \lambda_n$ be the eigenvalues of $A \in \mathbb{C}^{n\times n}$. Then there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that U^*AU is an upper-triangular matrix, i.e

$$\begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

Theorem 1.2.7. *Spectral Decomposition[139]*
Let A be a matrix in $\mathbb{C}^{n\times n}$ with eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$. Then A is normal if and only if there exists a unitary matrix U such that

$$U^*AU = diag(\lambda_1, \lambda_2, ..., \lambda_n).$$

Moreover, A is Hermitian if and only if the λ_i are all real. A is positive semi-definite if and only if the λ_i are all nonnegative and A is positive definite if and only if the λ_i are all positive.

This leads us to the real version of spectral decomposition for symmetric matrices that we will need for our purposes in Chapter 3. For convenience, we will state it in the following (see [43, Corollary 2.5.11]).

Theorem 1.2.8. *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix with eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$. Then, there exits a real orthogonal matrix P such that*

$$P^TAP = diag(\lambda_1, \lambda_2, ..., \lambda_n).$$

The spectral decomposition is used to define the pth root of A, denoted by $A^{1/p}$ for $p \geq 2$. In fact, every nonnegative number has a unique nonnegative pth root for $p \in \{2, 3, ...\}$. An analogous result for positive semi-definite matrices holds as we can see in the following theorem which the proof can be found in [43, Theorem 7.2.6].

Theorem 1.2.9. *Let $A \in \mathbb{C}^{n\times n}$ be positive semi-definite and let $p \in \{2, 3, ...\}$. Then there exists a unique positive semi-definite matrix B such that $B^p = A$.*

Next, we will define a way to construct a positive semi-definite matrix by considering an inner product space. Let $x_1, ..., x_n$ be vectors in an inner product space X with inner product $< ., . >\cdot$ Then, the Gram matrix of the vectors $x_1, ..., x_n$ with respect to the inner product $< ., . >$ is $M = (< x_i, x_j >)_{1\leq i,j\leq n}$. Therefore, we have the following theorem which can be found in [43, Theorem 7.2.10].

Theorem 1.2.10. *[43] Let $x_1,...,x_n$ be vectors in an inner product space $\mathcal{X}$ with inner product $<.,.>$, and let $M=(<x_i,x_j>)_{i,j}\in\mathbb{C}^{n\times n}$. Then,*

- *M is positive semi-definite*
- *M is positive definite if and only if the vectors $x_1,...,x_n$ are linearly independent.*

Moreover, every positive semi-definite matrix can be written as a Gram matrix. In fact, consider a positive semi-definite matrix $A\in\mathbb{C}^{n\times n}$. Partition $A^{1/2}=[x_1,...,x_n]$ according to its columns and notice that $A=A^{1/2}A^{1/2}=(A^{1/2})^*A^{1/2}=(<x_i,x_j>_E)_{i,j}$, in which $<.,.>_E$ is the Euclidean inner product on $\mathbb{C}^n$. Hence, A is a Gram matrix.

Next, we give the following definition.

Definition 1.2.11. Let $A=[a_{ij}]$ and $B=[b_{ij}]\in\mathbb{C}^{n\times n}$ then their *Hadamard product* (or *Schur product*) denoted by $A\circ B$ is the matrix whose entries are given by $(A\circ B)_{i,j}=a_{i,j}b_{i,j}$.

Next, we state the Schur product theorem [43, Theorem 7.5.3] concerning the positive semi-definitness of the Schur product of two positive semi-definite matrices.

Theorem 1.2.12. *[43] Let $A,B\in\mathbb{C}^{n\times n}$ be two positive semi-definite matrices (respectively positive definite matrices). Then, $A\circ B$ is also positive semi-definite (respectively positive definite).*

Finally, we end our section, with the following definitions that we will need for our purposes.

Definition 1.2.13. A Hermitian matrix $A\in\mathbb{C}^{n\times n}$ is said to be two-positive (resp. three-positive) semi-definite if all of its principal submatrices of size two (resp. three) are positive semi-definite.

Definition 1.2.14. Let $A,B\in\mathbb{C}^{n\times n}$. We write $A\succeq B$ (respectively $A\succ B$) if A and B are Hermitian and $A-B$ is positive semi-definite (respectively positive definite).

Excellent accounts of the theory of positive semi-definite matrices can be found in [8, 43, 139].

1.3 Nonnegative and doubly stochastic matrices

This section will give a general introduction to some of the basic facts of the theory of nonnegative matrices and doubly stochastic matrices. For more details on this subject, see Berman and Plemmons [7], Minc [77] and Bapat and Raghavan [4]. First, we recall the following definitions.

Definition 1.3.1. Let X and Y be two real square matrices. X is said to be cogredient to Y if there exists a permutation matrix Q such that $X = Q^T Y Q$.

Definition 1.3.2. A nonnegative $n \times n$ matrix $A(n > 2)$ is called reducible if it is cogredient to a matrix of the form $\begin{pmatrix} B & D \\ 0 & C \end{pmatrix}$ where B and C are square matrices. Otherwise, A is irreducible.

The Perron-Frobenius theory of nonnegative matrices (see [77, Chapter 1] and [43, Chapter 8]) can be summarised as follows.

Theorem 1.3.3. *Let A be an $n \times n$ nonnegative matrix and let $\rho(A)$ be the spectral radius of A. Then $\rho(A)$ is an eigenvalue of A, called the Perron eigenvalue of A, with a corresponding nonnegative eigenvector, called the Perron eigenvector of A. In addition, if A is irreducible, then*

- $\rho(A) > 0$
- *there is an $x > 0$ such that $Ax = \rho(A)x$.*
- $\rho(A)$ *is a simple eigenvalue of A (that is, it has algebraic multiplicity 1).*

A class of nonnegative matrices with important applications in many areas of mathematics and informatics is the class of doubly stochastic matrices.

Definition 1.3.4. A real square matrix is said to be doubly quasi-stochastic if each of its row and column sum is equal to 1. A nonnegative doubly quasi-stochastic matrix is called doubly stochastic.

It follows that a nonnegative matrix A of size n is doubly stochastic if and only if

$$Ae_n = e_n$$

and

$$e_n A^T = e_n$$

or equivalently

$$AJ_n = J_n A = J_n.$$

From this characterization, we can conclude that the product of doubly stochastic matrices is doubly stochastic, and in particular, any power of a doubly stochastic matrix is also doubly stochastic.

Finally, we should note that doubly stochastic matrices represent a special subset of stochastic ones.

Definition 1.3.5. A nonnegative $n \times n$ matrix A is said to be a stochastic (or row stochastic) matrix if each row sum is equal to 1. A stochastic matrix A such that A^T is also stochastic is said to be doubly stochastic.

1.4 Inverse eigenvalue problem

The nonnegative inverse eigenvalue problem (NIEP) asks the necessary and sufficient conditions for the components in an n-tuplets, also called list, $\sigma = (\lambda_1, ..., \lambda_n)$ in $\mathbb{C}^n$ to be the spectrum of an $n \times n$ nonnegative matrix A. Although the nonnegative inverse eigenvalue problem has been studied extensively, it is still unsolved for $n \geq 5$.

We say that $\sigma = (\lambda_1, ..., \lambda_n)$ in $\mathbb{C}^n$ is realizable if there is an $n \times n$ nonnegative matrix A with spectrum $\{\lambda_1, ..., \lambda_n\} \subset \mathbb{C}^n$. Then A is called a realising matrix for σ . For any $n \times n$ nonnegative matrix A with spectrum $\{\lambda_1, .., \lambda_n\}$, and for any positive integer k, we denote the trace of A^k by

$$s_k(\sigma) = \lambda_1^k + ... + \lambda_n^k.$$

Some well known necessary conditions for a list $\sigma = (\lambda_1, ..., \lambda_n)$ to be realizable by an $n \times n$ nonnegative matrix A are:

1. $\rho(\sigma) = max|\lambda_i| \in \sigma$, by Perron-Frobenius ' theorem.
2. $s_k(\sigma) \geq 0$, $\forall k \in \mathbb{N}$ as A^k is also a nonnegative matrix.
3. σ is closed under complex conjugation, i.e. $\overline{\sigma} := (\bar{\lambda_1}, \bar{\lambda_2}, ..., \bar{\lambda_n}) = \sigma$.
4. $s_k(\sigma)^m \leq n^{m-1} s_{km}(\sigma)$ for all positive integers k and m.

Condition (3) follows from the fact that the characteristic polynomial of A has real coefficients. The inequalities in (4) are proved by Loewy and London [64], and independently by Johnson [50]. They are known by the JLL conditions.

When σ is a list of n real numbers, we have then the following two problems:

- The real nonnegative inverse eigenvalue problem (RNIEP) that asks which lists of n real numbers are realizable by a $n \times n$ nonnegative matrix A.
- The symmetric nonnegative inverse eigenvalue problem (SNIEP) that asks which lists of n real numbers are realizable by an $n \times n$ symmetric nonnegative matrix A.

If there exists a nonnegative symmetric matrix A with spectrum σ, then σ is said to be symmetrically realizable by A. A large body of work on the SNIEP / RNIEP can be found in

the literature, giving several different sufficient conditions[68–70]. Many partial results for the three problems are known. We encourage the reader to see also Johnson et al. [51] and the references within for an extensive survey on the NIEP.

These problems become more challenging when extra properties are imposed to the nonnegative matrices. For example, the realizabilty of the spectra by nonnegative circulant matrices is studied in [104], by nonnegative integral matrices in [55], by companion matrices in [57] and recently by permutative matrices in [3, 63, 66, 92]. Indeed, when NIEP involves doubly stochastic matrices, we are talking about the doubly stochastic inverse eigenvalue problem denoted by (DIEP), which is the problem of determining the necessary and sufficient conditions for a complex n-tuple to be the spectrum of an $n \times n$ doubly stochastic matrix. When σ is a list of real numbers the problem of determining necessary and sufficient conditions for σ to be the spectrum of a doubly stochastic matrix is called the real doubly stochastic inverse eigenvalue problem (RDIEP). When we are searching for a symmetric doubly stochastic matrix to be the the realising matrix of a list of real numbers, we are talking about the symmetric doubly stochastic inverse eigenvalue problem (SDIEP). For convenience, we will define it explicitly in the following.

Problem 1. *The symmetric doubly stochastic inverse eigenvalue problem (SDIEP) asks which sets of n real numbers can be the spectrum of an $n \times n$ symmetric doubly stochastic matrix.*

All the above problems have drawn considerable interest but all the results are partial [50, 76, 78, 85, 97, 96, 95, 121, 122]. In this thesis, we will be dealing only with SDIEP. So far, SDIEP have only been solved for the case $n = 3$ by Perfect and Mirsky [97] and remains open for the cases $n \geq 4$ (see [45, 50, 58, 79, 81, 82, 85, 97, 99, 122] for a collection of most sufficient conditions for the SDIEP). In Chapter 4, we mention several sufficient conditions for the SDIEP, and we give some new results of this type.

1.5 Distance and similarity

In this section, we collect some useful definitions concerning distances.

Definition 1.5.1. Let $\mathcal{X}$ be any set. A function $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is called a *metric*, if for any $x, y, z \in \mathcal{X}$, it satisfies the following conditions:

(1) $d(x,y) \geq 0$ (non-negativity),

(2) $d(x,y) = d(y,x)$ (symmetry),

(3) $d(x,y) \leq d(x,z) + d(z,y)$ (triangle inequality),

(4) $d(x,y) = 0$ if and only if $x = y$ (identity of indiscernibles).

There are some variations of a metric function, which satisfy weaker properties. Property (4) can be split into (4)' and (4)", the properties of:

- (4)' Reflexivity : $d(x,x) = 0$ for all $x \in \mathcal{X}$,
- (4)" Separability: $d(x,y) = 0 \Longrightarrow x = y$.

If d verifies the Conditions (1), (2), (3) and (4)' , then d is called a *pseudo-metric*. If d satisfies only the Conditions (1), (2) and (4)', then d is called a *dissimilarity* .

Next, we provide some examples of distance functions, which can be applied to numerical data:

1. Euclidean: $d(x,y) = \sum_{i=1}^{p} \sqrt{(x_i - y_i)^2}$,
2. Weighted Euclidean: $d(x,y) = \sum_{i=1}^{p} \sqrt{\alpha_i (x_i - y_i)^2}$,
3. Mahalanobis: $d(x,y) = \sqrt{(x-y)^T \Sigma^{-1} (x-y)}$,

where p denotes the data dimension, $x = (x_i)_{i=1,\ldots,p}$ and $y = (y_i)_{i=1,\ldots,p}$ two data points belonging to $\mathbb{R}^p$, $(\alpha_i)_{i=1,\ldots,p}$ a vector of positive weights, and Σ the data covariance matrix.

Complementary to dissimilarity (or distance) is the concept of similarity. As an alternative to a dissimilarity function, we can define a similarity measure. A dissimilarity is a measure that quantify how different two objects are. A high value of a dissimilarity indicates that the objects are very different and a small value indicates that they are close. For a similarity coefficient the interpretation is the other way around. A high value indicates great resemblance between the objects whereas a low value indicates great dissimilarity.

From this interpretation, we can define similarly a similarity measure as a mapping from $\mathcal{X} \times \mathcal{X}$ to $\mathbb{R}$ as follows.

Definition 1.5.2. A *similarity* is a mapping S from $\mathcal{X} \times \mathcal{X}$ into $\mathbb{R}$ such that:

1. $S(x,y) \geq 0$ for all $x, y \in \mathcal{X}$,
2. $S(x,y) = S(y,x)$ for all $x, y \in \mathcal{X}$,
3. $S(x,y) \leq S(x,x)$ for all $x, y \in \mathcal{X}$.

Note that we will see more about types of similarity measures and their properties in Chapter 2. Similarity measures quantify how much two objects are alike. However, when we are confronted to more than two objects, we will define a similarity matrix. As we will see in Chapter 2, such matrix is real symmetric. Our purpose is to prove that it is also positive semi-definite for a large number of similarity measures found in the literature because this property is very important to derive distances between data set in many applications.

For an extensive survey and comparison of various known distances and similarities, we encourage the reader to consult [23].

1.6 About this thesis

We address three main issues in this thesis which are the positive semi-definite property of similarity matrices, the positive semi-definite pth roots of positive semi-definite doubly stochastic matrices, and the symmetric doubly stochastic inverse eigenvalue problem.

Motivated by its widespread applications, our work starts with studying the positive semi-definite property of similarity matrices. Similarity matrices are fundamental objects of applied sciences [19]. They can be used to measure the resemblance between objects of a data set. Moreover, PSD similarity matrices are of a special importance by their ability to define distance metrics. However, few similarity matrices are proved to be positive semi-definite [35, 36]. Chapter 2 is concerned with the analysis of the properties of similarity matrices and more specifically the positive semi-definite property. We start by defining similarity measures and matrices, then we collect the most frequent ones found in the literature. Using a different approach on defining similarity matrices as theoretical operations between subsets of a finite set, we will prove that many useful similarity matrices are positive semi-definite. This property as we will show, and other stronger and weaker properties such as the positive definite and the three-positive semi-definite properties (see Sections 2.6 and 2.7), will pave the way to define new distance metrics and pseudo-metrics between data set, which is important in many applications of machine learning. That we will see more in details in the final section of Chapter 2.

One particular type of similarity matrix can be a doubly stochastic matrix. Beside their applications in combinatorics, the doubly stochastic matrices which represent a special subset of stochastic matrices, are a very useful tool in probability and statistics. In fact, a stochastic matrix $B = (b_{ij})$ can be thought of as a transition matrix of a Markov chain where b_{ij} is the probability of going from state i to j. They are often used to model the behaviour of evolutionary systems that are encountered in the analysis of queuing networks and computer systems, discrete economic models and in many stochastic models found in biological and

social science. One of many problems that is studied in the theory of stochastic matrices, is finding the pth roots of a stochastic matrix [37, 38] as the computation of such roots has led to many applications in the area of financial mathematics and chronic diseases. Motivated by their many applications, we extend the analysis to the doubly stochastic case and in particular to PSD doubly stochastic matrices. Our purpose in Chapter 3, is to characterise conditions under which the pth root of a doubly stochastic positive semi-definite matrix, is doubly stochastic. More particularly, we study some geometrical properties of the set of all $n \times n$ positive semi-definite doubly stochastic matrices with nonnegative pth root for any integer p, denoted by $K_n^{1/p}$. Moreover, we construct elements that belongs to $K_n^{1/p}$ via the use of eigenvalues and the theory of M-matrices.

In addition, we strongly believe that there is a close connection between finding nonnegative roots of PSD doubly stochastic matrices and SDIEP. More precisely, any solution of SDIEP should in principle lead to a new algorithm for a solution in finding nonnegative pth roots of doubly stochastic PSD matrices (see Section 3.6). Motivated by the close relation between the problem of finding roots and the inverse eigenvalue problem, the third part of the thesis is devoted to the study of the symmetric doubly stochastic inverse eigenvalue problem (SDIEP). In Chapter 4, we focus on deriving new sufficient conditions for a list of real numbers to be the spectrum of a doubly stochastic matrix using a recursive method of constructing new doubly stochastic matrices from old ones. By many examples and observations, we will show how these new sufficient conditions can be of a great importance in finding new independent partial solutions for SDIEP. Finally, Chapter 5 is devoted to study a special case of SDIEP, the normalised Suleimanova symmetric doubly stochastic inverse eigenvalue problem that asks which sets of a normalised Suleimanova spectrum occurs as a spectrum of a symmetric doubly stochastic matrix.

Chapter 2

Positive semi-definite similarity matrices and some applications

2.1 Introduction

The notion of similarity is a fundamental concept in many research areas and applications. In practice, similarities are evaluated by a measure that quantifies the extent to which objects, or variables, resemble one another. We can distinguish different measures according to the type of data they apply to, i.e binary, numerical or structured data. A similarity measure gives rise to a similarity matrix when considering pairs of variables or sample units. This chapter is concerned with the analysis of the properties of binary similarity matrices. More specifically, we focus on their positive semi-definite property, which is important to derive useful distances between data sets [11, 36, 127]. In fact, a similarity coefficient is used to quantitatively measure the resemblance between elements of a data set in many applications fields such as information integration [5], theory of hypergraphs [9, 10], image analysis [15], biology [44], or more generally in pattern analysis problem [109]. However, for some applications like clustering and pattern recognition, it is useful to have dissimilarities and more particularly metric distance [23]. The simplest way to derive a dissimilarity coefficient from a similarity measure s is to apply a decreasing function on s. Moreover, many similarity measures will become a metric, or even Euclidean, if transformed according to $d = \sqrt{1-s}$. In fact, Gower and Legendre [36] proved that if S is a positive semi-definite similarity matrix, then the dissimilarity matrix given by $D = \sqrt{1-S}$ is Euclidean.

Most machine learning algorithms rely on the development of an effective distance metric based on the concept of similarity. For example, we will cite the K-nearest neighbour (KNN) classifier [21] that needs a suitable distance metric, through we can identify neighbouring

data points. The most frequently used distance to compare numerical data is the Euclidean distance that assumes that each feature of the data points is equally important and independent from others. This assumption may not be always satisfied in real applications. A good distance metric with good quality should identify relevant from irrelevant features. However, one can consider a more general linear transformation of the data, defined by $d_A(x,y) = \sqrt{(x-y)^T A(x-y)}$ where $x,y \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m\times m}$ is a positive semi-definite matrix. For instance in Statistics, an effective method to determine the similarity between data sets is the Mahalonobis distance. This distance ($d_{\sigma^{-1}}$), thanks to the inverse of the covariance matrix σ, takes into account correlations between data. It may be generalised by replacing the correlation matrix by another matrix, whose role is to integrate information on the sets under study [136]. To define this type of metrics (respectively pseudo-metrics), we need PD matrices (respectively PSD matrices). The similarity matrices are very good candidates for this. In addition, PSD similarity matrices are very popular since they can also be used as kernels, in the now popular kernel-based classification methods [20, 110], such as *Support Vector Machines*. The positive semi-definiteness of kernels is essential to ensure that the kernel algorithms, like SVMs, converge.

Motivated by the importance of the positive semi-definite property of similarity matrices in many applications, we will focus of this chapter on binary similarity measures that have been proposed in various fields [23, 36, 109] and play a critical role in pattern analysis problems such as classification and clustering. While few similarity matrices are proved to be positive semi-definite in [35, 36], we will use a different approach on defining similarity matrices as theoretical operations between subsets of a finite set to show that most known binary similarity matrices are positive semi-definite. This chapter is also designed to give answers for the following questions.

1. Under what conditions, PSD similarity matrices are strictly positive definite?

2. When we are unable to prove that a particular similarity matrix is PSD, can we obtain a weaker conclusion such as the three-positive semi-definite property ?

3. Why is it important for a similarity matrix to be positive semi-definite (and definite)?

But first, we will give some basic information about similarity coefficients in Section 2.2. In Section 2.3, we will collect many binary similarity measures defined in the literature that we will consider in the rest of this chapter. In Section 2.4, general results concerning PSD matrices are proven from which we can then derive that many similarity matrices are indeed positive semi-definite in Section 2.5. In Section 2.6, the strictly positive definite condition of similarity matrices is investigated . In Section 2.6.1, we give a sufficient condition for all

positive semi-definite similarity matrices introduced in Section 2.5 to be positive definite. Then, the positive definiteness of the Tversky's family of similarity matrices is proven in Section 2.6.2. Section 2.7 deals with giving a sufficient condition for a similarity matrix to be three-positive semi-definite via the use of the formal definition of a similarity metric [18]. For a complete study of the known similarity coefficients, regarding the positive semi-definite property, Section 2.8 gives examples of non-PSD similarity matrices. Section 2.9 explains how to generalise the results on the similarity matrices taken from binary data to numerical data. Finally, in Section 2.10, we illustrate the importance of the positive semi-definite and definite property in many applications of machine learning.

2.2 General definitions

Informally, similarity measures (also named similarity coefficients or resemblance measures) are functions that quantify the extent to which objects resemble one another. They are, in some sense, the inverse of distance metrics. They take large values for similar objects and either zero or negative values for dissimilar objects.

Although there do not exist a single unified definition of a similarity measure, certain criteria for defining similarity measures are more useful than others; especially in the applications. Denoting $\mathcal{X}$ the data space, a similarity measure is usually defined as follows.

Definition 2.2.1. A *similarity* S is a mapping from $\mathcal{X} \times \mathcal{X}$ into $\mathbb{R}$ such that:

1. $S(x,y) \geq 0$ for all $x,y \in \mathcal{X}$,
2. $S(x,y) = S(y,x)$ for all $x,y \in \mathcal{X}$,
3. $S(x,y) \leq S(x,x)$ for all $x,y \in \mathcal{X}$.

Condition 1 states that the similarity between any two elements x and y is nonnegative. Condition 2 states that $S(x,y)$ is symmetric. Condition 3 states that for any x, the self similarity is no less than the similarity between x and any y. Other properties can be required, as for instance it may be necessary to normalise such a measure to take values in the interval $[0,1]$ which essentially amounts to taking $S(x,x) = 1$ for all $x \in \mathcal{X}$.

However, for some similarity measures defined in the literature (mostly known as coefficients of association or correlation), the positivity constraint is rejected. Such coefficients measure the strength of the relationship between the two variables in the interval $[-1,1]$. The two extreme values reveal maximum strength differing only in direction and the value of the similarity coefficient is zero in the absence of association between the two variables. For this

purpose, we can adopt a more general definition of a similarity measure that does not verify the positivity condition as follows.

Definition 2.2.2. A mapping R from $\mathcal{X} \times \mathcal{X}$ into $\mathbb{R}$ is said to be a similarity if R satisfies the following:

a) $R(x,y) = R(y,x)$ for all $x, y \in \mathcal{X}$,

b) $R(x,x) \geq 0$ for all $x \in \mathcal{X}$,

c) $|R(x,y)| \leq R(x,x)$ for all $x, y \in \mathcal{X}$.

R is said to be a real normalized similarity measure if $R(x,x) = 1$ for all $x \in \mathcal{X}$.

The similarity measures can determine the relationships between two variables (or two objects). However, under most circumstances, we are confronted with more than two variables. For this purpose we define the similarity matrix for any data set as follows.

Definition 2.2.3. Let O_1, O_2,...,O_m be m elements of a data set $\mathcal{X}$. And let $S : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$ be a similarity measure defined on the data set $\mathcal{X}$. Then a similarity matrix between O_1, O_2,...,O_m is defined as

$$M_S = \begin{pmatrix} S(O_1,O_1) & S(O_1,O_2) & .... & S(O_1,O_m) \\ S(O_2,O_1) & S(O_2,O_2) & .... & S(O_2,O_m) \\ \vdots & \vdots & \ddots & \vdots \\ S(O_m,O_1) & S(O_m,O_2) & .... & S(O_m,O_m) \end{pmatrix},$$

where $S(O_i,O_j)$ represents the similarity value between O_i and O_j.

Since $S(O_i,O_j) = S(O_j,O_i)$ for all $i = 1,...,m$, then the similarity matrix M_S is a nonnegative symmetric matrix. Moreover, if S is a normalised similarity measure (i.e $S(O_i,O_i) = 1$ for all i), then, M_S is a nonnegative symmetric matrix with 1 as its diagonal elements. Note that if we consider a real normalised similarity measure R, then the corresponding similarity matrix M_R is symmetric with 1 as its diagonal elements. However, M_R can have negative entries.

The main objective of this chapter is to prove that for many similarity measures found in the literature, the corresponding similarity matrices are PSD. Many similarity coefficients have been proposed for different type of data : binary, numerical and structural data, and in various fields. Moreover, a same similarity measure may have been proposed by several authors independently, with different names. In this thesis, we focus on the case of binary data, also called set data, i.e., data represented by the presence or absence of characteristics.

We will exploit more about binary similarity measures in the following section. First, let us give the following notation.

Notation. We will use S as a general symbol for a similarity coefficient with positive values and R as a general symbol for a similarity coefficient with real values.

2.3 Similarity measures for binary data

A large variety of data can be represented by binary variables, which express two possible states of the data, i.e. presence/absence, true/false, yes/no, male/female. Generally speaking, a variable corresponds to an object or individual and the binary scores reflect the presence or absence of certain characteristics of the object. For example, in archaeology, binary data may denote if certain species types is found or not in a specific location; in psychology, binary data may denote if a person has a specific psychological behaviour. As the binary representation is one of the most common representations of patterns, many binary similarity measures were proposed to determine the amount of agreement between variables in many problems of clustering, classification and data analysis. For example, the binary similarity measure were applied in ecology [47], taxonomy [117], ethnology [25], image retrieval [115] and geology [42]. We are not concerned here with recommending what coefficients should be used on what circumstances neither doing a comparative study collecting the wide variety of binary similarity measures. For more details on grouping different existing binary similarity measures, we encouraged the interested reader to consult [19, 44, 128, 48].

2.3.1 Definition and Notation

We are interested in the case of $\mathcal{X} = \{0,1\}^n$, i.e. data described by n binary scores. Such data can encode data described by the set of present characteristics, from a predefined list. For each characteristic, a score is defined, that takes value 1 if the object possesses the characteristic, and 0 otherwise.

Given two objects described by two variables $x = (x_1,...,x_n)$ and $y = (y_1,...,y_n)$ both belonging to $\mathcal{X}$, the binary similarity measures are commonly calculated using the four dependant counts a, b, c and d presented in Table 2.1, called the 2×2 contingency table .

In Table 2.1,

- $a =$ the number of 1 's that the variables share in the same positions, meaning the number of attributes common to x and y.

Table 2.1 Bivariate counts table for binary variables.

Variable 1	Variable 2		
	value 1	value 0	Total
value 1	a	b	$a+b$
value 0	c	d	$c+d$
Total	$a+c$	$b+d$	n

- b =the number of 1 's in the first variable and 0's in the second variable in the same positions, meaning the number of attributes present in x but not in y.
- c = the number of 0 's in the first variable and 1's in the second variable in the same positions, meaning the number of attributes present in y but not in x.
- d = the number of 0 's that the variables share in the same positions, meaning the number of attributes in neither x nor y .

Clearly, $n = a+b+c+d$ is the total number of features (attributes) or dimension of each binary vector. The diagonal sum $a+d$ represent the total number of matches between x and y, and the other diagonal sum $b+c$ represents the total number of mismatches between x and y.

2.3.2 Similarity measures with nonnegative entries

Since binary data are very common, the literature abounds in coefficients developed specifically for binary variables (see [23] and [131]). In this section, we will make a list of many binary nonnegative similarity measures. The problem that arises when selecting a binary similarity measure to compare our data objects is whether we should include the value d, which is the number of double zeros in the variables. We can therefore wonder if d influences the comparisons? Sokal and Sneath [116, 117] among others, make distinction between coefficients that do and those that do not include the quantity d. When we are investigating the presence or absence of a list of characteristics, then d reflects the number of negatives matches, which is generally felt not to contribute to resemblance. Indeed, an infinite number of characteristics may be missing in both objects. However, in a piece of data where the two binary states are of equal importance, d can be as important as a. For example, if the binary data encodes a nominal variable (female/male, vertebrate/invertebrate...), the a reflects the number of matches on the first attribute and d the number of matches on the second one. In such cases, d is as meaningful as a, and coefficients that involves d should be used. Hence, no general rule can be given as to the incorporation of d in a similarity measure. It depends

on the data itself, the nature of the study and the intended type of analysis.

Similarity measures ignoring d

Table 2.2 recalls binary similarity measures ignoring d that we will consider in this chapter.

Table 2.2 Similarity measures ignoring d

Similarity measure	Notation	Definition
Jaccard [47]	S_{Jac}	$\frac{a}{a+b+c}$
Gleason [34]	S_{Gleas}	$\frac{2a}{2a+b+c}$
Sorensen [119]	S_{Sor}	$\frac{4a}{4a+b+c}$
Anderberg [2]	S_{And}	$\frac{8a}{8a+b+c}$
Sneath and Sokal 2 [116]	S_{SS2}	$\frac{a}{a+2b+2c}$
Ochiai [90]	S_{Och}	$\frac{a}{\sqrt{a+b}\sqrt{a+c}}$
Kulczynski 2 (1927)	S_{Ku2}	$\frac{1}{2}(\frac{a}{a+b}+\frac{a}{a+c})$
Braunn-Blanquet [13]	S_{BB}	$\frac{a}{\max(a+b,a+c)}$
Simpson [111]	S_{Sim}	$\frac{a}{\min(a+b,a+c)}$
Forbes [32]	S_{Forbes}	$\frac{n.a}{(a+b)(a+c)}$
Sorgenfrei [120]	S_{Sorg}	$\frac{a^2}{(a+b)(a+c)}$
Faith [28]	S_{Fai}	$\frac{2a+d}{2n}$

It is worthy to mention here that the name of the above similarity coefficients can vary according to authors. To each similarity measure, we point out the reader to a suitable reference.

Similarity measures including d

Table 2.3 recalls main binary similarity measures including d that will be our elements of study in this chapter.

Table 2.3 Similarity measures including d

Similarity measure	Notation	Definition
Rogers and Tanimoto [101]	S_{RT}	$\frac{a+d}{a+2(b+c)+d}$
Sokal and Michener [118]	S_{SM}	$\frac{a+d}{a+b+c+d}$
Sokal and Sneath 1 [117]	S_{SS1}	$\frac{a+d}{a+d+\frac{1}{2}(b+c)}$
Russel and Rao [105]	S_{RR}	$\frac{a}{a+b+c+d}$
Sokal and Sneath 3 [117]	S_{SS3}	$\frac{ad}{\sqrt{(a+b)(a+c)(b+d)(c+d)}}$
Rogot and Goldberg [102]	S_{RG}	$\frac{a}{a+b+a+c}+\frac{d}{b+d+c+d}$
Hawkins and Dotson [40]	S_{HD}	$\frac{1}{2}(\frac{a}{a+b+c}+\frac{d}{b+c+d})$
Doolittle [24]	S_{Doo}	$\frac{(ad-bc)^2}{(a+b)(a+c)(b+d)(c+d)}$
Pearson [93]	$S_{Pearson}$	$\frac{n(ad-bc)^2}{(a+b)(a+c)(b+d)(c+d)}$

2.3.3 Coefficients of correlation and association

In this section, we consider a list of similarity measures that may have negative values as shows Table 2.4. Most of them have the covariance in the numerator. The covariance between two binary variables is given by $(ad-bc)$. Coefficients with quantity $(ad-bc)$ are mostly known as coefficients of correlation and association. Most coefficients of association and correlation measure the strength of the relationship in the interval of $[-1,1]$, an exception being the Dispersion similarity matrix. For example, if we consider the most popular coefficient of correlation, Phi similarity coefficient, a value of 0 indicates no relationship between x and y, whereas a value of 1 represents the perfect positive correlation. It occurs when x and y perfectly coincide. However -1 is the perfect negative correlation. It occurs when x and the negative of y perfectly coincide.

2.4 Main results concerning PSD matrices

In this section, we are mainly concerned with proving that certain classes of matrices are PSD. This in turn will be the basis for proving that a number of similarity matrices are positive semi-definite in Section 2.5. First, we will exploit some standard PSD matrices before constructing more original ones basically built for our purposes.

Table 2.4 Similarity measures with negative values

Similarity measure	Notation	Definition
Dispersion similarity [23]	R_{Dis}	$\frac{ad-bc}{n^2}$
Phi (Yule [138], Pearson and Heron[94])	R_{Phi}	$\frac{ad-bc}{\sqrt{(a+b)(a+c)(b+d)(c+d)}}$
Maxwell and Pilliner [71]	R_{MP}	$\frac{2(ad-bc)}{(a+b)(b+d)+(a+c)(c+d)}$
Loevinger [62]	R_{Loe}	$\frac{ad-bc}{\min(a+b)(b+d),(a+c)(c+d)}$
Fleiss [31]	R_{Fleiss}	$\frac{(ad-bc)((a+b)(b+d)+(a+c)(c+d))}{2(a+b)(a+c)(b+d)(c+d)}$
Yule 1 [137]	R_{YuQ}	$\frac{ad-bc}{ad+bc}$
Yule 2 [138]	R_{YuY}	$\frac{\sqrt{ad}-\sqrt{bc}}{\sqrt{ad}+\sqrt{bc}}$
Hamann [39]	R_{Ham}	$\frac{a-b-c+d}{a+b+c+d}$
McConnaughey [72]	R_{McC}	$\frac{a^2-bc}{(a+b)(a+c)}$

We begin with a lemma that deals with two known PSD matrices for which the proof can be found in [8] for example.

Lemma 2.4.1. *Let $x_1, x_2, \ldots, x_m \in \mathbb{R}$. Then*

i) The matrix whose (i,j)-entry is given by $(x_i.x_j)_{1\leq i,j\leq m}$ is PSD.

ii) If in addition, $x_1, x_2, \ldots, x_m > 0$, then for all $\alpha \in \mathbb{R}^{+*}$, the matrix $\left(\frac{1}{(x_i+x_j)^\alpha}\right)_{1\leq i,j\leq m}$ is PSD. In particular, the Cauchy matrix $(\frac{1}{x_i+x_j})_{1\leq i,j\leq m}$ is PSD.

Proof.

i) Let P be an $m\times m$ matrix whose (i,j)-entry is given by $(x_i.x_j)_{1\leq i,j\leq m}$, and let $v=(v_1,v_2,...,v_m)^T \in \mathbb{R}^m$. Then, $v^TPv = \sum_{i=1}^m \sum_{j=1}^m v_j x_j x_i v_i = \sum_{i=1}^m \sum_{j=1}^m x_i v_i x_j v_j = (\sum_{i=1}^m x_i v_i)^2 \geq 0$.

ii) See [8, page 25].

□

As a conclusion, we prove the following lemma.

Lemma 2.4.2. *Let $x_1, x_2, \ldots, x_m > 0$, then the following holds.*

i) The matrix whose (i,j)-entry is given by $\left(\frac{1}{\max\{x_i,x_j\}}\right)_{1\leq i,j\leq m}$ is PSD.

ii) The matrix whose (i,j)-entry is $\left(\min\{x_i,x_j\}\right)_{1\leq i,j\leq m}$ is also PSD.

Proof.

i) For all i,j, we know that $\max\{x_i,x_j\} = \lim_{n\to\infty}(x_i^n + x_j^n)^{\frac{1}{n}}$. Hence we obtain $\frac{1}{\max\{x_i,x_j\}} = \lim_{n\to\infty}\frac{1}{(x_i^n+x_j^n)^{\frac{1}{n}}}$. In view of the preceding lemma, the matrices $\left(\frac{1}{(x_i^n+x_j^n)^{\frac{1}{n}}}\right)_{1\leq i,j\leq m}$ are PSD for all n. It follows that for any $v = (v_1, v_2, ..., v_m)^T \in \mathbb{R}^m$ and for all n,

$$v^T . \left(\frac{1}{(x_i^n + x_j^n)^{\frac{1}{n}}}\right)_{1\leq i,j\leq m} .v \geq 0.$$

This implies that

$$\lim_{n\to\infty} \quad v^T . \left(\frac{1}{(x_i^n + x_j^n)^{\frac{1}{n}}}\right)_{1\leq i,j\leq m} .v \geq 0,$$

that is,

$$v^T .(\max\{x_i,x_j\})_{1\leq i,j\leq m}.v \geq 0.$$

And the proof is completed.

ii) It is enough to notice that $\min\{x_i,x_j\} = \frac{1}{\max\{\frac{1}{x_i},\frac{1}{x_j}\}}$ and then the proof is complete by using the first part.

□

While the matrix $(\min\{x_i,x_j\})_{1\leq i,j\leq m}$ is PSD, the matrix $(\max\{x_i,x_j\})_{1\leq i,j\leq m}$ is not necessarily PSD. The next proposition deals with this issue.

Proposition 2.4.3. *Let $a_1,\ldots,a_m$ be m positive real numbers with $a = \max\{a_i\}$. Let A be the matrix defined by $A = (\max\{a_i,a_j\})_{1\leq i,j\leq m}$ and whose minimum and maximum eigenvalues are denoted by $\lambda_{\min}$ and $\lambda_{\max}$ respectively. Let I_m be the $m\times m$ identity matrix, and $\mathbb{J}_m$ to be the $m\times m$ matrix whose all entries are equal to 1. Then*

i) $0 < a \leq \lambda_{\max}$.

ii) If $a_i = a$ for all i, then $\lambda_{\min} = 0$ and $A = a\mathbb{J}_m \succeq 0$.

iii) Otherwise, $\lambda_{\min} < 0$ *and*

$$\lambda_{\min} I_m \preceq A \preceq a\mathbb{J}_m.$$

Moreover, A *and* 0 *are not comparable with respect to* $\preceq$.

Proof.

i) By appealing to the Courant-Fischer theorem (see [43, page 236]) which gives a variational formulation of the eigenvalues of a symmetric matrix, we know that

$$\lambda_{\min} \leq \frac{V^T AV}{V^T V} \leq \lambda_{\max} \quad \text{for all nonzero } V \in \mathbb{R}^n. \tag{2.1}$$

Suppose now that $a = a_i$ for some i, and in (2.1) if we let $V = (0,\dots,1,\dots 0)^T$ with 1 in i-th position and zero elsewhere, then we clearly obtain $V^T V = 1$ and $V^T AV = a_i = a$, so the proof of the first part is complete.

ii) The matrix $a\mathbb{J}_m$ possesses only two distinct eigenvalues $\lambda = 0$ and $\lambda = am$, which are nonnegative, hence $A = a\mathbb{J}_m \succeq 0$.

iii) First, $\lambda_{\min} < 0$, since otherwise $A \succeq 0$ and for every $i < j$, if we let

$$V = (0,\dots,0,1,0,\dots,0,-1,0,\dots,0)^T$$

with 1 in the i-th position, -1 in the j-th position and zero elsewhere, then we obtain $0 \leq V^T AV = a_i + a_j - 2\max\{a_i, a_j\}$. However, this would imply that $a_i = a_j$ for all i, j which is a contradiction. On the other hand, in view of (2.1), it is easy to see that $V^T AV \geq \lambda_{\min} V^T V$, hence $V^T(A - \lambda_{\min} I_m)V \geq 0$, $\forall V \in \mathbb{R}^n$, and the first inequality is proved. Next if we let $b = a + \frac{1}{n}$ where $n \geq 1$, then obviously $b > a_i$ for every i. Thus we get $b\mathbb{J} - A = (b - \max\{a_i, a_j\})_{i,j}$ = $(\min\{b - a_i, b - a_j\})_{i,j}$ which is PSD by the preceding lemma. By letting n tends to infinity, the proof is complete.

□

Finally, we give the following lemma that treats the convergence of sequences of PSD matrices $(V_n)_{n\geq 0}$, for which the proof can be found in [8].

Lemma 2.4.4. *Let* $V = (v_{ij})$ *be an* $m \times m$ *real matrix and for each natural number* k*, define the* k*-Hadamard product of* V *by* $V^{\circ k} = V \circ \dots \circ V = (v_{ij}^k)$*. Consider the sequence of matrices* $(V_n)_{n\geq 0}$ *which is defined by* $V_n = \sum_{k=0}^{n} V^{\circ k}$*. If* V *is PSD then the matrices* V_n *are PSD for all* n*, and* $\sum_{k=0}^{\infty} V_n = \lim_{n\to\infty} \sum_{k=0}^{n} V_n$ *is PSD.*

We now focus on constructing interesting classes of PSD matrices for which many similarity matrices will be deduced based on them.

Theorem 2.4.5. *Let $E = \{e_1, \ldots, e_n\}$ be a non-empty finite set, μ a non-trivial finite (positive) measure on E, and $(A_i)_{1\leq i\leq m}$ be m subsets of E. Then the following holds.*

1) The matrix $A = (\mu(A_i \cap A_j))_{1\leq i,j\leq m}$ is PSD. More generally, for any real numbers a and b the matrix

$$\left(a^2.\mu(A_i \cap A_j) + ab.\mu(A_i \triangle A_j) + b^2.\mu({A_i}^c \cap {A_j}^c)\right)_{1\leq i,j\leq m}$$

is PSD, where A_i^c is the complement of A_i, and $A_i \triangle A_j$ denotes the symmetric difference between the two sets.

2) If $\mu(A_i) \neq 0$ for all i and $a \geq b > 0$, then the following 4 matrices are PSD:

i) $B = \left(\frac{1}{a.\mu(A_i \triangle A_j)+b.\mu(A_i \cap A_j)}\right)_{1\leq i,j\leq m}$

ii) $C = \left(\frac{\mu(A_i \cap A_j)}{a.\mu(A_i \triangle A_j)+b.\mu(A_i \cap A_j)}\right)_{1\leq i,j\leq m}$

iii) $D = \left(\frac{1}{a.\mu(A_i \triangle A_j)+b.\mu((A_i \triangle A_j)^c)}\right)_{1\leq i,j\leq m}$

iv) $E = \left(\frac{\mu((A_i \triangle A_j)^c)}{a.\mu(A_i \triangle A_j)+b.\mu((A_i \triangle A_j)^c)}\right)_{1\leq i,j\leq m}$.

Proof.
1) Consider the Hilbert space $\mathcal{L}^2(E, \mu, \mathbb{R})$ with its inner product $<,>$ given by:

$$< f, g >= \int_E f.g \, d\mu, \tag{2.2}$$

for any $f, g \in \mathcal{L}^2(E, \mu, \mathbb{R})$. Note that for a any set $S \subseteq E$, $1_S(x)$ is the characteristic function of S returning 1 for $x \in S$. For any element x not contained in S, $1_S(x) = 0$. In addition, we have

$$\mu(S) = \int_E 1_S \, d\mu.$$

For any subsets A_i and A_j of E, the characteristic functions 1_{A_i} and $1_{Aj} \in \mathcal{L}^2(E, \mu, \mathbb{R})$. It follows that

$$< 1_{A_i}, 1_{A_j} >= \int_E 1_{A_i}.1_{A_j} \, d\mu = \int_E 1_{A_i \cap A_j} \, d\mu = \mu(A_i \cap A_j).$$

Therefore, the matrix A is a Gram-matrix, and thus it is PSD. Now in (2.2), if we let $f = a1_{A_i} + b1_{A_i^c}$ and $g = a1_{A_j} + b1_{A_j^c}$ then the proof can be achieved by noticing that

$$< a1_{A_i} + b1_{A_i^c}, a1_{A_j} + b1_{A_j^c} >= a^2.\mu(A_i \cap A_j) + ab.\mu(A_i \triangle A_j) + b^2.\mu({A_i}^c \cap {A_j}^c).$$

2) i) Let $b_{ij} = \dfrac{1}{a.\mu(A_i \triangle A_j) + b.\mu(A_i \cap A_j)}$. Then obviously we can write

$$\begin{aligned}
\frac{1}{b_{ij}} &= a.\mu\left((A_i \cup A_j) \setminus (A_i \cap A_j)\right) + b.\mu(A_i \cap A_j) \\
&= a.\mu(A_i \cup A_j) + (b-a)\mu(A_i \cap A_j) \\
&= a.\mu(E) - a.\mu(A_i^c \cap A_j^c) + (b-a)\mu(A_i \cap A_j) \\
&= a.\mu(E)(1 - v_{ij}),
\end{aligned}$$

where

$$v_{ij} = \frac{a.\mu(A_i^c \cap A_j^c) + (a-b)\mu(A_i \cap A_j)}{a.\mu(E)}. \tag{2.3}$$

As $a \geq b$ then $v_{ij} \geq 0$, and since $b > 0$ then, for each i, j, it holds that:

$$v_{ij} < \frac{a.\mu(A_i^c \cap A_j^c) + a.\mu(A_i \cap A_j)}{a.\mu(E)} = \frac{\mu((A_i \cup A_j)^c) + \mu(A_i \cap A_j)}{\mu(E)} \leq 1.$$

Now if we define the matrix $V = (v_{ij})$, then using (2.3), it is easy to see that the matrix V is PSD as it is the sum of two PSD matrices. Next, for each natural number n, define the matrix $V_n = \left(\sum_{k=0}^{n} v_{i,j}^k\right)_{1 \leq i,j \leq m}$, and consider the sequence $(V_n)_{n \geq 0}$. As V is PSD, then obviously V_n is PSD for each n and therefore in view of Lemma 2.4.4, $\lim\limits_{n \to \infty} V_n = \left(\frac{1}{1-v_{ij}}\right)_{1 \leq i,j \leq m}$ is PSD. To complete the proof, it suffices to notice that $b_{ij} = \frac{1}{a\mu(E)}\left(\frac{1}{1-v_{ij}}\right)$ which ensures that the matrix $B = (b_{ij})$ is PSD.

ii) It suffices to notice that the matrix $C = A \circ B$ and hence it is PSD since both A and B are PSD.

iii) Let $d_{ij} = \dfrac{1}{a\mu(A_i \triangle A_j) + b\mu((A_i \triangle A_j)^c)}$. Then

$$\begin{aligned}\frac{1}{d_{ij}} &= a.\mu((A_i \cup A_j) \setminus (A_i \cap A_j)) + b(\mu(A_i \cap A_j) + \mu(A_i^c \cap A_j^c)) \\ &= a.\mu(E) - a.\mu(A_i^c \cap A_j^c) - a.\mu(A_i \cap A_j) + b(\mu(A_i \cap A_j) + \mu(A_i^c \cap A_j^c)) \\ &= a.\mu(E) + (b-a)\mu(A_i^c \cap A_j^c) + (b-a)\mu(A_i \cap A_j) = a.\mu(E)(1-n_{ij})\end{aligned}$$

where $n_{ij} = \frac{(a-b)\mu(A_i^c \cap A_j^c) + (a-b)\mu(A_i \cap A_j)}{a.\mu(E)}$. Consequently we obtain

$$d_{ij} = \frac{1}{a.\mu(E)} \left(\frac{1}{1-n_{ij}} \right)$$

and thus the proof can be completed by employing a similar argument as the one used in Part i).

iv) Let M be the $m \times m$ matrix whose (i,j)-entry is given by $(A_i^c \cap A_j^c))_{1 \leq i,j \leq m}$. M is PSD by Part i) and thus the proof can be completed by noticing that $E = A \circ D + M \circ D$.

□

In particular, with the choice $\mu(\{e_i\}) = 1$, we obtain the following result.

Corollary 2.4.6. *Let $E = \{e_1, \ldots, e_n\}$ be a non-empty finite set, and $(A_i)_{1 \leq i \leq m}$ be m subsets of E. Then the following holds.*

1) *The matrix $(|A_i \cap A_j|)_{1 \leq i,j \leq m}$ is PSD, and for any real numbers a and b the matrix*

$$\left(a^2.|A_i \cap A_j| + ab.|A_i \triangle A_j| + b^2.|{A_i}^c \cap {A_j}^c|\right)_{1 \leq i \leq m}$$

is also PSD.

2) *If $\mu(A_i) \neq 0$ for all i and $a \geq b > 0$, then the following 4 matrices are PSD:*

   i) $\left(\frac{1}{a.|A_i \triangle A_j| + b.|A_i \cap A_j|}\right)_{1 \leq i,j \leq m}$

   ii) $\left(\frac{|A_i \cap A_j|}{a.|A_i \triangle A_j| + b.|A_i \cap A_j|}\right)_{1 \leq i,j \leq m}$

   iii) $\left(\frac{1}{a.|A_i \triangle A_j| + b.|(A_i \triangle A_j)^c|}\right)_{1 \leq i,j \leq m}$

   iv) $\left(\frac{|(A_i \triangle A_j)^c|}{a.|A_i \triangle A_j| + b.|(A_i \triangle A_j)^c|}\right)_{1 \leq i,j \leq m}$.

The next result shows a class of PSD matrices with negative values.

Theorem 2.4.7. *Let $E = \{e_1, \dots, e_n\}$ be a non-empty finite set, μ a non-trivial finite (positive) measure on E, and $(A_i)_{1\leq i\leq m}$ be m subsets of E. Then the matrix*

$$R = \left(\mu(A_i \cap A_j)\mu(A_i^c \cap A_j^c) - \mu(A_i \cap A_j^c)\mu(A_i^c \cap A_j)\right)_{1\leq i,j\leq m}$$

is PSD.

Proof. If we let $R = (r_{ij})$, then we can write:

$$\begin{aligned}
r_{ij} &= \mu(A_i \cap A_j)\mu(A_i^c \cap A_j^c) - \mu(A_i \cap A_j^c)\mu(A_i^c \cap A_j) \\
&= \mu(A_i \cap A_j)\mu(A_i^c \cap A_j^c) - (\mu(A_i) - \mu(A_i \cap A_j))(\mu(A_j) - \mu(A_i \cap A_j)) \\
&= \mu(A_i \cap A_j)\mu(A_i^c \cap A_j^c) - \mu(A_i)\mu(A_j) + \mu(A_i \cap A_j)(\mu(A_i) + \mu(A_j)) - (\mu(A_i \cap A_j))^2 \\
&= \mu(A_i \cap A_j)(\mu(A_i^c \cap A_j^c) + \mu(A_i) + \mu(A_j) - \mu(A_i \cap A_j)) - \mu(A_i)\mu(A_j) \\
&= \mu(A_i \cap A_j)(\mu(A_i^c \cap A_j^c) + \mu(A_i \cup A_j)) - \mu(A_i)\mu(A_j) \\
&= \mu(E)\mu(A_i \cap A_j) - \mu(A_i)\mu(A_j).
\end{aligned}$$

Noticing that

$$< 1_{A_i} - \frac{\mu(A_i)}{\mu(E)}1_E, 1_{A_j} - \frac{\mu(A_j)}{\mu(E)}1_E >= \mu(A_i \cap A_j) - \frac{\mu(A_i)\mu(A_j)}{\mu(E)},$$

so that $\frac{1}{\mu(E)}R =< 1_{A_i} - \frac{\mu(A_i)}{\mu(E)}1_E, 1_{A_j} - \frac{\mu(A_j)}{\mu(E)}1_E >$ is a Gram matrix and hence it is PSD. Therefore, R is a PSD matrix. □

Similarly, with the choice $\mu(\{e_i\}) = 1$, we obtain the following.

Corollary 2.4.8. *Let $E = \{e_1, \dots, e_n\}$ be a non-empty finite set, and $(A_i)_{1\leq i\leq m}$ be m subsets of E. Then the matrix*

$$\left(|A_i \cap A_j||A_i^c \cap A_j^c| - |A_i \cap A_j^c||A_i^c \cap A_j|\right)_{1\leq i,j\leq m}$$

is PSD.

2.5 A list of PSD similarity matrices

Although numerous binary similarity measures have been proposed in different fields [131], few are proved to be positive semi-definite [35, 36]. Moreover, similarity measures can also be defined based on sets of arbitrary elements. Our approach in this section is to consider many similarity matrices as cardinals of the intersection, union, symmetric difference and other set theoretical operations between subsets of a finite set, to show that our previous results concerning certain positive semi-definite matrices will pave the way to prove that these similarity matrices are positive semi-definite.

2.5.1 Notations

Suppose that $x = (x_1, x_2, ..., x_n)$ and $y = (y_1, y_2, ..., y_n)$ are two binary n sequences, and let $X = \{1 \leq i \leq n : x_i = 1\}$ and $Y = \{1 \leq i \leq n : y_i = 1\}$ be the set of attributes present in objects x and y respectively. Then, the binary similarity measures can be expressed using the cardinals of the following four quantities

- $a = |X \cap Y|$
- $b = |X \cap \bar{Y}|$
- $c = |\bar{X} \cap Y|$
- $d = |\bar{X} \cap \bar{Y}|$.

It follows that all similarity measures are now defined based on the cardinality of theoretical operations between subsets of a finite set. Hence, without loss of a generality, we can suppose that $\mathcal{X} = P(E)$, the power set of a finite set E of cardinal n and a similarity measure S is defined from $P(E) \times P(E)$ into $\mathbb{R}$. Given m non empty subsets $(A_i)_{1 \leq i \leq m}$ of E, the similarity measure S gives rise to a similarity matrix $M_S = \big(S(A_i, A_j)\big)_{1 \leq i,j \leq m}$.

2.5.2 Family of PSD similarity matrices

It is worth mentioning here that similarity measures may be members of some sort of a parameter family. For instance, Tversky [129] proposed a model of similarity measures where he rejected the symmetry constraint. He introduced the following contrast model

$$S_{Tve}(x, y) = \frac{a}{a + \alpha.b + \beta.c},$$

i.e,

$$S_{Tve}(x,y) = \frac{|X \cap Y|}{|X \cap Y| + \alpha|X - Y| + \beta|X - Y|},$$

where α and β are any two real numbers. Contrary to others, Tversky argues that similarity should not be treated as a symmetric relation. He explains about the directional nature of the similarity relation of the form "x is like y". Such statement has a subject x and a referent y. Hence the two objects have different roles and non-symmetric measure are necessary ([129]). With non-symmetric measures, we obtain asymmetric similarity matrices that we will not consider in this thesis. Symmetry is obtained when $\alpha = \beta$.

Moreover, in [36], Gower and Legendre define for two binary variables x and y, two parameter families of similarity coefficients S_θ and T_θ as follows:

$$T_\theta(x,y) = \frac{|X \cap Y|}{|X \cap Y| + \theta|X \triangle Y|},$$

and,

$$S_\theta(x,y) = \frac{|(X \triangle Y)^c|}{|(X \triangle Y)^c| + \theta|X \triangle Y|},$$

where θ is a real number . We note that T_θ is the special case of Tversky index, when $\alpha = \beta$.

By taking $\theta \geq 0$ to avoid the possibility of negative similarity coefficients, and making use of parts ii) and iv) of Corollary 2.4.6 for $a = \theta$ and $b = 1$, we have the following conclusion.

Corollary 2.5.1. *Let $E = \{e_1, \ldots, e_n\}$ be a non-empty finite set, μ a non-trivial finite (positive) measure on E, and $(A_i)_{1 \leq i \leq m}$ be m subsets of E such that $\forall i\, A_i \neq \emptyset$. Assume that $\theta \geq 1$, then the following similarity matrices:*

i) $\big(T_\theta(A_i,A_j)\big)_{1 \leq i,j \leq m} = \left(\frac{|A_i \cap A_j|}{|A_i \cap A_j| + \theta|A_i \triangle A_j|}\right)_{1 \leq i,j \leq m}$ *and*

ii) $\big(S_\theta(A_i,A_j)\big)_{1 \leq i,j \leq m} = \left(\frac{|(A_i \triangle A_j)^c|}{|(A_i \triangle A_j)^c| + \theta|A_i \triangle A_j|}\right)_{1 \leq i,j \leq m}$

are PSD.

Next, we give examples of known similarity matrices that belongs to the previous families.

Example 1.

1. $M_{S_{Jac}} = \left(\frac{|A_i \cap A_j|}{|A_i \cup A_j|}\right)_{1\leq i,j\leq m} = \left(\frac{|A_i \cap A_j|}{|A_i \triangle A_j| + |A_i \cap A_j|}\right)_{1\leq i,j\leq m} = (T_1(A_i,A_j))_{1\leq i,j\leq m}.$
2. $M_{S_{SS2}} = \left(\frac{|A_i \cap A_j|}{2|A_i \triangle A_j| + |A_i \cap A_j|}\right)_{1\leq i,j\leq m} = (T_2(A_i,A_j))_{1\leq i,j\leq m}.$
3. $M_{S_{SM}} = \left(\frac{|A_i \cap A_j| + |A_i^c \cap A_j^c|}{n}\right)_{1\leq i,j\leq m} = (S_1(A_i,A_j))_{1\leq i,j\leq m}.$
4. $M_{S_{RT}} = \left(\frac{|(A_i \triangle A_j)^c|}{|(A_i \triangle A_j)^c| + n}\right)_{1\leq i,j\leq m} = (S_2(A_i,A_j))_{1\leq i,j\leq m}.$

The following remarks exhibits the case of θ less than 1.

Remark 2. It is worthy to mention here that for $\theta < 1$, our knowledge concerning the positive semi-definiteness of the elements of T_θ and S_θ is less. For example, by taking T_θ with $\theta = \frac{1}{2}$, we obtain

$$T_{\frac{1}{2}} = \left(\frac{2|A_i \cap A_j|}{2|A_i \cap A_j| + |(A_i \triangle A_j)|}\right)_{1\leq i,j\leq m} = \left(\frac{2|A_i \cap A_j|}{|A_i| + |A_j|}\right)_{1\leq i,j\leq m},$$

which we will prove to be positive semi-definite in the following section. However, for $\theta < \sqrt{2}-1$, we can show by a counterexample that T_θ is not PSD. Indeed, consider the three binary variables x, y and $z \in \{0,1\}^4$, defined by $x = (1,0,0,0), y = (0,0,0,1)$ and $z = (1,0,0,1)$. Then by a simple computation the similarity matrix T_θ between x, y and z is given by

$$T_\theta(x,y,z) = \begin{pmatrix} 1 & 0 & \frac{1}{\theta+1} \\ 0 & 1 & \frac{1}{\theta+1} \\ \frac{1}{\theta+1} & \frac{1}{\theta+1} & 1 \end{pmatrix}.$$

By a simple check, $T_\theta(x,y,z)$ has eigenvalues $1, \frac{\theta+\sqrt{2}+1}{1+\theta}, \frac{\theta-\sqrt{2}+1}{1+\theta}$. The proof can be achieved by noticing that for $\theta < \sqrt{2}-1$, $T_\theta(x,y,z)$ has a negative eigenvalue.

2.5.3 Other similarity matrices

Next, our main objective is to show that many binary similarity matrices introduced in Section 2.3, can be rewritten as cardinals of theoretical operations between subsets of a finite set and consequently they are PSD using the results of the previous section.

1. $\mathbf{M}_{S_{RR}} = \left(\frac{|A_i \cap A_j|}{n}\right)_{1\leq i,j\leq m}$ is PSD by Corollary 2.4.6 part 1).

2. $\mathbf{M}_{S_{Fai}} = \left(\frac{|A_i \cap A_j| + |(A_i \triangle A_j)^c|}{2n}\right)_{1 \leq i,j \leq m}$ is PSD since

$$|(A_i \triangle A_j)^c| = |E| - (|A_i \cup A_j| - |A_i \cap A_j|) = |E \setminus A_i \cap E \setminus A_j| + |A_i \cap A_j|$$

and then making use of Corollary 2.4.6 part 1).

3. $\mathbf{M}_{S_{For}} = \left(\frac{n.|A_i \cap A_j|}{|A_i|.|A_j|}\right)_{1 \leq i,j \leq m}$ is PSD by Corollary 2.4.6 part 1), Lemma 2.4.1 part i) with $x_i = \frac{1}{|A_i|}$ and $x_j = \frac{1}{|A_j|}$, and Schur's theorem.

4. $\mathbf{M}_{S_{Sorg}} = \left(\frac{|A_i \cap A_j|^2}{|A_i|.|A_j|}\right)_{1 \leq i,j \leq m}$ is PSD for the same reason stated in the preceding case and then by applying Schur's theorem twice.

5. $\mathbf{M}_{S_{Och}} = \left(\frac{|A_i \cap A_j|}{\sqrt{|A_i|.|A_j|}}\right)_{1 \leq i,j \leq m}$ is PSD by Corollary 2.4.6 part 1), Lemma 2.4.1 part i) with $x_i = \frac{1}{\sqrt{|A_i|}}$ and $x_j = \frac{1}{\sqrt{|A_j|}}$, and Schur's theorem.

6. $\mathbf{M}_{S_{SS3}} = \left(\frac{|A_i \cap A_j|.|(A_i \cup A_j)^c|}{\sqrt{|A_i|.|A_i^c|.|A_j|.|A_j^c|}}\right)_{1 \leq i,j \leq m}$ is PSD by Corollary 2.4.6 part 1), Lemma 2.4.1 part i) with $x_i = \frac{1}{\sqrt{|A_i||A_i^c|}}$, and Schur's theorem applied twice.

7. $\mathbf{M}_{S_{Glea}} = \left(\frac{2|A_i \cap A_j|}{2|A_i \cap A_j| + |A_i \cap A_j^c| + |A_i^c \cap A_j|}\right)_{1 \leq i,j \leq m} = \left(\frac{2|A_i \cap A_j|}{|A_i| + |A_j|}\right)_{1 \leq i,j \leq m}$ is PSD by Corollary 2.4.6 part 1), Lemma 2.4.1 part ii) with $x_i = |A_i|$, and Schur's theorem.

8. $\mathbf{M}_{S_{RG}} = \left(\frac{|A_i \cap A_j|}{|A_i| + |A_j|} + \frac{|A_i^c \cap A_j^c|}{|A_i^c| + |A_j^c|}\right)_{1 \leq i,j \leq m}$ is PSD by the same reasonning as in the preceding case.

9. $\mathbf{M}_{S_{BB}} = \left(\frac{(|A_i \cap A_j|}{\max(|A_i|, |A_j|)}\right)_{1 \leq i,j \leq m}$ is PSD by Corollary 2.4.6 part *i*), Lemma 2.4.2 and Schur's theorem.

10. $\mathbf{M}_{S_{HD}} = \frac{1}{2}\left(\frac{|A_i \cap A_j|}{|A_i \triangle A_j| + |A_i \cap A_j|} + \frac{|A_i^c \cap A_j^c|}{|A_i \triangle A_j| + |A_i^c \cap A_j^c|}\right)_{1 \leq i,j \leq m}$ is PSD as $\left(\frac{|A_i^c \cap A_j^c|}{|A_i \triangle A_j| + |A_i^c \cap A_j^c|}\right)_{1 \leq i,j \leq m}$ is PSD and then employing a similar proof to that used in proving Theorem 2.4.5 part 2*ii*).

11. $\mathbf{M}_{R_{Disp}} = \left(\frac{|A_i \cap A_j|.|A_i^c \cap A_j^c| - |A_i^c \cap A_j|.|A_i \cap A_j^c|}{n^2}\right)_{1 \leq i,j \leq m}$ is PSD by Corollary 2.4.8.

12. $\mathbf{M}_{R_{Phi}} = \left(\frac{|A_i \cap A_j|.|A_i^c \cap A_j^c| - |A_i^c \cap A_j|.|A_i \cap A_j^c|}{\sqrt{|A_i|.|A_i^c|.|A_j|.|A_j^c|}}\right)_{1 \leq i,j \leq m}$ is PSD by Corollary 2.4.8 and Lemma 2.4.1 with $x_i = \frac{1}{\sqrt{|A_i||A_i^c|}}$, and Schur's theorem.

13. $\mathbf{M}_{S_{Doo}} = \left(\frac{(|A_i \cap A_j|.|A_i^c \cap A_j^c| - |A_i^c \cap A_j|.|A_i \cap A_j^c|)^2}{|A_i|.|A_i^c|.|A_j|.|A_j^c|} \right)_{1 \leq i,j \leq m}$ is PSD since it is the Hadamard product of $\mathbf{M}_{R_{Phi}}$ by itself.

14. $\mathbf{M}_{S_{Pearson}} = \left(\frac{n(|A_i \cap A_j|.|A_i^c \cap A_j^c| - |A_i^c \cap A_j|.|A_i \cap A_j^c|)^2}{|A_i|.|A_i^c|.|A_j|.|A_j^c|} \right)_{1 \leq i,j \leq m}$ is PSD since it is the Hadamard product of $\mathbf{M}_{S_{Doo}}$ and the matrix whose all entries are equal to n.

15. $\mathbf{M}_{R_{MP}} = \left(\frac{2(|A_i \cap A_j|.|A_i^c \cap A_j^c| - |A_i^c \cap A_j|.|A_i \cap A_j^c|)}{|A_i|.|A_i^c| + |A_j|.|A_j^c|} \right)_{1 \leq i,j \leq m}$ is PSD by Corollary 2.4.8, Lemma 2.4.1 part ii) with $x_i = |A_i||A_i^c|$, and Schur's theorem.

16. $\mathbf{M}_{R_{Ham}} = \left(2\frac{|A_i \cap A_j| + |A_i^c \cap A_j^c|}{n} - 1 \right)_{1 \leq i,j \leq m} = \left(\frac{|A_i \cap A_j| + |A_i^c \cap A_j^c| - |A_i \triangle A_j|}{n} \right)_{1 \leq i,j \leq m}$, is PSD by Corollary 2.4.6 part 1) with $a = 1$ and $b = -1$.

2.6 Positive definiteness of similarity matrices

In this section, we will extend the previous results to a stronger property than the positive semi-definite property. In particular, we will focus on the definite property, which insures that a null distance between two objects guarantees they are strictly identical, or formally, $d(x_1, x_2) = 0 \Longleftrightarrow x_1 = x_2$. Although, the reflexivity property is always satisfied (see Section 1.5), the separability is rather often violated by distance measures. This is what distinguishes metric distances from pseudo-metric distances. The positive definiteness of a similarity matrix used as a weighting matrix in the generalised Euclidean distance will guaranty then that the associated Euclidean distance is a full metric. Among all similarity matrices proved to be positive semi-definite, the complete $(2^n - 1) \times (2^n - 1)$ Jaccard index of all pairs of subsets of a finite set E of size n (excluding the empty set), is by far the only similarity matrix proved to be positive definite for any positive integer n [11].

First, consider a non-empty finite set E of size n. Let $P(E)$ be the power set of E, and $P(E)' = P(E) \setminus \emptyset$. Since every principal submatrix of a positive definite matrix is positive definite, the Jaccard similarity formed by m arbitrary non-empty subsets of E is also positive definite. However, the positive definiteness of other similarity matrices depends on the subsets under study, as shows the following example.

Example 2. Let $E = \{e_1, e_2, e_3, e_4\}$ be a finite set of size 4.

- Consider the Ochiai-Otsuka similarity matrix O_1 formed by the 3 subsets $A_1 = \{e_1\}$, $A_2 = \{e_2\}$ and $A_3 = \{e_1, e_2\}$. Then with respect to the following order: A_1, A_2, A_3, we

get

$$O_1 = \begin{pmatrix} 1 & 0 & 1/\sqrt{2} \\ 0 & 1 & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} & 1 \end{pmatrix},$$

and O_1 is positive semi-definite and not positive definite since it has 0 as one of its eigenvalues.

- However, considering the Ochiai Otsuka similarity matrix O_2 formed by the 3 subsets $A_1 = \{e_1\}, A_2 = \{e_2, e_3\}$ and $A_3 = \{e_2, e_4\}$. Then with respect to the order A_1, A_2, A_3, we obtain

$$O_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1/2 \\ 0 & 1/2 & 1 \end{pmatrix},$$

and O_2 is positive definite since all its eigenvalues are strictly positive.

2.6.1 Sufficient condition for the positive definiteness of similarity matrices

As mentioned before, the positive definiteness of PSD similarity matrices depends on the data set under study. In the next theorem, we will give a sufficient condition on the subsets of E under study for which all similarity matrices given in Section 2.5.3, are positive definite. For this purpose, we state the following lemma for which the proof can be found in [43, Theorem 7.5.3].

Lemma 2.6.1. *Let $A, B \in M_n$ be positive semi-definite. If A is positive definite and every entry on the main diagonal of B is positive, then $A \circ B$ is positive definite.*

Theorem 2.6.2. *Consider the matrix $M = \left(|A_i \cap A_j|\right)_{1 \leq i \leq m}$ formed by m subsets of E such that for all $i \in \{1 \leq i \leq m\}$, there exists $x_i \in E$ with $x_i \in A_i \setminus \cup_{j \neq i} A_j$ Then M is positive definite.*

Proof.

By part i) of Theorem 2.4.5, $M = \left(< 1_{A_i}, 1_{A_j} >\right)_{1 \leq i \leq m}$. Now let $\alpha_1, \alpha_2, ..., \alpha_m \in \mathbb{R}$ such that

$$\alpha_1 1_{A_1} + \alpha_2 1_{A_2} + ... + \alpha_m 1_{A_m} = 0.$$

Hence for all $x \in E$, $(\alpha_1 1_{A_1} + \alpha_2 1_{A_2} + ... + \alpha_m 1_{A_m})(x) = 0$. In particular, for $x = x_i$ we have $\alpha_1 1_{A_1}(x_i) + \alpha_2 1_{A_2}(x_i) + ... + \alpha_m 1_{A_m}(x_i) = 0$. Thus $\alpha_i = 0$, since $1_{A_j}(x_i) = 0$ for all $j \neq i$, and so the vectors $1_{A_1}, 1_{A_2}, ..., 1_{A_m}$ are linearly independent. Therefore M is a Gram matrix of m

linearly independent vectors in the inner product space $\mathcal{L}^2(E,\mu,\mathbb{R})$, and consequently M is positive definite. □

The previous condition is also a sufficient condition for any similarity matrix considered in Section 2.5.3, to be strictly positive definite, as the next corollary shows.

Corollary 2.6.3. *Let E be a non-empty finite set of size n. Let S be any similarity matrix from Section 4.1, formed by m arbitrary subsets $(A_i)_i$ of E and such that for all $i \in \{1,2,...,m\}$, there exists $x_i \in E$ with $x_i \in A_i \setminus \cup_{j\neq i} A_j$. Then S is a positive definite.*

Proof. It suffices to see that every similarity matrix of Section 2.5.3 is the Hadamard product of $M = \left(|A_i \cap A_j|\right)_{1\leq i\leq m}$ and a positive semi-definite matrix with positive main diagonal entries. By Lemma 2.6.1, the proof is complete. □

The next simple example is given to illustrate the previous corollary.

Example 3. Consider the following four binary 9-sequences $x = (1,1,0,0,1,1,0,1,0)$, $y = (1,0,1,1,0,1,0,1,0)$, $z = (1,0,0,1,1,0,1,0,0)$ and $t = (1,0,0,1,1,1,0,1,1)$. Using the same notation as in Section 2.5, we have $X = \{1,2,5,6,8\}$, $Y = \{1,3,4,6,8\}$, $Z = \{1,4,5,7\}$, and $T = \{1,4,5,6,8,9\}$ which are four subsets that belong to the set $E = \{1,2,3,4,5,6,7,8,9\}$. It is easy to see that X, Y, Z and T satisfy the assumption of the previous corollary. This implies that the Ochiai-Otsuka similarity matrix O_3 formed by the 4 subsets X, Y, Z and T given by

$$O_3 = \begin{pmatrix} 1 & \frac{3}{5} & \frac{1}{\sqrt{5}} & \frac{4}{\sqrt{30}} \\ \frac{3}{5} & 1 & \frac{1}{\sqrt{5}} & \frac{4}{\sqrt{30}} \\ \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 1 & \frac{3}{2\sqrt{6}} \\ \frac{4}{\sqrt{30}} & \frac{4}{\sqrt{30}} & \frac{3}{2\sqrt{6}} & 1 \end{pmatrix},$$

is positive definite.

The condition of the previous corollary is achieved when considering a sample of objects such that each object has a proper characteristic (or attribute) not found in the others.

2.6.2 The complete Tversky similarity matrix is positive definite

For any $a \geq b > 0$, consider the Tversky similarity matrix of m arbitrary nonempty subsets $(A_i)_{1\leq i\leq m}$ of E given by

$$\left(\frac{|A_i \cap A_j|}{a|A_i \triangle A_j| + b|A_i \cap A_j|}\right)_{1\leq i,j\leq m}.$$

This subsection explores the positive definitness of the Tversky family of matrices given above. It is worthy to note here that the only similarity matrix proved by far to be strictly positive definite, is the Jaccard similarity matrix which is a special case of Tversky matrices ($a = b = 1$). First, let us introduce the following notation.

Notation. For a matrix M whose rows and columns are indexed by elements of $P(E)'$, the scalar $M(A,B)$ refers to the entry of M whose row is associated with the subset A and whose column is associated with B.

In [11], Bouchard et al. proved the following theorem.

Theorem 2.6.4. *The matrix **Jac** whose elements are the Jaccard indexes*

$$Jac(A,B) = \frac{|A \cap B|}{|A \cup B|},$$

of all pairs of subsets (excluding the empty set) A and B of a reference frame of size n is positive definite, for any integer $n > 1$.

For this purpose, **Jac** can be written as $\mathbf{Jac} = \lim_{r\to\infty} Jac_r$ where $Jac_r = \sum_{k=0}^{r} J_k$ and $J_k = \frac{Int}{n} \circ \frac{W_k}{n^k}$, with Int and W_k are matrices whose rows and columns are indexed by the elements of $P(E)'$ and such that $Int(A,B) = |A \cap B|$ and $W_k(A,B) = (n - |A \cup B|)^k$. In [11], the authors proved that for any nonzero vector $x \in \mathbb{R}^{2^n-1}$, we have either $x^T J_0 x > 0$ or $x^T J_{n-1} x > 0$. Our main objective here is to extend this result to the $(2^n-1) \times (2^n-1)$ Tversky similarity matrix **T** which is formed from all pairs of subsets of E excluding the empty set.

In view of Corollary 2.4.6 part 2) *ii*), the Tversky similarity matrix formed by m arbitrary nonempty subsets $(A_i)_{1\leq i\leq m}$ of E is PSD. Hence, the complete $(2^n-1) \times (2^n-1)$ Tversky similarity matrix **T** is PSD. In fact, it is (strictly) PD as the next theorem shows.

Theorem 2.6.5. *Let E be a non-empty finite set of size n. Let* **T** *be the complete $(2^n-1) \times (2^n-1)$ Tversky similarity matrix whose elements are the Tversky's coefficients for any pair of subsets A and B of E (excluding the empty set) defined by*

$$T(A,B) = \frac{|A \cap B|}{a|A \triangle B| + b|A \cap B|}.$$

Then **T** *is positive definite for any positive integer $n > 1$, and for all $a \geq b > 0$.*

Proof. Let $a \geq b > 0$, and consider A and B to be any two different elements in $P(E)'$. In addition, let L be the matrix whose rows and columns are indexed by $P(E)'$ and such

that $L(A,B) = \frac{1}{a|A\triangle B|+b|A\cap B|}$. By Theorem 2.4.5, L is positive semi-definite and $L(A,B) = \frac{1}{an}\left(\frac{1}{1-V(A,B)}\right)$ where V is the matrix indexed by $P(E)'$, and with $V(A,B) = \frac{a|A^c\cap B^c|+(a-b)|A\cap B|}{an}$. Moreover, since $\frac{1}{1-V(A,B)} = \lim_{r\to\infty} V_r(A,B)$ where $V_r = \sum_{k=0}^{r} V^{\circ k}$, then $L = \frac{1}{an}.\lim_{r\to\infty} V_r$, and hence

$$\mathbf{T} = Int \circ L = \frac{1}{an}\lim_{r\to\infty} Int \circ V_r.$$

Now clearly for any nonzero vector $x \in \mathbb{R}^{2^n-1}$, we have

$$x^T\mathbf{T}x = \frac{1}{an}\lim_{r\to\infty} x^T(Int \circ V_r)x.$$

Since $V_r = \sum_{k=0}^{r} V^{\circ k}$ is a sum of PSD matrices, then every term is greater than or equal to 0. Thus, it suffices to consider only two terms and consequently we can write

$$x^T\mathbf{T}x \geq \frac{1}{an}\left(x^T(Int \circ V^0)x + x^T(Int \circ V^{\circ(n-1)})x\right).$$

It is easy to see that $Int \circ V^0 = Int = nJ_0$. Next, if we let X and Y be two matrices such that for each of them, the rows and columns are indexed by $P(E)'$ and with $X(A,B) = a|A^c \cap B^c|$ and $Y(A,B) = (a-b)|A\cap B|$, then we obtain

$$V^{\circ(n-1)} = \frac{1}{(an)^{n-1}}(X+Y)^{\circ(n-1)}.$$

On the other hand, since $(X+Y)^{\circ(n-1)} = \sum_{k=0}^{n-1}\binom{n-1}{k} X^{\circ(n-1-k)} \circ Y^{\circ k}$ $\left(\text{where } \binom{n-1}{k} := \frac{(n-1)!}{(n-1-k)!k!}\right)$, then clearly

$$(X+Y)^{\circ(n-1)} - X^{\circ(n-1)} = \sum_{k=1}^{n-1}\binom{n-1}{k} X^{\circ(n-1-k)} \circ Y^{\circ k}.$$

Each term of the right-hand side of this last equality is positive semi-definite since it is a Hadamard product of positive semi-definite matrices. Obviously, we conclude that the left hand side is also positive semi-definite. Therefore, for any nonzero vector $x \in \mathbb{R}^{2^n-1}$, it holds that

$$x^T((X+Y)^{\circ(n-1)} - X^{\circ(n-1)})x \geq 0,$$

$$x^T((X+Y)^{\circ(n-1)})x \geq x^T X^{\circ(n-1)}x,$$

$$x^T V^{\circ(n-1)}x \geq \frac{1}{(an)^{n-1}}x^T X^{\circ(n-1)}x.$$

But this implies that

$$\begin{aligned} x^T(Int \circ V^{\circ(n-1)})x &\geq \frac{1}{(an)^{n-1}} x^T(Int \circ X^{\circ(n-1)})x \\ &= \frac{1}{n^{n-1}} x^T(Int \circ W_{n-1})x \\ &= n.x^T J_{n-1} x. \end{aligned}$$

Finally, we obtain

$$\begin{aligned} x^T \mathbf{T} x &\geq \frac{1}{an} x^T(nJ_0)x + \frac{1}{an} x^T(nJ_{n-1})x \\ &= \frac{1}{a}(x^T J_0 x + x^T J_{n-1} x) > 0. \end{aligned}$$

□

Next, we give an example of a complete Tversky similarity matrix where $|E| = 3$.

Example 4. Consider Sokal and Sneath similarity measure defined below by

$$S_{SS2} = \left(\frac{|A \cap B|}{2|A \triangle B| + |A \cap B|} \right),$$

for any pair of subsets A and B, of the finite set $E = \{e_1, e_2, e_3\}$. It is easy to see that S_{SS2} is a special case of Tversky matrix where $a = 2$ and $b = 1$. The complete Sokal and Sneath similarity matrix where the rows are indexed by the subsets of E according to the following order: $\{e_1\}$, $\{e_2\}$, $\{e_3\}$, $\{e_1, e_2\}$, $\{e_1, e_3\}$, $\{e_2, e_3\}$ and E, is given by

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{5} \\ 0 & 1 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{5} \\ 0 & 0 & 1 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{3} & 0 & 1 & \frac{1}{5} & \frac{1}{5} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{5} & 1 & \frac{1}{5} & \frac{1}{2} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{5} & \frac{1}{5} & 1 & \frac{1}{2} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}.$$

By Theorem 2.6.3, $\mathbf{T}$ is positive definite. Therefore, any principal submatrix of $\mathbf{T}$ is positive definite. This implies that for any combination of p binary 3-sequences among the 7 binary possible non zero 3-sequences, the Sokal and Sneath similarity matrix between them is positive definite.

In the previous theorem, we provided a formal proof of the positive definiteness of the Tversky similarity matrix. This property means that the Tversky distance defined by $d_T^2(m_1,m_2) = \sqrt{(m_1-m_2)^T T (m_1-m_2)}$, between belief functions m_1 and m_2 is a full distance for any size of the frame of discernment E. This result justify the use of the Tversky distance as a suitable distance between belief functions. We will discuss more about this in Section 2.9.

2.7 Three-positive semi-definite similarity matrices

Not all similarity matrices are PSD as we will see in Section 2.8. So in general the question of characterising which similarity matrices are PSD, is not a straightforward matter and in certain cases, it may be hard to prove that particular similarity matrices are PSD. So one can check if such matrices satisfy a weaker property such as the three-positive semi-definite property. In this section, we present a sufficient condition to a similarity matrix to be three-positive semi-definite and we investigate briefly the advantages of such property.

For this purpose, we recall the following definition from [18].

Definition 2.7.1. Given a set X, a real-valued function $s(x,y)$ on the cartesian product $X \times X$ is a similarity metric if, for any $x,y,x \in X$, it satisfies the following conditions:

1. $s(x,y) = s(y,x)$
2. $s(x,x) \geq 0$
3. $s(x,x) \geq s(x,y)$
4. $s(x,y) = s(x,x) = s(y,y)$ if and only if $x = y$.
5. $s(x,y) + s(y,z) \leq s(x,z) + s(y,y)$.

Next, we impose on Definition 2.2.1 of a similarity measure, the condition (5) that, in our terminology will become

$$S(X,Y) + S(Y,Z) \leq S(X,Z) + S(Y,Y), \tag{2.4}$$

for all $X,Y,Z \in P(E)$. Inequality (2.4) states that the similarity between X and Z through Y is no greater than the direct similarity between X and Z plus the self similarity of Y (which is somewhat equivalent to the triangle inequality in distance metrics). As mentioned earlier, it maybe very difficult to prove that such similarity matrix is PSD with this extra property. For this reason, we need the following definition.

Definition 2.7.2. Let $S : P(E) \times P(E) \to \mathbb{R}^+$(or $\mathbb{R}$) be a similarity. Then S is said to be three-positive semi-definite if for all $X,Y,Z \in P(E)$, the matrix

$$\begin{pmatrix} S(X,X) & S(X,Y) & S(X,Z) \\ S(X,Y) & S(Y,Y) & S(Y,Z) \\ S(X,Z) & S(Y,Z) & S(Z,Z) \end{pmatrix},$$

is a positive semi-definite matrix.

As a conclusion, we have the following lemma.

Lemma 2.7.3. *Let $S : P(E) \times P(E) \to \mathbb{R}^+$ ($\mathbb{R}$) be a similarity. Then S is three-positive semi-definite if and only if for all $X,Y,Z \in P(E)$,*

$$\begin{vmatrix} S(X,X) & S(X,Y) & S(X,Z) \\ S(X,Y) & S(Y,Y) & S(Y,Z) \\ S(X,Z) & S(Y,Z) & S(Z,Z) \end{vmatrix} \geq 0.$$

Proof. It is easy to see that the principal minors of size 1 and 2 are nonnegative from the definition of a similarity measure. □

We are ready now to identify a class of similarity matrices that are three-positive semi-definite in the following theorem.

Theorem 2.7.4. *Let $S : P(E) \times P(E) \to \mathbb{R}^+$ be a similarity coefficient such that $S(X,X) = k$, $k > 0$, for all $X \in P(E)$ and $S(X,Y) + S(Y,Z) \leq S(X,Z) + k$, for all $X,Y,Z \in P(E)$. Then S is three-positive semi-definite.*

Proof. By the preceding lemma, it suffices to prove that

$$k^3 + 2S(X,Y)S(Y,Z)S(X,Z) \geq k(S^2(X,Y) + S^2(X,Z) + S^2(Y,Z)).$$

For simplicity, we denote $a = S(X,Y)$, $b = S(Y,Z)$ and $c = S(X,Z)$. Then from Inequality (2.4), we know that the following 3 inequalities hold:

- $a + b \leq c + k$
- $a + c \leq b + k$
- $b + c \leq a + k$.

Now if we let $x = a+k-b-c$, $y = b+k-a-c$, and $z = c+k-a-b$, then obviously x, y and z are nonnegative with $x+y+z = 3k-a-b-c \leq 3k$, and

$$a = \frac{2k-y-z}{2}, \quad b = \frac{2k-x-z}{2}, \quad c = \frac{2k-x-y}{2}.$$

In order to prove that $k^3 + 2abc \geq k(a^2+b^2+c^2)$, we need to prove

$$\frac{k}{4}((2k-x-y)^2+(2k-x-z)^2+(2k-y-z)^2) \leq k^3 + \frac{1}{4}(2k-x-y)(2k-x-z)(2k-y-z)$$

which is equivalent to prove that

$$\begin{aligned} &k(4k^2+2x^2-8kx+2xy+4k^2+2y^2-8ky+2yz+4k^2+2z^2-8kz+2xz) \\ &\leq 12k^3+(2kx^2-8k^2x+6kxy)+(2ky^2-8k^2y+6kyz)+(2kz^2-8k^2z+6kxz) \\ &\quad -(x+y)(y+z)(x+z). \end{aligned}$$

That in turn is equivalent to show that

$$4k(xy+xz+yz) \geq (x+y)(x+z)(y+z).$$

Since $1 \geq \frac{x+y+z}{3k}$, then it is enough to prove that

$$\frac{4k}{3k}(x+y+z)(xy+xz+yz) \geq (x+y)(x+z)(y+z)$$

which is equivalent to show that

$$x^2y+x^2z+y^2x+y^2z+z^2x+z^2y+6xyz \geq 0.$$

However, this last inequality is true since x, y and z are nonnegative. □

As mentioned in the introduction, the use of PSD similarity matrices is important to derive distances between data set. The question that arise here is if a three-positive semi-definite matrix can be of a similar importance. That we will discuss briefly in the following two remarks.

Remark 3.

One of the advantages of considering a similarity coefficient S which is 3-PSD, is that this weaker condition on S allows us to obtain a function which is defined from S and fulfils the triangular inequality, that is a pseudo-metric. In fact, in [127], the authors proved that if

S is 3-PSD, we can then associate the two pseudo-metrics:

$$D_1 = \arccos(S) \qquad \text{and} \qquad D_2 = \sqrt{1-S}.$$

It is interesting to note that a pseudo-metric will have a different behaviour from that of the dissimilarity defined by $D = 1 - S$ because it satisfies the triangular inequality but it is far away from the Euclidean distance. We encourage the interested reader to consult [127] for numerical experiments and a comparison concerning this issue.

Remark 4. If $S : P(E) \times P(E) \to \mathbb{R}$ is a similarity with $S(X,X) = 1$ for all $X \in P(E)$, then we can conclude that $D' = 1 - S$ is a pseudo-metric if and only if S verifies inequality (2.4). This implies that S is three-positive semi-definite and that in turn means that $D = \sqrt{1-S}$ is a pseudo-metric. If in addition, S verifies the condition that $S(X,Y) = 1$ if and only if $X = Y$, then it holds that $D' = 1 - S$ is a metric if and only if S verifies inequality (2.4). Similarly this implies that S is three-positive semi-definite and which in turn means that in this case, $D = \sqrt{1-S}$ is a metric.

It is worth mentioning here that we were unable to find any matrix $A = (a_{ij})_{1 \leq i,j \leq m}$ with $0 \leq a_{ij} \leq k$, $a_{ii} = k$ and satisfying inequality (2.4) (i.e. $a_{ij} + a_{jl} \leq a_{il} + k$ for any i, j, l) which is not PSD. Thus, we state the following conjecture.

Conjecture 2.7.5. *Let M_S be a similarity matrix and $S : P(E) \times P(E) \to \mathbb{R}^+$ the corresponding similarity coefficient such that $S(X,X) = k$, for all $X \in P(E)$ and $S(X,Y) + S(Y,Z) \leq S(X,Z) + k$, for all $X,Y,Z \in P(E)$. Then M_S is positive semi-definite.*

In fact, we shall give a list of similarity matrices that verify the assumption of the preceding conjecture and are in fact PSD. For this purpose, we need the following theorem given in [18].

Theorem 2.7.6. *[18] Let $s(x,y)$ be a similarity metric, and f a concave function over $[0,\infty)$ satisfying $f(0) \geq 0$, $f(x) > 0$ if $x > 0$ and $f(x) \leq f(y)$ if $x < y$. Then, $\bar{s}(x,y) = \frac{s(x,y)}{f(s(x,x)+s(y,y)-s(x,y))}$ is a similarity metric.*

Example 5. Let E be a finite set and A and $B \in P(E)$. Let $F(A,B) = |A \cap B|$. Then F is a similarity metric by [18]. By Theorem 2.7.6, it follows that

$$F_0(A,B) = \frac{F(A,B)}{F(A,A) + F(B,B) - F(A,B)} = \frac{|A \cap B|}{|A \cup B|}$$

is also a similarity metric. In fact, it is easy to prove recursively that for any nonnegative integer p,

$$F_p(A,B) = \frac{|A \cap B|}{2^p |A \triangle B| + |A \cap B|}$$

is a similarity metric. Indeed for $p = 0$, clearly we obtain the Jaccard similarity $F_0(A,B)$. Suppose that the assumption is true for p. We will prove it for $p+1$. Since $F_{p+1}(A,B) = \frac{|A\cap B|}{|A\cap B|+2^{p+1}|A\triangle B|} = \frac{s_p(A,B)}{s_p(A,A)+s_p(B,B)-s_p(A,B)}$, then by Theorem 2.7.6, F_{p+1} is a similarity metric.

As a result, the Jaccard similarity matrix and the Sokal-Sneath similarity matrix as well as any member of the family T_θ with $\theta = 2^p$ (where p is any nonnegative integer) are all PSD and verify the assumption of Conjecture 2.7.5.

Example 6. Let E be a finite set of cardinal n and A and $B \in P(E)$. The simple matching similarity is defined by $G(A,B) = \frac{|A\cap B|+|A^c\cap B^c|}{n}$. It is easy to see that G verifies conditions 1 to 4 of a similarity metric. For condition 5, we have

$$\begin{aligned}
G(A,B)+G(B,C) &= \frac{|A\cap B|+|A^c\cap B^c|}{n} + \frac{|B\cap C|+|B^c\cap C^c|}{n} \\
&= \frac{|A\cap B|}{n} + \frac{|B\cap C|}{n} + \frac{|A^c\cap B^c|}{n} + \frac{|B^c\cap C^c|}{n} \\
&\leq \frac{|A\cap C|+|B|}{n} + \frac{|A^c\cap C^c|+|B^c|}{n} \\
&= \frac{|A\cap C|+|A^c\cap C^c|}{n} + \frac{|B|+|B^c|}{n} \\
&= \frac{|A\cap C|+|A^c\cap C^c|}{n} + 1 \\
&= G(A,C)+G(B,B).
\end{aligned}$$

Thus G is a similarity metric. Now by Theorem 2.7.6,

$$G_1(A,B) = \frac{G(A,B)}{G(A,A)+G(B,B)-G(A,B)} = \frac{|(A\triangle B)^c|}{|(A\triangle B)^c|+2|A\triangle B|}$$

is also a similarity metric. More generally, for any nonnegative integer p

$$G_p(A,B) = \frac{|(A\triangle B)^c|}{|(A\triangle B)^c|+2^p|A\triangle B|},$$

is a similarity metric. Consequently, the simple matching similarity matrix, Rogers-Tanimoto similarity matrix and any member of the family S_θ with $\theta = 2^p$, are all PSD and verify the assumption of Conjecture 2.7.5.

Remark 5. It is worth mentioning that while writing these thesis, the author in a recent paper [100] shows that the previous conjecture is indeed false in his current form by giving the

following counterexample. The similarity matrix

$$S = \begin{pmatrix} 1 & 0.4 & 0.6 & 0.2 & 0.8 \\ 0.4 & 1 & 0.8 & 0.4 & 0.6 \\ 0.6 & 0.8 & 1 & 0.6 & 0.4 \\ 0.2 & 0.4 & 0.6 & 1 & 0.4 \\ 0.8 & 0.6 & 0.4 & 0.4 & 1 \end{pmatrix},$$

verifies the assumption of the Conjecture 2.7.5, but it is not positive semi-definite.

2.8 Non-PSD similarity matrices

In order to make a complete study of the positive semi-definite property of similarity matrices, we prove by examples that some of similarity matrices that we consider in Section 2.3 are not PSD. Let $E = \{e_1, e_2, e_3, e_4\}$ be a non-empty finite set, and $(A_i)_{1\leq i\leq 3}$ be 3 subsets of E such that

- $A_1 = \{e_1\}$,
- $A_2 = \{e_2\}$,
- $A_3 = \{e_1, e_2\}$.

Next we consider similarity matrices M_t defined on the set $\{A_1, A_2, A_3\}$ and we prove that they are not PSD.

1. $M_{S_{Sor}} = \left(\frac{4|A_i \cap A_j|}{|A_i \triangle A_j| + 4|A_i \cap A_j|}\right)_{1\leq i,j\leq 3} = (T_{1/4}(A_i, A_j))_{1\leq i,j\leq 3}$ is not PSD since $\det(M_{S_{Sor}}) = -7/25 < 0$.

2. $M_{S_{And}} = \left(\frac{8|A_i \cap A_j|}{|A_i \triangle A_j| + 8|A_i \cap A_j|}\right)_{1\leq i,j\leq 3} = (T_{1/8}(A_i, A_j))_{1\leq i,j\leq 3}$ is not PSD since $\det(M_{S_{And}}) = -47/81 < 0$.

3. $M_{S_{Ku2}} = \frac{1}{2}\left(\frac{|A_i \cap A_j|}{|A_i|} + \frac{|A_i \cap A_j|}{|A_j|}\right)_{1\leq i,j\leq 3}$ is not PSD since $\det(M_{S_{Ku2}}) = -1/8 < 0$.

4. $M_{S_{Sim}} = \left(\frac{|A_i \cap A_j|}{min(|A_i|, |A_j|)}\right)_{1\leq i,j\leq 3}$ is not PSD since $\det(M_{S_{Sim}}) = -1 < 0$.

5. $M_{R_{Loe}} = \left(\frac{|A_i \cap A_j|.|A_i^c \cap A_j^c| - |A_i^c \cap A_j|.|A_i \cap A_j^c|}{min(|A_i|.|A_j^c|, |A_j|.|A_i^c|)}\right)_{1 \leq i,j \leq 3}$ is not PSD since $\det(M_{R_{Loe}}) = -16/9 < 0$.

6. $M_{R_{YuQ}} = \left(\frac{|A_i \cap A_j|.|A_i^c \cap A_j^c| - |A_i^c \cap A_j|.|A_i \cap A_j^c|}{|A_i \cap A_j|.|A_i^c \cap A_j^c| + |A_i^c \cap A_j|.|A_i \cap A_j^c|}\right)_{1 \leq i,j \leq 3}$ is not PSD since $\det(M_{S_{YuQ}}) = -4 < 0$.

7. $M_{R_{YuY}} = \left(\frac{\sqrt{|A_i \cap A_j|.|A_i^c \cap A_j^c|} - \sqrt{|A_i^c \cap A_j|.|A_i \cap A_j^c|}}{\sqrt{|A_i \cap A_j|.|A_i^c \cap A_j^c|} + \sqrt{|A_i^c \cap A_j|.|A_i \cap A_j^c|}}\right)_{1 \leq i,j \leq 3}$ is not PSD since $\det(M_{S_{YuY}}) = -4 < 0$.

8. $M_{R_{McC}} = \left(\frac{|A_i \cap A_j|^2 - |A_i^c \cap A_j|.|A_i \cap A_j^c|}{|A_i|.|A_j|}\right)_{1 \leq i,j \leq 3}$ is not PSD since $\det(M_{S_{McC}}) = -1 < 0$.

9. $M_{R_{Fleiss}} = \left(\frac{(|A_i \cap A_j|.|A_i^c \cap A_j^c| - |A_i^c \cap A_j|.|A_i \cap A_j^c|)(|A_i|.|A_j^c| + |A_j|.|A_i^c|)}{2|A_i|.|A_i^c|.|A_j|.|A_j^c|}\right)_{1 \leq i,j \leq 3}$ is not PSD since $\det(M_{R_{Fleiss}}) = -8/27 < 0$.

Remark 6. Note that from the non-PSD similarity measures cited in Section 2.3, Sokal and Sneath 1 similarity matrix is the only matrix that we couldn't show that it is not PSD by a counterexample. It is interesting to mention that Gower and Legendre [36] concluded this result by showing that $\sqrt{1 - S_\Theta}$ is not Euclidean for $\Theta < 1$ (see [36, Theorems 6 and 12]).

2.9 Similarity measures for numerical data

The focus of this chapter is the study of the positive semi-definite property of binary similarity measures. However, one would ask about vector-based similarity measures. Unfortunately, it seems difficult to identify classes of matrices that are PSD for numerical data. In this section, we introduce briefly the case of numerical data, i.e., data represented as real vectors and we give our approach to generalise the results on the similarity matrices taken from binary data to numerical data.

If p is the number of characteristics, then the data space is given now by $\mathcal{X} = \mathbb{R}^p$. Consider two non-zero vectors $x = (x_1, ..., x_p)$ and $y = (y_1, ..., y_p) \in \mathbb{R}^p$. In [59], the authors distinguish between two kinds of numerical similarity measures:

1. Measure derived from dissimilarity measure through a decreasing function.
2. Measure that is an increasing function of dot products. We recall that the scalar product between x and y depends on the angle between the two vectors and their norms

$$< x, y >= \|x\|.\|y\| \cos(x, y).$$

It cannot be used directly as a similarity measure because it doesn't verify the maximality property. However, we can consider the cosine of the angle between the two vectors as a similarity measure. Actually, consider $r(x,y) = \cos(x,y) = \frac{<x,y>}{\|x\|.\|y\|}$. It is easy to see that $r(x,x) = 1$, $-1 \leq r(x,y) = \cos(x,y) \leq 1$ and $r(x,y) = r(y,x)$. Hence, r is a similarity measure. Note that we can normalise r to take value in the interval $[0,1]$. Some generalised similarity measures for numerical data based on dot products are the following:

- Ellenberg similarity : $\frac{\sum(xi+yi)1_{x_iy_i\neq 0}}{\sum(xi+yi)(1+1_{x_iy_i=0})}$,
- Tanimoto similarity : $\frac{\sum xiyi}{\sum xi^2+\sum yi^2-\sum xiyi}$,
- Cosine similarity : $\frac{\sum xiyi}{\sqrt{\sum xi^2}.\sqrt{\sum yi^2}}$,

  and,
- Dice similarity : $\frac{\sum xiyi}{\sum xi^2+\sum yi^2}$.

The binary cases of Ellenberg and Tanimoto similarities coincide; it is the Jaccard similarity S_{Jac}. The binary case of Dice is the Gleason similarity S_{Gleas}.

Some numerical similarity measures can be proved to be positive semi-definite using the inner product of two real vectors. Consider the Cosine coefficient $S_{(1)}(x,y) = \frac{\langle x,y\rangle}{\|x\|.\|y\|}$, and the Dice coefficient $S_{(2)}(x,y) = \frac{2.\langle x,y\rangle}{\|x\|+\|y\|}$ that are proved to be three-positive semi-definite in [127]. However, we can prove easily that they are actually positive semi-definite. These are easy consequences of the results given in Section 2.4, applied to numerical data. In fact, given x_1, x_2,...,$x_m \in \mathbb{R}^n$, the matrix defined by $\left(\langle x_i,x_j\rangle\right)_{1\leq i,j\leq m}$ is PSD since it is a Gram matrix. It follows that the cosine similarity $\left(\frac{\langle x_i,x_j\rangle}{\|x_i\|.\|x_j\|}\right)_{1\leq i,j\leq m}$ is PSD as it is the Hadamard product of a Gram matrix and $\left(\frac{1}{\|x_i\|.\|x_j\|}\right)_{1\leq i,j\leq m}$ (PSD by Lemma 2.4.1). Similarly, the Dice similarity $\left(\frac{2.\langle x_i,x_j\rangle}{\|x_i\|+\|x_j\|}\right)_{1\leq i,j\leq m}$ is positive semi-definite since it is the Hadamard product of a Gram matrix and the Cauchy matrix.

Finally, we strongly believe that with a special choice of the measure μ defined on a finite set E in Theorems 2.4.5 and 2.4.7, we will be able to construct general PSD matrices that will pave the way to prove the positive semi-definite of many numerical similarity matrices. We hope that as well as providing insight into what makes this interesting and practically important problem so difficult our work will prove useful for further development of use of distances in machine learning area.

2.10 Applications

We devote this last section to discuss some applications and motivation for the positive semi-definite and definite property for similarity matrices.

1. The notion of metric between data points is important in machine learning, data mining and pattern recognition. For example, the k-Nearest Neighbor classifier [21] and the prominent K-Means [61], rely on the use of distance measurements between data points. The performance of these methods rely on the quality of the metric. General metrics exist (e.g. the Euclidean distance) but they sometimes fail to give sharp consequences about the peculiar behaviour of the data of interest. We can improve the results when the metric is designed specifically for the task at hand.

   For instance, one can use Mahalanobis distance which originally refers to a distance measure that incorporates the correlation between features:

   $$d_{\text{mah}}(x,y) = \sqrt{(x-y)^\top \Omega^{-1}(x-y)},$$

   where x and y are random vectors from the same distribution with covariance matrix Ω. In addition, it can also refer to more generalized quadratic distances, defined as

   $$d_M(x,y) = \sqrt{(x-y)^\top M(x-y)}, \tag{2.5}$$

   where M is any $n \times n$ symmetric positive semi-definite (PSD) matrix and x and $y \in \mathbb{R}^n$. The positive semi-definite property ensures that d_M satisfies the properties of a pseudo-metric. Note that if $rank(M) = r < n$, then it induces a linear projection of the data into a space of lower dimension r. Hence, when the original space is high- dimensional, it allows cheaper distance computations. These nice properties explain why deriving Mahalanobis distances from positive semi-definite matrices is interesting.

   In fact, a distance model that has been successfully applied to image databases [109] and that has the power to model dependencies between different components of features or histogram vectors is provided by the class of quadratic form distance functions (2.5), where $M = (m_{ij})$ and the weights m_{ij} denote the similarity between the components i and j of the vectors x and y, respectively.

   We note finally that learning Mahalanobis distance has attracted a lot of interest and it is a major component of metric learning [108, 126, 133, 136].

2. A dissimilarity measure d can be derived from a normalised similarity measure s, through decreasing functions. One can for instance consider $d = 1 - s$. However,

such dissimilarity does not verify, in general, the triangular inequality. So, it seems necessary to impose extra conditions on s with the use of other transformations. On the other hand, it has been shown in [36] that if a matrix $S = (s_{i,j})_{1\leq i,j\leq m}$ is a positive semi-definite similarity matrix with elements $0 \leq s_{ij} \leq 1$ and $s_{ii} = 1$ then, the dissimilarity matrix defined by

$$\left(d_{ij}\right)_{1\leq i,j\leq m} = \left(\sqrt{1-s_{i,j}}\right)_{1\leq i,j\leq m} \tag{2.6}$$

is Euclidean. Consequently for all matrices defined in Section 2.5 verifying $0 \leq s_{ij} \leq 1$ and $s_{ii} = 1$, their dissimilarity matrices defined above are Euclidean.

Consider the example of K-means clustering ([65]) which is essentially a method of vector quantisation, that is popular for cluster analysis in data mining (see [49] and the references within). More specifically, K-means clustering aims to partition n observations into K clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. Let $X = \{x_i \ : \ i = 1,2,...,n\}$ be a set of n d-dimensional points to be clustered into a set of K clusters, $c_k,\ k = 1,...,K$. Generally, a K-means algorithm aims at finding a partition such that the squared error between the empirical mean of a cluster and the points in the cluster is minimised. Explicitly, if μ_k denotes the mean of the cluster c_k, then the squared error between μ_k and the points in cluster c_k is defined as $J(c_k) = \sum_{x_i\in c_k} \| x_i - \mu_k \|^2$. So now the goal of K-means is to minimise the sum of the following squared error over all K clusters,

$$J(C) = \sum_{k=1}^{K} \sum_{x_i\in c_k} \| x_i - \mu_k \|^2 .$$

In addition, K-means are appropriate to use in combination with the Euclidean distance because the main objective of K-means is to minimise the sum of intra-cluster variances (i.e. J(C)), and the intra-cluster variance is calculated in the same way as the sum of Euclidean distances between all points in the cluster to the cluster centroid. So, there is a close link between K-means and the Euclidean distance as the algorithm is somewhat designed to calculate the mean of a set of data points, but the convergence of the clustering process is guaranteed by the mean only if the data points are reassigned to the nearest centroid by the use of the Euclidean distance.

Finally it is interesting to point out that positive semi- definite similarity measures can be converted to the Euclidean distance using the transformation (2.6). An example of that is the cosine similarity (See [106]).

It is worth mentioning that we can still use K-means with other distance measures. As in [112] where the author uses this algorithm with the Minkowski and Tchebychev distances. However, in these cases, convergence is not guaranteed.

3. In evidence theory, the use of a distance between belief functions can be practical as a definition of agreement between sources of information. Let X be a frame of discernment containing n distinct objects. For a subset $A \subseteq X$, called hypothesis, a belief function assigns a belief value to each hypothesis based on one or more pieces of evidence (see [11] for a formal definition of belief functions). Many distances between belief functions have been defined, among which generalised Euclidean distances play a major role. A generalised Euclidean distance between two belief functions m_1 and m_2, is of the form

$$d_W^2(m_1,m_2) = \sqrt{(m_1-m_2)^\top W(m_1-m_2)}, \tag{2.7}$$

where W is a weighted matrix. From the properties of W are derived the properties of the associated distance d_W^2. In particular, if W is positive definite , then d_W^2 is a full metric distance. However, if W is only positive semi-definite, then d_W^2 is a pseudo-metric, which means that two distinct belief functions may have a null distance. In the aim of defining a "full" metric between two belief functions which accounts for the interaction between focal elements, an Euclidean distance of the form of (2.7) was proposed in which the weighted matrix is **Jac**. It is worth noting that $d_{\mathbf{Jac}}^2$ is proved to be a metric distance (or a "full" metric) due to the positive definiteness of the complete Jaccard index **Jac** [11]. In addition, $d_{\mathbf{Jac}}^2$ can be extended by replacing **Jac** by any positive definite similarity matrix. By Section 2.6, any matrix that belongs to the Tversky family of similarity matrices is appropriate as a weighted matrix.

4. Kernel methods [110] are widely used in machine learning. The main idea behind these methods is based on kernels or kernel functions.

   Let $\mathcal{X}$ be a non-empty set. The idea is to define a kernel $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ such that for any two points x and $x' \in \mathcal{X}$, $K(x,x')$ be equal to an inner product of vectors $\Phi(x)$ and $\Phi(y)$,

   $$\forall x,x' \in \mathcal{X}, K(x,x') = < \Phi(x), \Phi(y) >,$$

   for some mapping $\Phi : \mathcal{X} \to \mathbb{H}$ to a Hilbert space $\mathbb{H}$. Since an inner product is a measure of similarity of two vectors, K is often interpreted as a similarity measure between elements of the input space $\mathcal{X}$. An important advantage of such a kernel K is that K is often more efficient to compute than Φ and an inner product in $\mathcal{H}$. In fact, there is

no need to explicitly define or compute a mapping Φ. The kernel K can be arbitrarily chosen so long as the existence of Φ is guaranteed, i.e. K satisfies Mercer's condition (see [110]). This condition is essential to guarantee the convexity of the optimisation problem for algorithms such as SVM's and thus convergence in this case is obtained.

On the other hand, a condition which is equivalent to Mercer's condition is that the kernel K be symmetric positive definite. A symmetric function $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is a positive definite kernel on $\mathcal{X}$ if

$$\sum_{i,j=1}^{n} c_i c_j K(x_i, x_j) \geq 0,$$

holds for any $n \in \mathbb{N}, x_1, ..., x_n \in \mathcal{X}$ and $c_1, ..., c_n \in \mathbb{R}$. Equivalently, a kernel K is positive definite if and only if for any $N \in \mathbb{N}$ and any set of points $(x_1, x_2, ..., x_n) \in \mathcal{X}^n$, the matrix $\mathbf{K} := [K(x_i, x_j)]_{ij} \in \mathbb{R}^{n \times n}$ is positive semi-definite. Noting that kernel methods are essentially algorithms that take such matrices as input, the requirement of positive semi-definiteness is important when seen from at least two perspectives as follows. First, the usage of positive definite matrices is a key assumption in convex programming [12]. In practice, the positive definiteness of kernel matrices ensures that kernel algorithms such as Gaussian processes or support vector machines (SVMs) converge to a relevant solution. Second, the positive definiteness assumption is also a key assumption in reproducing kernel Hilbert spaces (RKHS)[6, 107].

Any function K that creates a symmetric, positive definite matrix is a valid kernel. Hence, similarity matrices given in Section (2.5) can be used as kernels.

It is worthy to note that some methods deals with the problem of learning with a non- PSD similarity matrix in kernel machines. In Support Vector Machines (SVMs), these methods can be divided into two approaches: algorithmic [91] and spectrum-transformation [135].

Chapter 3

Positive semi-definite pth roots of positive semi-definite doubly stochastic matrices

We saw in Chapter 2, that the property of being positive semi-definite has many applications especially when we are considering similarity matrices. A particular form of a similarity matrix can be a doubly stochastic matrix [130]. Moreover, some algorithms are proposed to improve clustering by learning a doubly stochastic matrix from a data similarity matrix [130]. In addition, exploring the structure of the spectrum of symmetric doubly stochastic matrices can be helpful in many applications such as data clustering. We take the example of a data clustering method introduced in [75] where the clustering algorithm takes as input the consensus similarity matrix S created from whatever combination of clustering methods . Then, S is converted into the doubly stochastic matrix P using the Sinkhorn-Knopp algorithm [114]. Next, all the eigenvalues of P are computed, and the Perron cluster of P is identified. Generally, the number of eigenvalues near 1 is the number of clusters. Note that stochastic consensus clustering works on several well-known test data sets and can give better results [75].

In this chapter, our problem of interest is to combine the positive semi-definite property and the doubly stochastic property in some matrix A, then to consider the solution of the equation $X^p = A$, where p is a positive integer .

3.1 Introduction of the problem of interest

One of the most intriguing problems in matrix analysis is the calculus of matrix functions. In particular, pth roots of matrices play an important role in many applications. Indeed, finding roots of certain classes of matrices, and especially nonnegative matrices, are widely

studied (see, e.g., [73, 74, 89, 125]). Recently, finding the pth roots of a stochastic matrix have been discussed in [37, 38] as the computation of such roots has led to many interesting applications in the area of financial mathematics[46, 132]. In fact, stochastic matrices arise in Markov chain models. A stochastic matrix $B = (b_{ij})$ appears as a transitive matrix over a certain time interval to describe the transition probabilities from state i to state j. Often, a transition matrix over a shorter time interval is needed but only a stochastic matrix over a longer time interval is available. Such a transition matrix can be obtained by calculating pth roots of the original matrix [17]. We will extend the analysis of the finding roots problem to PSD doubly stochastic matrices.

We recall that for every $A \succeq 0$ and for every positive integer p, there exists a unique $X \succeq 0$ so that $X^p = A$ (see Theorem 1.2.9). Then the matrix X is called the positive semi-definite *pth root* of A and is denoted by $A^{1/p}$. The positive semi-definite pth root of a positive semi-definite doubly stochastic matrix is quasi-doubly stochastic (see Lemma 3.2.2 below), but need not to be nonnegative. This leads us to the following problem.

Problem 2. *What are the necessary and sufficient conditions for a positive semi-definite doubly stochastic matrix to have its positive semi-definite pth root, doubly stochastic ?*

While the computation of the pth roots of a (not necessarily positive semi-definite) stochastic (or doubly stochastic) matrix involves a lot of problems [41], like under what conditions does a given stochastic matrix have a stochastic pth root? how many roots are there? and how they can be computed, our problem of interest here is more restrictive, since we are interested in the unique positive semi-definite doubly stochastic pth root of a positive semi-definite doubly stochastic matrix. In addition, as mentioned earlier, such a matrix is always doubly quasi-stochastic, so the problem is reduced to studying the nonnegativity of the matrix. To the best of our knowledge, this problem was only considered by Marcus and Minc [67] for the particular case $p = 2$ (i.e. square root) and their results are given in Theorem 3.2.1 below.

We shall begin by establishing some notations.

Notations. Let $\mathbb{M}_s(n)$ be the class of all $n \times n$ symmetric real matrices and $\mathbb{M}_s^+(n)$ denote the convex cone of all nonnegative elements in $\mathbb{M}_s(n)$. In addition, the set of all $n \times n$ symmetric doubly stochastic matrices will be denoted by Δ_n^s. Recall here that Δ_n^s is a convex polytope of dimension $\frac{1}{2}n(n-1)$, where its vertices were determined in [54] (see also [22]). On the other hand, we shall denote by H_n to be the closed convex cone of all $n \times n$ real positive semi-definite matrices, and we also denote by $K_n := H_n \bigcap \Delta_n^s$ to be the convex set of all $n \times n$ positive semi-definite doubly stochastic matrices. Finally, for any positive integer p, we define

$$K_n^{1/p} := \{A \in K_n : A^{1/p} \text{ is doubly stochastic}\}.$$

This chapter is organised as follows. Section 2 is concerned with giving new sufficient conditions for the square root of an $n \times n$ positive semi-definite doubly stochastic matrix to be nonnegative. In Section 3, we prove some geometrical properties of the set $K_n^{1/p}$. More precisely, our main result deals with proving that for $n \geq 3$, the set $K_n^{1/p}$ is not convex but star convex with respect to J_n. Next, we will identify a large convex set that sits inside $K_n^{1/p}$. In Section 4, we shall use the theory of M-matrices to present a method for constructing elements in $K_n^{1/p}$. Then, in Section 5, we investigate in depth the way of finding elements in $K_n^{1/p}$ via the use of eigenvalues. Finally, the last section deals with the square root of all positive semi-definite doubly stochastic matrices of order 3. In addition, a generalisation to order n gives us a family of elements that belong to $K_n^{1/2}$.

3.2 Square root of a PSD doubly stochastic matrix

In this section, we shall give some new sufficient conditions for a positive semi-definite doubly stochastic matrix of order n to have a doubly stochastic square root. As mentioned earlier, the square root of a positive semi-definite doubly stochastic matrix is not, in general, doubly stochastic as we can see in the example below.

Example 7. Consider the positive semi-definite doubly stochastic matrix defined by $A = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{7}{12} & \frac{1}{12} \\ \frac{1}{3} & \frac{1}{12} & \frac{7}{12} \end{pmatrix}$. The square root of A is the doubly quasi-stochastic matrix

$$A^{1/2} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3}+\frac{1}{4}\sqrt{2} & \frac{1}{3}-\frac{1}{4}\sqrt{2} \\ \frac{1}{3} & \frac{1}{3}-\frac{1}{4}\sqrt{2} & \frac{1}{3}+\frac{1}{4}\sqrt{2} \end{pmatrix}$$

with its (2,3)-entry $\frac{1}{3}-\frac{1}{4}\sqrt{2}$ is negative. □

Remark 7. Note that in the previous example, A has a positive semi-definite doubly quasi-stochastic pth root with negative entries for all even integer $p \geq 2$. We can prove that by contradiction. Suppose that there exists an even integer p_0 such that A^{1/p_0} is doubly stochastic. As p_0 is even, $p_0 = 2k$ with $k > 1$ a positive integer. Hence, $(A^{1/p_0})^k = A^{1/2}$ is doubly stochastic, which is a contradiction.

In [67], Marcus and Minc gave the following sufficient condition for Problem 2 to be solvable for the case $p = 2$.

Theorem 3.2.1. *[67] The square root of a positive semi-definite n-square doubly stochastic matrix $A = (a_{ij})$ is doubly quasi-stochastic. If $a_{ii} \leqslant \frac{1}{n-1}$, for $i = 1, ..., n$, then $A^{1/2}$ is doubly stochastic.*

Proof. See [67, Theorem 2] or [77, Chapter 5, Theorem 4.2]. □

Before doing a refinement of the previous theorem, we will prove the following lemma.

Lemma 3.2.2. *Let $A = (a_{ij})$ be an $n \times n$ positive semi-definite doubly quasi-stochastic matrix. Then, A has a unique positive semi-definite doubly quasi-stochastic pth root $B = (b_{ij})$ for any positive integer $p \geq 1$.*

Proof. Let $1, \lambda_2, ..., \lambda_n$ be the nonnegative eigenvalues of the positive semi-definite matrix A. Then by spectral theorem, there exists a real orthogonal matrix $V = (e_n, v_2, ..., v_n)$ whose first column is e_n and each of the last $n-1$ remaining column sums is zero, and such that

$$V^T AV = \text{diag}(1, \lambda_2, ..., \lambda_n).$$

It is easy to show that a pth root of A is given by

$$B = (b_{ij}) = V\text{diag}(1, \sqrt[p]{\lambda_2}, ..., \sqrt[p]{\lambda_n})V^T,$$

since $B^p = A$. Moreover, B being symmetric with nonnegative eigenvalues, is the unique positive semi-definite matrix with $A = B^p$. Next, a simple check shows that

$$\begin{aligned} Be_n &= V\text{diag}(1, \sqrt[p]{\lambda_2}, ..., \sqrt[p]{\lambda_n})V^T e_n \\ &= V\text{diag}(1, \sqrt[p]{\lambda_2}, ..., \sqrt[p]{\lambda_n}).(1,0,...,0)^T \\ &= V(1,0,...,0)^T = e_n. \end{aligned}$$

Similarly, $e_n^T B = e_n^T V\text{diag}(1, \sqrt[p]{\lambda_2}, ..., \sqrt[p]{\lambda_n})V^T = e_n^T$. Thus, B is doubly quasi-stochastic. □

In order to make a refinement of Theorem 3.2.1, let us firstly recall that if $A = (a_{ij}) \in M_n(\mathbb{C})$ is positive semi-definite, then

$$a_{ii} = (Af_i, f_i) \geqslant 0, \qquad \forall 1 \leqslant i \leqslant n,$$

with f_i the ith column of the identity matrix I_n.

Theorem 3.2.3. *Let $A = (a_{ij})$ be a positive semi-definite n-square doubly stochastic matrix with at least $(n-1)$ main diagonal entries $\leqslant \frac{1}{n-1}$. Then there exists a doubly stochastic matrix $B = (b_{ij})$ such that $B^2 = A$.*

Proof. By taking $p = 2$ in Lemma 3.2.2, we conclude that B is doubly quasi-stochastic. Since $b_{ii} \geq 0$ for all $i = 1, ..., n$, in order to complete the proof, it remains to prove that $b_{ij} \geqslant 0$, for all $1 \leqslant i \neq j \leqslant n$. Suppose that $b_{pq} < 0$ for some $p \neq q$, then

$$\begin{aligned} a_{pp} &= \sum_{j=1}^{n} b_{pj}^2 \\ &> \sum_{j\neq q} b_{pj}^2 && \text{(since } b_{pq} < 0\text{)} \\ &\geq \frac{1}{n-1}\left(\sum_{j\neq q} b_{pj}\right)^2 && \text{(by the Cauchy-Shwarz inequality)} \\ &> \frac{1}{n-1} && \text{(since } \sum_{j\neq q} b_{pj} = \sum_{j=1}^{n} b_{pj} - b_{pq} = 1 - b_{pq} > 1\text{)}. \end{aligned}$$

Since B is symmetric, then $b_{qp} = b_{pq} < 0$ and so by repeating the same process on b_{qp} instead of b_{pq}, we also get $a_{qq} > \frac{1}{n-1}$. Since $p \neq q$, this contradicts the hypothesis. □

Next, we give the following remarks.

Remark 8. Let A be a positive semi-definite doubly stochastic matrix and $B := (b_{ij})$ be the square root of A. It is worthy to mention here that $a_{ii} \geq \frac{1}{n} > 0$ and $b_{ii} \geq \frac{1}{n} > 0$.

In fact, from the equation $A = B^2$, we get that $a_{ii} = \sum_{j=1}^{n} b_{ij}^2$ for all $i = 1, ..., n$. Now, using the Cauchy-Schwarz inequality, we obtain

$$\left(\sum_{j=1}^{n} b_{ij}\right)^2 \leqslant n \sum_{j=1}^{n} b_{ij}^2,$$

for every $i = 1, ..., n$. Since B is always doubly quasi-stochastic matrix (see the proof of the preceding theorem), then the preceding inequality becomes

$$a_{ii} = \sum_{j=1}^{n} b_{ij}^2 \geq \frac{1}{n},$$

for all $i = 1, ..., n$. The same process can be repeated but with B is now playing the role previously played by A.

Remark 9. It is worthy to observe that in the case of Marcus and Minc's result given in Theorem 3.2.1, the positive semi-definite doubly stochastic matrix $A = (a_{ij})$ is taken such that

$$1 \leq Tr(A) \leq 1 + \frac{1}{n-1},$$

since $\frac{1}{n} \leqslant a_{ii} \leqslant \frac{1}{n-1}$, for all $i = 1, ..., n$. In our case (Theorem 3.2.3), since $\frac{1}{n} \leqslant a_{ii} \leqslant \frac{1}{n-1}$ for all $i \neq r$ and $\frac{1}{n} \leqslant a_{rr} \leqslant 1$, then A is taken with the condition that

$$1 \leq Tr(A) \leq 2.$$

Of course, these conditions are not sufficient to have a doubly stochastic square root, as can be easily seen in Example 7. In there, $Tr(A) = \frac{18}{12}$ but A does not have a doubly stochastic square root.

The next theorem gives a sufficient condition on the trace.

Theorem 3.2.4. *Let $A = (a_{ij})$ be a n-square positive semi-definite doubly stochastic matrix such that $Tr(A) \leqslant \frac{n^2-n+2}{n^2-n}$. Then, there exists a doubly stochastic matrix B such that $B^2 = A$.*

Proof. Let the square root $B = (b_{ij})$ of A be defined as in the proof of Lemma 3.2.2. Therefore B is positive semi-definite quasi-doubly stochastic. Suppose that $b_{pq} = b_{qp} < 0$ for some p and q with $p < q$. Then,

$$\begin{aligned}
Tr(A) &= a_{11} + ... + a_{pp} + ... + a_{qq} + ... + a_{nn} \\
&= \sum_{k=1}^{n} b_{1k}^2 + ... + \sum_{k=1}^{n} b_{pk}^2 + ... + \sum_{k=1}^{n} b_{qk}^2 + ... + \sum_{k=1}^{n} b_{nk}^2 \\
&= \sum_{k=1}^{n} b_{pk}^2 + \sum_{k=1}^{n} b_{qk}^2 + \sum_{i \neq \{p,q\}} \sum_{k=1}^{n} b_{ik}^2 \\
&> \sum_{k \neq q} b_{pk}^2 + \sum_{k \neq p} b_{qk}^2 + \sum_{i \neq \{p,q\}} \sum_{k=1}^{n} b_{ik}^2 \\
&\geqslant \frac{1}{n-1} \left(\sum_{k \neq q} b_{pk} \right)^2 + \frac{1}{n-1} \left(\sum_{k \neq p} b_{qk} \right)^2 + \sum_{i \neq \{p,q\}} \frac{1}{n} \left(\sum_{k=1}^{n} b_{ik} \right)^2 \\
&> \frac{1}{n-1}(1) + \frac{1}{n-1}(1) + \frac{n-2}{n} \\
&= \frac{n^2 - n + 2}{n(n-1)},
\end{aligned}$$

which contradicts the assumption of the hypothesis. □

Obviously, the trace condition of the preceding theorem is not necessary as shows the following example.

Example 8. Consider $n = 4$. The square root of the matrix $B = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$ is B itself. However, $Tr(B) = 2 > \frac{7}{6}$. □

3.3 Some geometrical properties of the set $K_n^{1/p}$

In this section, we shall explore some geometrical properties of the set $K_n^{1/p}$ where all the positive semi-definite doubly stochastic matrices that have positive semi-definite doubly stochastic pth roots, lie. But first, we shall start with the following definition and propositions which are needed for our purposes.

Definition 3.3.1. A subset Γ of $\mathbb{M}_s^+(n)$ is said to be *star convex* with respect to a point $x \in \mathbb{M}_s^+(n)$, if the line from any point in the set Γ to x is also contained in Γ.

Proposition 3.3.2. *Let p, n and m be any positive integers. If $X \in K_n^{1/p}$ and $Y \in K_m^{1/p}$, then $X \oplus Y \in K_{n+m}^{1/p}$.*

Proof. As $X \in K_n^{1/p}$ and $Y \in K_m^{1/p}$, then $X^{1/p}$ and $Y^{1/p}$ are doubly stochastic. Finally, noting that

$$X \oplus Y = \underbrace{(X^{1/p} \oplus Y^{1/p})(X^{1/p} \oplus Y^{1/p}).....(X^{1/p} \oplus Y^{1/p})}_{p \text{ times}},$$

we have $(X \oplus Y)^{1/p} = (X^{1/p} \oplus Y^{1/p})$. Then, we conclude that $(X \oplus Y)^{1/p}$ is also doubly stochastic since it is the product of p doubly stochastic matrices. Hence, $X \oplus Y \in K_{n+m}^{1/p}$. □

As a consequence, we have the following corollary.

Corollary 3.3.3. *Let p, n and m be any positive integers such that $m < n$. Then the set defined by*

$$K_m^{1/p} \oplus K_{n-m}^{1/p} := \{X \oplus Y \ \text{ where } X \in K_m^{1/p} \ \text{ and } Y \in K_{n-m}^{1/p}\}$$

is a subset of $K_n^{1/p}$.

Proposition 3.3.4. *Let p be any positive integer. If $X \in K_n^{1/p}$, then, $Q^T X Q \in K_n^{1/p}$ for any $n \times n$ permutation matrix Q.*

Proof. Let $X \in K_n^{1/p}$, then $X^{1/p}$ is doubly stochastic, this implies that $Q^T X^{1/p} Q$ is doubly stochastic, for any permutation matrix Q, since it is the product of three doubly stochastic matrices. By noticing that $\left(Q^T X^{1/p} Q\right)^p = Q^T X Q$, we conclude that $Q^T X Q \in K_n^{1/p}$. □

Finally, we have the following lemma.

Lemma 3.3.5. *Let*

$$y = y(x) := \frac{1}{3} + \frac{1}{6} 4^{-x^{-1}} - \frac{1}{2} 3^{x^{-1}} 4^{-x^{-1}}$$

be a function of x which is defined on $[2, +\infty[$. *Then,*

1. $y(2) < 0$,
2. *y is negative on* $[3, +\infty[$.

Proof.

1. A simple check shows that $y(2) = -0.0163460353 < 0$.
2. The derivative of y is given by:

$$\begin{aligned} y' &= \frac{1}{3} 4^{-x^{-1}} \ln(2) x^{-2} + \frac{1}{2} 3^{x^{-1}} \ln(3) 4^{-x^{-1}} x^{-2} - 3^{x^{-1}} 4^{-x^{-1}} \ln(2) x^{-2} \\ &= x^{-2} \left(\frac{\ln(2)}{3} 4^{-x^{-1}} + \frac{\ln(3)}{2} 3^{x^{-1}} 4^{-x^{-1}} - \ln(2) 3^{x^{-1}} 4^{-x^{-1}} \right) \\ &= \frac{1}{2x^2} \left(\frac{\ln(4)}{3} 4^{-\frac{1}{x}} + \ln\left(\frac{3}{4}\right) 3^{\frac{1}{x}} 4^{-\frac{1}{x}} \right) \\ &= \frac{4^{-\frac{1}{x}}}{2x^2} \left(\frac{\ln(4)}{3} - \ln\left(\frac{4}{3}\right) 3^{\frac{1}{x}} \right). \end{aligned}$$

   Since the first term of the product in y' is positive, then y' is positive if and only if $\frac{\ln(4)}{3} - \ln\left(\frac{4}{3}\right) 3^{\frac{1}{x}} \geq 0$, which implies that $\frac{\ln(4)}{3\ln\left(\frac{4}{3}\right)} = 1.606280561.. \geq 3^{\frac{1}{x}}$ which in turn is equivalent to saying that $x \geq \frac{\ln(3)}{\ln\left(\frac{\ln(4)}{3\ln\left(\frac{4}{3}\right)}\right)} = 2.318132354...$, so that clearly y is strictly increasing in $[3, +\infty[$. Moreover, $y(3) = -0.0166623670...$ and y goes to zero as x tends to infinity, then we conclude that y is always negative in $[3, +\infty[$.

□

Next, we give the main result of this section which deals with certain geometrical properties of the sets K_n and $K_n^{1/p}$.

Theorem 3.3.6. *Let p be any positive integer. Then the following statements hold.*

1. *K_n is convex for any $n \geq 1$.*
2. *$K_2^{1/p}$ is convex.*
3. *For $n \geq 3$, $K_n^{1/p}$ is not convex.*
4. *$K_n^{1/p}$ is star convex with respect to J_n.*

Proof.

1. It suffices to note that $K_n = H_n \bigcap \Delta_n^s$ where H_n and Δ_n^s are both convex in $\mathbb{M}_s^+(n)$. Hence, their intersection K_n is convex.
2. It suffices to prove that $K_2^{1/p}$ is the line segment $[I_2, J_2]$. In fact, Δ_2^s is equal to the line segment $[I_2, C_2]$ where $C_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ hence obviously, K_2 is equal to the line segment $[I_2, J_2]$. Now let $X \in K_2$, then $X = aI_2 + (1-a)J_2$ for some a with $0 \leq a \leq 1$. Consider now the matrix $V_a = \sqrt[p]{a}I_2 + (1 - \sqrt[p]{a})J_2$. Using the binomial formula, we have,

$$\begin{aligned}
V_a^p &= \left(\sqrt[p]{a}I_2 + (1-\sqrt[p]{a})J_2\right)^p \\
&= \sum_{k=0}^{p} \binom{p}{k} \left(\sqrt[p]{a}I_2\right)^k \left((1-\sqrt[p]{a})J_2\right)^{p-k} \qquad \left(\text{where } \binom{p}{k} := \frac{p!}{(p-k)!k!}\right) \\
&= (a)I_2^p + \sum_{k=0}^{p-1} \binom{p}{k} \left(\sqrt[p]{a}\right)^k \left((1-\sqrt[p]{a})\right)^{p-k} I_2^k J_2^{p-k} \\
&= aI_2 + \sum_{k=0}^{p-1} \binom{p}{k} \left(\sqrt[p]{a}\right)^k \left((1-\sqrt[p]{a})\right)^{p-k} J_2 \\
&= aI_2 - aJ_2 + \sum_{k=0}^{p} \binom{p}{k} \left(\sqrt[p]{a}\right)^k \left((1-\sqrt[p]{a})\right)^{p-k} J_2 \\
&= aI_2 - aJ_2 + \left(\sqrt[p]{a} + (1-\sqrt[p]{a})\right)^p J_2 \\
&= aI_2 - aJ_2 + J_2 \\
&= aI_2 + (1-a)J_2.
\end{aligned}$$

Then, $X^{1/p} = \sqrt[p]{a}I_2 + (1-\sqrt[p]{a})J_2$. It follows that $X^{1/p} \in [I_2, J_2]$. Therefore, $K_2^{1/p}$ is the line segment $[I_2, J_2]$.

3. For $n = 3$, consider the following two positive semi-definite doubly stochastic matrices:

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix} \text{ and } S = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

A simple check shows that $R^{1/p} = R$ and $S^{1/p} = S$. However

$$M := \frac{R+S}{2} = \begin{pmatrix} 3/4 & 1/4 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1/4 & 3/4 \end{pmatrix}$$

has its positive semi-definite pth root with negative entries. Indeed, it is easy to check that if Q is the orthogonal matrix given by

$$Q := \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \end{pmatrix},$$

and

$$D := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 3/4 \end{pmatrix}$$

then $Q^T M Q = D$ and hence $M^{1/p} = Q D^{1/p} Q^T$. More explicitly,

$$M^{1/p} = \begin{pmatrix} \frac{1}{3} + \frac{1}{6}4^{-p^{-1}} + \frac{1}{2}3^{p^{-1}}4^{-p^{-1}} & \frac{1}{3} - \frac{1}{3}4^{-p^{-1}} & \frac{1}{3} + \frac{1}{6}4^{-p^{-1}} - \frac{1}{2}3^{p^{-1}}4^{-p^{-1}} \\ \frac{1}{3} - \frac{1}{3}4^{-p^{-1}} & \frac{1}{3} + \frac{2}{3}4^{-p^{-1}} & \frac{1}{3} - \frac{1}{3}4^{-p^{-1}} \\ \frac{1}{3} + \frac{1}{6}4^{-p^{-1}} - \frac{1}{2}3^{p^{-1}}4^{-p^{-1}} & \frac{1}{3} - \frac{1}{3}4^{-p^{-1}} & \frac{1}{3} + \frac{1}{3}4^{-p^{-1}} + \frac{1}{3}3^{p^{-1}}4^{-p^{-1}} \end{pmatrix}.$$

Now by noticing that the (1,3)-entry of $M^{1/p}$ is equal to $y(p)$ with y the function given in Lemma 3.3.5, we conclude that $M^{1/p}$ has its (1,3)-entry negative for all $p = 2, 3, \ldots$.

For general n, consider the matrices $R_n = R \oplus I_{n-3}$ and $S_n = S \oplus I_{n-3}$. It is easy to see that $R_n^{1/p} = R_n$, $S_n^{1/p} = S_n$ and $M_n := \frac{R_n+S_n}{2} = M \oplus I_{n-3}$. But then $M_n^{1/p} = M^{1/p} \oplus I_{n-3}$ has a negative entry for all $p = 2, 3, \ldots$, which means that M_n is not in $K_n^{1/p}$ for all $p = 2, 3, \ldots$. This concludes the proof that $K_n^{1/p}$ is not convex.

4. Let $A \in K_n^{1/p}$, then there exists a positive semi-definite doubly stochastic matrix X such that $A = X^p$. For $0 \leq a \leq 1$, consider the matrix

$$W_a := \left(\sqrt[p]{1-a}X + (1-\sqrt[p]{1-a})J_n \right).$$

Using again the binomial formula, we can write

$$\begin{aligned}
W_a^p &= \left(\sqrt[p]{1-a}X + (1-\sqrt[p]{1-a})J_n \right)^p \\
&= \sum_{k=0}^{p} \binom{p}{k} \left(\sqrt[p]{1-a}X \right)^k \left((1-\sqrt[p]{1-a})J_n \right)^{p-k} \\
&= (1-a)X^p + \sum_{k=0}^{p-1} \binom{p}{k} \left(\sqrt[p]{1-a} \right)^k \left((1-\sqrt[p]{1-a}) \right)^{p-k} X^k J_n^{p-k} \\
&= (1-a)X^p + \sum_{k=0}^{p-1} \binom{p}{k} \left(\sqrt[p]{1-a} \right)^k \left((1-\sqrt[p]{1-a}) \right)^{p-k} J_n \\
&\qquad \Big(\text{as } X^k \text{ is doubly stochastic so } X^k J_n^{p-k} = X^k J_n = J_n \Big) \\
&= (1-a)X^p - (1-a)J_n + \sum_{k=0}^{p} \binom{p}{k} \left(\sqrt[p]{1-a} \right)^k \left((1-\sqrt[p]{1-a}) \right)^{p-k} J_n \\
&= (1-a)X^p - (1-a)J_n + \left(\sqrt[p]{1-a} + (1-\sqrt[p]{1-a}) \right)^p J_n \\
&= (1-a)X^p - (1-a)J_n + J_n \\
&= (1-a)A + aJ_n.
\end{aligned}$$

Each point on the line-segment $[J_n A]$ has a doubly stochastic pth root on the line segment $[J_n X]$. This completes the proof.

□

Using a similar proof, we obtain the following conclusion.

Corollary 3.3.7. *Let p, n and m be any positive integers such that $m < n$. Then the subset $K_m^{1/p} \oplus K_{n-m}^{1/p}$ of $K_n^{1/p}$, is star convex with respect to $J_m \oplus J_{n-m}$.*

Next, we include an example in the sake of illustrating the previous corollary.

Example 9. Consider the following 7×7 positive semi-definite doubly stochastic matrix

$$M = \begin{pmatrix} \frac{2}{3} & \frac{1}{6} & \frac{1}{6} & 0 & 0 & 0 & 0 \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & 0 & 0 & 0 & 0 \\ \frac{1}{6} & \frac{1}{6} & \frac{2}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{3}{8} & \frac{1}{8} & \frac{3}{8} & \frac{1}{8} \\ 0 & 0 & 0 & \frac{1}{8} & \frac{3}{8} & \frac{1}{8} & \frac{3}{8} \\ 0 & 0 & 0 & \frac{3}{8} & \frac{1}{8} & \frac{3}{8} & \frac{1}{8} \\ 0 & 0 & 0 & \frac{1}{8} & \frac{3}{8} & \frac{1}{8} & \frac{3}{8} \end{pmatrix}.$$

A simple check shows that

$$M = \frac{1}{2}(R \oplus B) + \frac{1}{2}(J_3 \oplus J_4),$$

where

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix} \text{ and } B = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}.$$

Using Example 8 and the proof of Theorem 3.3.6 Part 3) , we know that $R \in K_3^{1/p}$ and $B \in K_4^{1/p}$ for every positive integer p. In view of the preceding corollary, we conclude that M is an element of $K_3^{1/p} \oplus K_4^{1/p}$, which implies that $M \in K_7^{1/p}$ for every positive integer p, by Corollary 3.3.3.

Next, we identify a large convex subset in $K_n^{1/p}$.

Theorem 3.3.8. *Let $X_1 := J_n$, $X_2 := 1 \oplus J_{n-1}$,..., $X_{n-1} := I_{n-2} \oplus J_2$, $X_n := I_n$, and define T to be the convex hull of the matrices*

$$\{X_1, X_2, ..., X_n\}.$$

Then T is contained in $K_n^{1/p}$ for any positive integer p.

Proof. Consider the following orthogonal matrix U_n defined by

$$U_n = \begin{pmatrix} \frac{1}{\sqrt{n}} & \frac{-(n-1)}{\sqrt{n(n-1)}} & 0 & 0 & \cdots & 0 \\ \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n(n-1)}} & \frac{-(n-2)}{\sqrt{(n-1)(n-2)}} & 0 & \cdots & 0 \\ \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{(n-1)(n-2)}} & \frac{-(n-3)}{\sqrt{(n-2)(n-3)}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{(n-1)(n-2)}} & \frac{1}{\sqrt{(n-2)(n-3)}} & \cdots & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{(n-1)(n-2)}} & \frac{1}{\sqrt{(n-2)(n-3)}} & \cdots & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Let $\Lambda_i = \text{diag}(\underbrace{1,1,...,1}_{i \text{ times}},0,...,0)$, for each $i = 1,...,n$. A simple check shows that $U_n^T X_i U_n = \Lambda_i$ for all $i = 1,...,n$. Hence, if $X := x_1X_1 + ... + x_nX_n$, then clearly, we have

$$\begin{aligned} X &= x_1U_n\Lambda_1U_n^T + ... + x_nU_n\Lambda_nU_n^T \\ &= U_n(x_1\Lambda_1 + x_2\Lambda_2 + ... + x_n\Lambda_n)U_n^T \end{aligned}$$

which implies

$$X = U_n\left(\text{diag}\left(\lambda_1,\lambda_2,...,\ \lambda_{n-1},\lambda_n\right)\right)U_n^T, \tag{3.1}$$

where $\lambda_i = \sum_{k=i}^n x_k$. Noticing that for each $i = 1,...,n-1$, the coefficient x_i is equal to $\lambda_i - \lambda_{i-1}$ and the last one of course is x_n. Thus, (3.1) is equivalent to the following

$$X := (\lambda_1 - \lambda_2)X_1 + (\lambda_2 - \lambda_3)X_2 + ... + (\lambda_{n-1} - \lambda_n)X_{n-1} + x_nX_n \tag{3.2}$$

Now suppose that A belongs to T, then there exist nonnegative numbers $\alpha_1,....,\alpha_n$ with $\sum_{i=1}^n \alpha_i = 1$ and such that

$$A := \alpha_1X_1 + \alpha_2X_2 + \alpha_3X_3 + ... + \alpha_{n-1}X_{n-1} + \alpha_nX_n.$$

Clearly, A is a symmetric doubly stochastic by construction. Using the fact that $\sum_{i=1}^n \alpha_i = 1$, then by (3.1), an equivalent way of writing A is the following

$$\begin{aligned} A &= U_n\left(\text{diag}\left(\sum_{i=1}^n \alpha_i,\ \sum_{i=2}^n \alpha_i, \sum_{i=3}^n \alpha_i, ...,\ \sum_{i=n-1}^n \alpha_i, \alpha_n\right)\right)U_n^T \\ &= U_n\text{diag}\left(1,\mu_2,\mu_3,,...,\mu_{n-1},\alpha_n\right)U_n^T, \end{aligned}$$

with $\mu_i = \sum_{k=i}^{n} \alpha_k$. But then, the pth root of A is clearly given by

$$A^{1/p} = U_n \text{diag}\left(1, \mu_2^{1/p}, \mu_3^{1/p}, ..., \mu_{n-1}^{1/p}, \alpha_n^{1/p}\right) U_n^T.$$

Hence, by (3.2), an equivalent way of writing $A^{1/p}$ is the following

$$A^{1/p} = \left(1 - \mu_2^{1/p}\right) X_1 + \left(\mu_2^{1/p} - \mu_3^{1/p}\right) X_2 + ... + \left(\mu_{n-1}^{1/p} - \alpha_n^{1/p}\right) X_{n-1} + \alpha_n^{1/p} X_n.$$

Finally, noticing that all the coefficients in the linear combination of $A^{1/p}$ are nonnegative and they sum up to 1, then we conclude that $A^{1/p}$ is doubly stochastic, and the proof is complete. □

As a result, we have the following corollary.

Corollary 3.3.9. *Let $\Lambda = \{\lambda_1, ..., \lambda_n\} \subset \mathbb{R}$ be such that $1 = \lambda_1 \geqslant \lambda_2 ... \geqslant \lambda_n \geqslant 0 = \lambda_{n+1}$. Then there exists a positive semi-definite doubly stochastic matrix A with spectrum Λ and having a doubly stochastic pth root for every positive integer $p \geq 2$.*

Proof. It suffices to consider $X := x_1 X_1 + ... + x_n X_n \in T$, where $x_i = \lambda_i - \lambda_{i+1}$, for $i = 1, .., n$. □

3.4 Generating elements in $K_n^{1/p}$ using M-matrices

In this section, we shall identify an interesting class of positive definite doubly stochastic matrices with doubly stochastic pth roots by making use of the theory of M-matrices. Recall ([30]) that $A \in \mathbb{R}^{n \times n}$ is said to be an M-matrix if $A = sI - B$, where B is a nonnegative matrix and $s \geq \rho(B)$, the spectral radius of B. Moreover, when $s > \rho(B)$ then A is necessarily nonsingular. It is a standard property that the inverse of an M-matrix is nonnegative[7, Chapter 6]. Furthermore, $A^{1/p}$ is also an M-matrix for all positive integers p by a result of Fiedler and Schneider [30].

First, we begin by recalling the following result from Higham and Lin [41, Theorem 3.6] concerning the existence of a stochastic principal pth root of A.

Theorem 3.4.1. *If the stochastic matrix $A \in \mathbb{R}^{n \times n}$ is the inverse of an M-matrix, then the principal pth root of A denoted by $A^{1/p}$ exists and is stochastic for all positive integers p.*

We can conclude the same for positive semi-definite doubly stochastic matrices as follows.

Theorem 3.4.2. *If the positive semi-definite doubly stochastic matrix $A \in \Delta_n^s$ is the inverse of an M-matrix, then the unique positive semi-definite pth root of A, $A^{1/p}$, is doubly stochastic for all positive integers p.*

Proof. First it is worthy to note that since A is positive semi-definite, the principal p^{th} root of A is the unique positive semidefinite p^{th} root of A (see [60, Chapter 1]). By Theorem 3.4.1, the principal pth root, $A^{1/p}$ exists and is stochastic for all p. Take $M = A^{-1}$. Since M is an M-Matrix, then $M = sI - B$ with B nonnegative and $s > \rho(B)$. Clearly, $M^T = sI - B^T$ with B^T nonnegative and $s > \rho(B^T) = \rho(B)$. Hence, M^T is an M-Matrix. It follows that $M^T = (A^{-1})^T = (A^T)^{-1}$. Hence, $(A^T)^{1/p}$ exists and is stochastic for all p. Now the proof can be achieved by using the fact that $(A^{1/p})^T = (A^T)^{1/p}$. □

Next, we give the following example.

Example 10. Consider the matrix

$$A = \begin{pmatrix} \frac{1}{3} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} \end{pmatrix},$$

for which $A^{-1} = 6I_5 - S$, where S is the 5×5 matrix whose all entries are 1. Clearly, A^{-1} is an M-matrix. Hence, A has a doubly stochastic pth root for all positive integers p.

Next, we use the preceding theorem to present a simple but essential algorithm for constructing elements in $K_n^{1/p}$.

Theorem 3.4.3. *For any $c > 1$ and for any $n \times n$ symmetric doubly stochastic matrix B, let M be the matrix defined by $M := cI_n - (c-1)B$. Then M^{-1} is a positive definite doubly stochastic whose pth root is doubly stochastic, for every positive integer p.*

Proof. Since $c > \rho(B) = 1$, then clearly M is an invertible M-matrix. At this point, we aim to find its inverse M^{-1}. First, we start with some elementary auxiliary material. For all $k > 0$, it holds that

$$\left(I_n - \frac{c-1}{c}B\right)\left(I_n + \frac{c-1}{c}B + \ldots + \left(\frac{c-1}{c}B\right)^k\right) = I_n - \left(\frac{c-1}{c}B\right)^{k+1}.$$

Then, it follows that

$$\lim_{k\longrightarrow\infty}\left(I_n-\frac{c-1}{c}B\right)\left(I_n+\frac{c-1}{c}B+\ldots+\left(\frac{c-1}{c}B\right)^k\right)$$
$$=\lim_{k\longrightarrow\infty}\left(I_n-\left(\frac{c-1}{c}B\right)^{k+1}\right),$$

that is,

$$\left(I_n-\frac{c-1}{c}B\right)\left(\sum_{k=0}^{\infty}\left(\frac{c-1}{c}B\right)^k\right)=I_n,$$

since $\rho\left(\frac{c-1}{c}B\right)<1$. It follows that

$$\left(I_n-\frac{c-1}{c}B\right)^{-1}=\sum_{k=0}^{\infty}\left(\frac{c-1}{c}B\right)^k.$$

Since B is doubly stochastic, we deduce that the series in the right hand side of the preceding equation, has each row and column sum equal to

$$1+\frac{c-1}{c}+\frac{(c-1)^2}{c^2}+\ldots=\sum_{k=0}^{\infty}\left(\frac{c-1}{c}\right)^k=\frac{1}{1-\frac{c-1}{c}}=c.$$

Finally, we note that M can be rewritten as $M=c(I_n-\frac{c-1}{c}B)$, so that $M^{-1}=\frac{1}{c}(I_n-\frac{c-1}{c}B)^{-1}$. Thus M^{-1} is positive definite doubly stochastic matrix. In view of Theorem 3.4.2, M^{-1} has a positive definite doubly stochastic pth, for all positive integers p. □

The following examples illustrates the previous theorem.

Example 11. For any $c>1$, consider the $n\times n$ M-matrix M_c defined by

$$M_c:=cI_n-(c-1)J_n.$$

Its inverse M_c^{-1} is given by $M_c^{-1} = \frac{c-1}{c}J_n + \frac{1}{c}I_n$, as shows the following inspection:

$$\begin{aligned} M_c M_c^{-1} &= (cI_n - (c-1)J_n)\left(\frac{c-1}{c}J_n + \frac{1}{c}I_n\right) \\ &= (c-1)J_n + I_n - \frac{(c-1)^2}{c}J_n - \frac{(c-1)}{c}J_n \\ &= (c-1)J_n + I_n - (c-1)J_n \\ &= I_n. \end{aligned}$$

Thus M_c^{-1} is an element of $K_n^{1/p}$ for every positive integer p. □

Example 12. For any $c > 1$, and for any $n \times n$ symmetric permutation matrix P, consider the $n \times n$ M-matrix M_{cP} defined by

$$M_{cQ} := cI_n - (c-1)P.$$

From the proof of the preceding theorem, we know that

$$\begin{aligned} \left(I_n - \frac{c-1}{c}P\right)^{-1} &= \sum_{k=0}^{\infty}\left(\frac{c-1}{c}\right)^k P^k \\ &= \sum_{k=0,\ k \text{ odd}}^{\infty}\left(\frac{c-1}{c}\right)^k P^k + \sum_{k=0,\ k \text{ even}}^{\infty}\left(\frac{c-1}{c}\right)^k P^k \\ &= \sum_{s=0}^{\infty}\left(\frac{c-1}{c}\right)^{2s+1} P^{2s+1} + \sum_{s=0}^{\infty}\left(\frac{c-1}{c}\right)^{2s} P^{2s} \\ &= \sum_{s=0}^{\infty}\left(\frac{c-1}{c}\right)^{2s+1} P + \sum_{s=0}^{\infty}\left(\frac{c-1}{c}\right)^{2s} I_n \quad (\text{since } P^2 = I_n) \\ &= \sum_{s=0}^{\infty}\left(\frac{c-1}{c}\right)^{2s}\left[\frac{c-1}{c}P + I_n\right] \\ &= \left[\frac{c-1}{c}P + I_n\right]\sum_{s=0}^{\infty}\left(\left(\frac{c-1}{c}\right)^2\right)^s \\ &= \left[\frac{c-1}{c}P + I_n\right]\frac{1}{1-\left(\frac{c-1}{c}\right)^2} \qquad \left(\text{since } \left(\frac{c-1}{c}\right)^2 < 1\right) \\ &= \frac{c^2}{2c-1}\left[\frac{c-1}{c}P + I_n\right]. \end{aligned}$$

As a result, for any $c > 1$ and for any permutation matrix P, it holds that the matrix $M_{cP}^{-1} = (1/c)\frac{c^2}{2c-1}\left[\frac{c-1}{c}P + I_n\right] = \frac{c}{2c-1}\left[\frac{c-1}{c}P + I_n\right]$ is an element of $K_n^{1/p}$ for every positive integer p. □

As a consequence, we have the following.

Corollary 3.4.4. *For any $0 \leq a \leq \frac{1}{2}$, and for any permutation matrix P, it holds that $aP + (1-a)I_n$ is in $K_n^{1/p}$ for every positive integer p.*

Proof. It suffices to exploit the matrix $M_{cP}^{-1} = \frac{c}{2c-1}\left[\frac{c-1}{c}P + I_n\right]$ given in the previous example. Indeed, in this matrix, if we first let c tend to infinity, we obtain the matrix $\frac{1}{2}P + \frac{1}{2}I_n$ while if we let c tend to 1, we get the identity matrix I_n. By continuity, the proof is complete. □

3.5 Constructing elements of $K_n^{1/p}$ via the use of eigenvalues.

In this section, we will construct elements in $K_n^{1/p}$ by making connection with the symmetric doubly stochastic inverse eigenvalue problem (SDIEP) (see Section (1.4)). First, we recall that if A is a symmetric doubly stochastic matrix, then by spectral theorem, there exists an orthogonal matrix V whose first column is e_n and such that $A = V\text{diag}(1, \lambda_2, ..., \lambda_n)V^T$. Therefore, if A is positive semi-definite, then its unique positive semi-definite pth root is given by

$$A^{1/p} = V\text{diag}(1, \lambda_2^{1/p}, ..., \lambda_n^{1/p})V^T.$$

Thus, for a fixed such V, one could explore the relations that the eigenvalues $\{1, \lambda_2, ..., \lambda_n\}$ should satisfy in order for A and $A^{1/p}$ to be doubly stochastic.

In connection with this, recall that a principal method to solve the (SDIEP), relies on taking a real diagonal matrix $\Lambda = \text{diag}(1, \lambda_2, ..., \lambda_n)$, with $-1 \leq \lambda_i \leq 1$ for $i = 2, ..., n$, and an orthogonal matrix V with first column e_n, and then exploring the conditions under which $A = V\Lambda V^T$ is doubly stochastic. As a conclusion, one can see an obvious intersection between (SDIEP) and constructing elements in $K_n^{1/p}$.

Among all such orthogonal matrices V, those of interest to us are the so-called *Soules matrices*. Indeed, Soules [122] used the same procedure just described with a particular matrix V_s, later known as a Soules matrix, and obtained the following theorem (see [122, Theorem 2.5]).

Theorem 3.5.1. *[122] If $1 \geq \lambda_2 \geq ... \geq \lambda_n \geq -1$ and*

$$\frac{1}{n} + \frac{n-m-1}{n(m+1)}\lambda_2 + \sum_{k=1}^{m} \frac{\lambda_{n-2k+2}}{(k+1)k} \geq 0, \tag{3.3}$$

holds with $n = 2m+2$ if n even and $n = 2m+1$ if n odd, then there exists an $n \times n$ symmetric doubly stochastic matrix D such that D has eigenvalues $1, \lambda_2, ..., \lambda_n$.

It follows from the preceding theorem that for any $1 \geq \lambda_2 \geq ... \geq \lambda_n \geq 0$, then (3.3) is valid and hence the symmetric doubly-stochastic inverse eigenvalue problem has solutions i.e. there exists an $n \times n$ symmetric doubly-stochastic matrix X such that $X = V_s \text{diag}(1, \lambda_2, ..., \lambda_n) V_s^T$. But then as $1 \geq \lambda_2^{1/p} \geq ... \geq \lambda_n^{1/p} \geq 0$ for any integer positive p, then $Y = V_s \text{diag}(1, \lambda_2^{1/p}, ..., \lambda_n^{1/p}) V_s^T$ is also a symmetric doubly-stochastic matrix with $Y = X^p$. Thus, any $n \times n$ positive semi-definite doubly stochastic matrix X which can be obtained from the preceding theorem, is an element of $K_n^{1/p}$ for any positive integer p.

In [27], Elsener et al. took a further step and gave the following definition.

Definition 3.5.2. Let $S \in \mathbb{R}^{n \times n}$ be an orthogonal matrix with columns $(r_1, r_2, ..., r_n)$. The set $\{r_1, ..., r_n\}$ is called a *Soules basis* and S is called a *Soules matrix* if the following 2 conditions are satisfied:

- r_1 is positive,
- for every diagonal matrix $\Lambda := \text{diag}(\lambda_1, \lambda_2, ..., \lambda_n)$ with $\lambda_1 \geqslant \lambda_2 \geqslant .... \geqslant \lambda_n \geqslant 0$, the matrix $A_\Lambda = S \Lambda S^T$ is nonnegative.

Hence, this leads naturally to the following theorem.

Theorem 3.5.3. *Let S be an $n \times n$ Soules matrix whose first column is $r_1 = e_n$. If $1 \geqslant \lambda_2 \geqslant ... \geqslant \lambda_n \geqslant 0$, then the matrix given by $A = S \text{diag}(1, \lambda_2, ..., \lambda_n) S^T$ is in $K_n^{1/p}$ for any positive integer p.*

Proof. As $r_1 = e_n$, A is doubly stochastic by [27]. The p^{th} root of A is given by

$$A^{1/p} = S \text{diag}(1, \lambda_2^{1/p}, ..., \lambda_n^{1/p}) S^T.$$

Since S is a Soules matrix, and $1 \geqslant \lambda_2^{1/p} \geqslant ... \geqslant \lambda_n^{1/p} \geqslant 0$, then $A^{1/p}$ is also doubly stochastic. Hence, A has a positive semi-definite doubly stochastic p^{th} root for every integer $p \geq 1$. □

In the same paper [27], the authors also characterised all Soules matrices starting with a fixed positive vector $x \in \mathbb{R}^n$. Thus, for $x := e_n$ one can find all corresponding Soules matrices and hence, in principle this leads to finding all elements in $K_n^{1/p}$ (for any positive integer p) that can be obtained in this fashion. We illustrate this, for the case $n = 3$. First, in order to state their characterisation, we require the following definition which can be found in [26].

Definition 3.5.4. Let $\mathcal{N} = (\mathcal{N}_1, \mathcal{N}_2, ..., \mathcal{N}_n)$ be a sequence of partitions of $< n >= \{1, 2, ..., n\}$ with $\mathcal{N}_i = \{\mathcal{N}_{i,1}, \mathcal{N}_{i,2}, ..., \mathcal{N}_{i,i}\}$, for each $i \in \{1, 2, ..., n\}$.

If for each $i \in \{2, ..., n\}$, there exist indices j, k, l with $1 \leq j \leq i-1$ and $1 \leq k < l \leq i$, such that

$$\mathcal{N}_{i-1}/\mathcal{N}_{i-1,j} = \mathcal{N}_i/\{\mathcal{N}_{i,k}, \mathcal{N}_{i,l}\},$$

and

$$\mathcal{N}_{i-1,j} = \mathcal{N}_{i,k} \cup \mathcal{N}_{i,l},$$

i.e. $\mathcal{N}_i$ is constructed from $\mathcal{N}_{i-1}$ by splitting exactly one of the sets in $\mathcal{N}_{i-1}$ into two subsets, then $\mathcal{N}$ is said to be a Soules-type sequence.

Example 13. For the case $n = 3$, we have three possible Soules-type sequences.

- Case 1: $\mathcal{N} = (\mathcal{N}_1, \mathcal{N}_2, \mathcal{N}_3)$, where

$$\mathcal{N}_1 = \{1, 2, 3\},$$

$$\mathcal{N}_2 = \{\{1, 2\}, \{3\}\},$$

and

$$\mathcal{N}_3 = \{\{1\}, \{2\}, \{3\}\}.$$

- Case 2: $\mathcal{N} = (\mathcal{N}_1, \mathcal{N}_2, \mathcal{N}_3)$, where

$$\mathcal{N}_1 = \{1, 2, 3\},$$

$$\mathcal{N}_2 = \{\{1, 3\}, \{2\}\},$$

and,

$$\mathcal{N}_3 = \{\{1\}, \{3\}, \{2\}\}.$$

- Case 3: $\mathcal{N} = (\mathcal{N}_1, \mathcal{N}_2, \mathcal{N}_3)$, where

$$\mathcal{N}_1 = \{1, 2, 3\},$$

$$\mathcal{N}_2 = \{\{2, 3\}, \{1\}\},$$

and,

$$\mathcal{N}_3 = \{\{2\}, \{3\}, \{1\}\}.$$

We can now expose the characterisation of Soules matrices due to Elsner, Nabben and Neumann [27, Theorem 2.2]. First, we will define for any $x \in \mathbb{R}^n$ and for a any subset $\mathcal{N}$ of $\{1,2,...,n\}$, $x_{\mathcal{N}} = \begin{cases} x_i & i \in \mathcal{N} \\ 0 & \text{elsewhere} \end{cases}$

Theorem 3.5.5. *[27] Let $x \in \mathbb{R}^n$ be a positive vector and let S be a Soules matrix with columns r_1 , r_2 , . . . , r_n , where $r_1 = x$. Then there exists a Soules-type sequence $\mathcal{N}$ of partitions of $\{1,2,...,n\}$ such that r_i is given(up to a factor of $\pm$ 1) by*

$$r_i = \frac{1}{\sqrt{\| x_{\mathcal{N}_{i,t}} \|_2^2 + \| x_{\mathcal{N}_{i,s}} \|_2^2}} \left(\frac{\| x_{\mathcal{N}_{i,t}} \|_2^2}{\| x_{\mathcal{N}_{i,s}} \|_2^2} x_{\mathcal{N}_{i,s}} - \frac{\| x_{\mathcal{N}_{i,s}} \|_2^2}{\| x_{\mathcal{N}_{i,t}} \|_2^2} x_{\mathcal{N}_{i,t}} \right), \tag{3.4}$$

where s and t are those indices in $\{1,2,...,i\}$ for which sets $\mathcal{N}_{i,s}$ and $\mathcal{N}_{i,t}$ do not coincide with one of the sets $\mathcal{N}_{i-1,j}, j = 1,...,i-1$.

Conversely, if $x \in \mathbb{R}^n$ is a positive vector with $\| x \|_2 = 1$ and $\mathcal{N}$ is a Soules-type sequence of partitions of $< n >$, then the matrix $S = [r_1, r_2, ..., r_n]$ with $r_1 = x$ and r_2, r_3,..., r_n given by (3.4) is a Soules matrix.

Example 14. For the case $n = 3$, and for $r_1 = e_3 > 0$, we now find all elements in $K_n^{1/p}$ that can be obtained from the technique just described. First, using the characterisation giving in Theorem 3.5.5, we construct all possible Soules matrices S with columns (r_1, r_2, r_3). Indeed, there are only 3 Soules matrices, up to $\pm r_2$ and $\pm r_3$, giving by

$$S_1 := \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & -1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & 2/\sqrt{6} & 0 \end{pmatrix}, \; S_2 := \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & 2/\sqrt{6} & 0 \\ 1/\sqrt{3} & -1/\sqrt{6} & -1/\sqrt{2} \end{pmatrix},$$

and

$$S_3 := \begin{pmatrix} 1/\sqrt{3} & 2/\sqrt{6} & 0 \\ 1/\sqrt{3} & -1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & -1/\sqrt{6} & -1/\sqrt{2} \end{pmatrix}.$$

As a result, by Theorem 3.5.3, the sets defined by

$$E_i := \left\{ S_i \text{diag}(1, \lambda_2, \lambda_3) S_i^T \text{ with } 1 \geqslant \lambda_2 \geqslant \lambda_3 \geqslant 0 \right\},$$

for $i = 1,2,3$ are subsets of $K_3^{1/p}$ for all integers $p \geqslant 1$.

However, if the eigenvalues $\{1, \lambda_2, ..., \lambda_n\}$ are not taken in the decreasing order, then our knowledge concerning constructing elements of $K_n^{1/p}$ using this procedure, is certainly less and more exploration in this direction is needed.

In the case when considering an orthogonal matrix R which is not a Soules matrix, it is necessary to investigate the conditions under which the matrix $R\text{diag}(1, \lambda_2, ..., \lambda_n)R^T$ has a doubly stochastic pth root and for what value of $p \geq 1$, this may happen. It is worthy to mention here that this can be done on a case by case basis as it appears that there is no systematic way of dealing with a general case. For illustration, we will give the following example.

Example 15. Let $1 \geqslant \lambda_2 \geqslant \lambda_3 \geqslant 0$, and consider the following orthogonal matrix

$$R = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \end{pmatrix}.$$

Using the characterisation of Elsener et. al given in Theorem3.5.5, R is obviously not a Soules matrix, and a simple matrix multiplication shows that

$$\begin{aligned} R\Lambda R^T &= R\text{diag}(1, \lambda_2, \lambda_3)R^T \\ &= \begin{pmatrix} \frac{1}{3} + \frac{1}{2}\lambda_2 + \frac{1}{6}\lambda_3 & \frac{1}{3} - \frac{1}{2}\lambda_2 + \frac{1}{6}\lambda_3 & \frac{1}{3} - \frac{1}{3}\lambda_3 \\ \frac{1}{3} - \frac{1}{2}\lambda_2 + \frac{1}{6}\lambda_3 & \frac{1}{3} + \frac{1}{2}\lambda_2 + \frac{1}{6}\lambda_3 & \frac{1}{3} - \frac{1}{3}\lambda_3 \\ \frac{1}{3} - \frac{1}{3}\lambda_3 & \frac{1}{3} - \frac{1}{3}\lambda_3 & \frac{1}{3} + \frac{2}{3}\lambda_3 \end{pmatrix} \end{aligned}$$

and

$$\begin{aligned} R\Lambda^{1/p}R^T &= R\text{diag}(1, \lambda_2^{\frac{1}{p}}, \lambda_3^{\frac{1}{p}})R^T \\ &= \begin{pmatrix} \frac{1}{3} + \frac{1}{2}\lambda_2{}^{\frac{1}{p}} + \frac{1}{6}\lambda_3{}^{\frac{1}{p}} & \frac{1}{3} - \frac{1}{2}\lambda_2{}^{\frac{1}{p}} + \frac{1}{6}\lambda_3{}^{\frac{1}{p}} & \frac{1}{3} - \frac{1}{3}\lambda_3{}^{\frac{1}{p}} \\ \frac{1}{3} - \frac{1}{2}\lambda_2{}^{\frac{1}{p}} + \frac{1}{6}\lambda_3{}^{\frac{1}{p}} & \frac{1}{3} + \frac{1}{2}\lambda_2{}^{\frac{1}{p}} + \frac{1}{6}\lambda_3{}^{\frac{1}{p}} & \frac{1}{3} - \frac{1}{3}\lambda_3{}^{\frac{1}{p}} \\ \frac{1}{3} - \frac{1}{3}\lambda_3{}^{\frac{1}{p}} & \frac{1}{3} - \frac{1}{3}\lambda_3{}^{\frac{1}{p}} & \frac{1}{3} + \frac{2}{3}\lambda_3{}^{\frac{1}{p}} \end{pmatrix} \end{aligned}$$

At this point, the matrix defined by $R\text{diag}(1,\lambda_2,\lambda_3)R^T$ is doubly stochastic and has a positive semi-definite doubly stochastic pth root if and only if the following system holds

$$\begin{cases} \frac{1}{3}-\frac{1}{2}\lambda_2+\frac{1}{6}\lambda_3 \geq 0 \\ \frac{1}{3}-\frac{1}{2}\lambda_2{}^{\frac{1}{p}}+\frac{1}{6}\lambda_3{}^{\frac{1}{p}} \geq 0 \\ 1 \geqslant \lambda_2 \geqslant \lambda_3 \geqslant 0. \end{cases}$$

It is worth noting here that for the list $(1,\frac{1}{4},0)$, a simple check shows that

$$R\text{diag}\left(1,\left(\frac{1}{4}\right)^{1/p},0\right)R^T \geq 0$$

for $p=1,2,3$. However

$$R\text{diag}\left(1,\left(\frac{1}{4}\right)^{1/p},0\right)R^T$$

has negative entries for all $p \geq 4$. While, for the list $(1,\frac{1}{4},\frac{1}{4})$, an inspection shows that for all $p \geq 1$, the matrix

$$R\text{diag}\left(1,\left(\frac{1}{4}\right)^{1/p},\left(\frac{1}{4}\right)^{1/p}\right)R^T \geq 0.$$

Finally, for the list $(1,\frac{1}{2},0)$, a simple check shows that

$$R\text{diag}\left(1,\left(\frac{1}{2}\right),0\right)R^T \geq 0,$$

however

$$R\text{diag}\left(1,\left(\frac{1}{2}\right)^{1/p},0\right)R^T$$

has negative entries for all $p \geq 2$.

It is worthy to mention here the following observation.

Observation 3.5.6. *If S is any Soules basis, then the process of permuting its rows results in another Soules matrix. However, permuting its columns may not result in a Soules matrix.*

Proof. Suppose that for every diagonal matrix $\Lambda := \text{diag}(\lambda_1,\lambda_2,...,\lambda_n)$ with $\lambda_1 \geqslant \lambda_2 \geqslant .... \geqslant \lambda_n \geqslant 0$, the matrix $A_\Lambda = S\Lambda S^T$ is nonnegative. Therefore, for any permutation matrix Q, it holds that $QS\Lambda S^T Q^T$ is nonnegative. Thus, QS is a Soules matrix. However, SQ may not be a Soules matrix as can be shown in the following example. Indeed, for the case $n=3$ and

with $1 \geq \lambda_2 \geq \lambda_3 \geq 0$, consider the matrix S given by

$$S_1 = \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & -1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & 2/\sqrt{6} & 0 \end{pmatrix},$$

which is a Soules matrix given in Example 14. Now permuting columns 2 and 3 in S_1 amounts to multiplying S_1 to the right by $Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$. Another simple check shows that

$$\begin{aligned} S_1 Q \Lambda Q^T S_1^T &= S_1 Q \mathrm{diag}(1, \lambda_2, \lambda_3) Q^T S_1^T \\ &= \begin{pmatrix} \frac{1}{3} + \frac{1}{2}\lambda_2 + \frac{1}{6}\lambda_3 & \frac{1}{3} - \frac{1}{2}\lambda_2 + \frac{1}{6}\lambda_3 & \frac{1}{3} - \frac{1}{3}\lambda_3 \\ \frac{1}{3} - \frac{1}{2}\lambda_2 + \frac{1}{6}\lambda_3 & \frac{1}{3} + \frac{1}{2}\lambda_2 + \frac{1}{6}\lambda_3 & \frac{1}{3} - \frac{1}{3}\lambda_3 \\ \frac{1}{3} - \frac{1}{3}\lambda_3 & \frac{1}{3} - \frac{1}{3}\lambda_3 & \frac{1}{3} + \frac{2}{3}\lambda_3 \end{pmatrix} \end{aligned}$$

and hence for $\lambda_2 = 1$ and $\lambda_3 = 0$, the matrix $SQ\Lambda Q^T S_1^T$ has negative entries. This proves that $S_1 Q$ is not a Soules matrix. In fact, $S_1 Q = R$, where R is the matrix given in Example 15. □

As a conclusion, we have the following observation concerning Soules matrices for the case $n = 3$.

Remark 10. By example 14 and using the previous observation, it is enough to take into account only S_1 as a Soules matrix since S_2 and S_3 can be obtained from S_1 by permuting some of their rows. Therefore, the sets defined by

$$E_1 := \left\{ S_1 \mathrm{diag}(1, \lambda_2, \lambda_3) S_1^T \text{ with } 1 \geqslant \lambda_2 \geqslant \lambda_3 \geqslant 0 \right\}$$

and

$$E_{1Q} := \left\{ Q E_1 Q^T, \text{ where } Q \text{ runs over all } 3 \times 3 \text{ permutation matrices} \right\}$$

are subsets of $K_3^{1/p}$ for any positive integer p. Thus, the union $E_1 \bigcup E_{1Q}$ is the only subset of $K_3^{1/p}$ that can be obtained using the concept of Soules bases. □

As mentioned before, it appears that there is no systematic way to deal with a general case of an orthogonal matrix O to define the sets of all positive semi-definite doubly stochastic matrices, given by $O\mathrm{diag}(1, \lambda_2, ..., \lambda_n) O^T$ with $1 \geqslant \lambda_2 \geqslant .... \geqslant \lambda_n \geqslant 0$, that belongs to $K_n^{1/p}$ for a positive integer p. Our objective in the next section is to define the set $K_n^{1/p}$ by making

use of general orthogonal matrices for the particular case of $n = 3$ and $p = 2$ then make a generalisation to obtain at least a family of matrices that belongs to $K_n^{1/2}$ for any n.

3.6 A family of $K_n^{1/2}$ via the use of eigenvalues.

In this section, we will make use of the eigenvalues to characterise $K_3^{1/2}$. A generalisation for any n, gives us a family of PSD doubly stochastic matrices having doubly stochastic square root. Firstly, let us denote by

$$a_t = \left(-\frac{1}{2}t^2 - \frac{1}{2}t\sqrt{-3t^2+2}\right),$$

$$b_t = -a_t - \frac{1}{3},$$

$$c_t = \left(-\frac{1}{2}t^2 + \frac{1}{2}t\sqrt{-3t^2+2}\right)$$

and,

$$d_t = -c_t - \frac{1}{3},$$

for any $t \in \left[-\frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{3}\right]$.

3.6.1 A method to define $K_3^{1/2}$

Next, we give orthogonal matrices of order 3 that suffices to diagonalize all the positive semi-definite doubly stochastic matrices. For the purpose, consider the following orthogonal matrices given by

$$O_{1t} = \begin{pmatrix} \frac{1}{\sqrt{3}} & t & \sqrt{2/3-t^2} \\ \frac{1}{\sqrt{3}} & \frac{-t-\sqrt{2-3t^2}}{2} & \frac{-\sqrt{2/3-t^2}-\sqrt{3t^2}}{2} \\ \frac{1}{\sqrt{3}} & \frac{-t+\sqrt{2-3t^2}}{2} & \frac{-\sqrt{2/3-t^2}+\sqrt{3t^2}}{2} \end{pmatrix}$$

where $t \in \left[-\frac{\sqrt{6}}{3}, 0\right]$ and,

$$O_{2t} = \begin{pmatrix} \frac{1}{\sqrt{3}} & t & -\sqrt{2/3-t^2} \\ \frac{1}{\sqrt{3}} & \frac{-t-\sqrt{2-3t^2}}{2} & \frac{\sqrt{2/3-t^2}-\sqrt{3t^2}}{2} \\ \frac{1}{\sqrt{3}} & \frac{-t+\sqrt{2-3t^2}}{2} & \frac{\sqrt{2/3-t^2}+\sqrt{3t^2}}{2} \end{pmatrix},$$

where $t \in \left[0, \frac{\sqrt{6}}{3}\right]$.

Proposition 3.6.1. *O_{1t} with $t \in \left[-\frac{\sqrt{6}}{3}, 0\right]$ and O_{2t} with $t \in \left[0, \frac{\sqrt{6}}{3}\right]$ diagonalize all symmetric doubly stochastic matrices of order 3.*

Proof. Let $U = [e_3 \| v_2 \| v_3]$ be an orthogonal matrix of order 3, with first column $e_3 = \frac{1}{\sqrt{3}}(1,1,1)$. If $v_2 = (a,b,c)^T$, then $a+b+c = 0$ since $e^T.v_2 = 0$, and $a^2+b^2+c^2 = 1$. Without loss of generality, suppose that a is equal to some parameter t, then we have two cases:

1. $b = \frac{-t-\sqrt{2-3t^2}}{2}$, and $c = \frac{-t+\sqrt{2-3t^2}}{2}$ with $t \in \left[-\frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{3}\right]$.

2. $b = \frac{-t+\sqrt{2-3t^2}}{2}$, and $c = \frac{-t-\sqrt{2-3t^2}}{2}$ with $t \in \left[-\frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{3}\right]$.

Case 1: $v_2 = (t, \frac{-t-\sqrt{2-3t^2}}{2}, c = \frac{-t+\sqrt{2-3t^2}}{2})^T$. By a similar manner, we have two possible cases for v_3.

i) If $v_3 = (x, \frac{-x-\sqrt{2-3x^2}}{2}, \frac{-x+\sqrt{2-3x^2}}{2})^T$, with $x \in [-\frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{3}]$. By a simple computation of $v_2^T.v_3 = 0$, we obtain $x = \sqrt{2/3-t^2}$ if $t \in [-\frac{\sqrt{6}}{3}, 0]$ and, $x = -\sqrt{2/3-t^2}$ if $t \in [0, \frac{\sqrt{6}}{3}]$. Hence, in this case

$$U = O_{1t},$$

for $t \in [-\frac{\sqrt{6}}{3}, 0]$ and

$$U = O_{2t},$$

for $t \in [0, \frac{\sqrt{6}}{3}]$.

ii) If $v_3 = (x, \frac{-x+\sqrt{2-3x^2}}{2}, \frac{-x-\sqrt{2-3x^2}}{2})^T$. By a similar computation, we obtain $x = -\sqrt{2/3-t^2}$ if $t \in [-\frac{\sqrt{6}}{3}, 0]$ and, $x = \sqrt{2/3-t^2}$ if $t \in [0, \frac{\sqrt{6}}{3}]$. Therefore,

$$U = W_{1t} = \begin{pmatrix} \frac{1}{\sqrt{3}} & t & -\sqrt{2/3-t^2} \\ \frac{1}{\sqrt{3}} & \frac{-t-\sqrt{2-3t^2}}{2} & \frac{\sqrt{2/3-t^2}+\sqrt{3t^2}}{2} \\ \frac{1}{\sqrt{3}} & \frac{-t+\sqrt{2-3t^2}}{2} & \frac{\sqrt{2/3-t^2}-\sqrt{3t^2}}{2} \end{pmatrix}$$

for $t \in \left[-\frac{\sqrt{6}}{3}, 0\right]$ and,

$$W_{2t} = \begin{pmatrix} \frac{1}{\sqrt{3}} & t & \sqrt{2/3-t^2} \\ \frac{1}{\sqrt{3}} & \frac{-t-\sqrt{2-3t^2}}{2} & \frac{-\sqrt{2/3-t^2}+\sqrt{3t^2}}{2} \\ \frac{1}{\sqrt{3}} & \frac{-t+\sqrt{2-3t^2}}{2} & \frac{-\sqrt{2/3-t^2}-\sqrt{3t^2}}{2} \end{pmatrix},$$

for $t \in \left[0, \frac{\sqrt{6}}{3}\right]$. However, W_{1t} is obtained from O_{1t} by multiplying its third column by -1, and W_{2t} is obtained from O_{2t} by multiplying its third column by -1.Therefore, it suffices in this case to consider O_{1t} and O_{2t}.

Case 2: $v_2 = (t, \frac{-t+\sqrt{2-3t^2}}{2}, c = \frac{-t-\sqrt{2-3t^2}}{2})^T$. By a similar way as described in the first case, It suffices here to consider the two following orthogonal matrices:

$$Z_{1t} = \begin{pmatrix} \frac{1}{\sqrt{3}} & t & \sqrt{2/3-t^2} \\ \frac{1}{\sqrt{3}} & \frac{-t+\sqrt{2-3t^2}}{2} & \frac{-\sqrt{2/3-t^2}+\sqrt{3t^2}}{2} \\ \frac{1}{\sqrt{3}} & \frac{-t-\sqrt{2-3t^2}}{2} & \frac{-\sqrt{2/3-t^2}-\sqrt{3t^2}}{2} \end{pmatrix}$$

where $t \in \left[-\frac{\sqrt{6}}{3}, 0\right]$ and,

$$Z_{2t} = \begin{pmatrix} \frac{1}{\sqrt{3}} & t & -\sqrt{2/3-t^2} \\ \frac{1}{\sqrt{3}} & \frac{-t+\sqrt{2-3t^2}}{2} & \frac{\sqrt{2/3-t^2}+\sqrt{3t^2}}{2} \\ \frac{1}{\sqrt{3}} & \frac{-t-\sqrt{2-3t^2}}{2} & \frac{\sqrt{2/3-t^2}-\sqrt{3t^2}}{2} \end{pmatrix},$$

where $t \in \left[0, \frac{\sqrt{6}}{3}\right]$. Multiply the second column of Z_{1t} by -1 and then take $k = -t$. Therefore $k \in \left[0, \frac{\sqrt{6}}{3}\right]$, and the matrix obtained is W_{2t}. Similarly, multiply the second column of Z_{2t} by -1 and take $k = -t$. Therefore $k \in \left[-\frac{\sqrt{6}}{3}, 0\right]$, and the matrix obtained is W_{1t}. As previously,in this case, it suffices to consider O_{1t} and O_{2t}.

And the proof is complete. □

Next, we give the following remark.

Remark 11. Note that $O_{1(t=0)}$ is obtained from $O_{2(t=0)}$ by multiplying its third column by -1. Hence,

$$O_{1(t=0)}\text{diag}(1,\lambda_2,\lambda_3)O_{1(t=0)}^T = O_{2(t=0)}\text{diag}(1,\lambda_2,\lambda_3)O_{2(t=0)}^T.$$

Note also that $O_{1(t=-\frac{\sqrt{6}}{3})}$ is obtained from $O_{2(t=\frac{\sqrt{6}}{3})}$ by multiplying its second column by -1. Hence,

$$O_{1(t=-\frac{\sqrt{6}}{3})}\text{diag}(1,\lambda_2,\lambda_3)O^T_{1(t=-\frac{\sqrt{6}}{3})} = O_{2(t=\frac{\sqrt{6}}{3})}\text{diag}(1,\lambda_2,\lambda_3)O^T_{2(t=\frac{\sqrt{6}}{3})}.$$

Therefore, given $(1,\lambda_2,\lambda_3)$ with $1 \geq \lambda_2 \geq \lambda_3 \geq 0$, the two family of matrices

$$H_1 = \left\{ O_{1t}\text{diag}(1,\lambda_2,\lambda_3)O^T_{1t},\ t \in \left]-\frac{\sqrt{6}}{3},0\right] \right\}$$

and

$$H_2 := \left\{ O_{2t}\text{diag}(1,\lambda_2,\lambda_3)O^T_{2t},\ t \in \left]0,\frac{\sqrt{6}}{3}\right] \right\}$$

are positive semi-definite doubly quasi-stochastic matrices with spectrum $(1,\lambda_2,\lambda_3)$.

To characterise the doubly stochastic matrices among $H_1 \cup H_2$, we need to investigate when

$$O_{1t}\text{diag}(1,\lambda_2,\lambda_3)O^T_{1t} \geq 0, \qquad t \in \left]-\frac{\sqrt{6}}{3},0\right]$$

and

$$O_{2t}\text{diag}(1,\lambda_2,\lambda_3)O^T_{2t} \geq 0, \qquad t \in \left]0,\frac{\sqrt{6}}{3}\right].$$

For the purpose, we prove the following propositions.

Proposition 3.6.2. *Let* $1 \geq \lambda_2 \geq \lambda_3 \geq 0$ *and* $t \in \left]-\frac{\sqrt{6}}{3},0\right]$. *Then* $O_{1t}diag(1,\lambda_2,\lambda_3)O^T_{1t} \geq 0$ *if and only if*

- *For* $t \in \left]-\frac{\sqrt{6}}{3},-\frac{\sqrt{6}}{6}\right[$,
$$\frac{1}{3} + c_t\lambda_2 + d_t\lambda_3 \geqslant 0.$$
- *For* $t \in \left[-\frac{\sqrt{6}}{6},0\right]$,
$$\frac{1}{3} + (t^2 - \frac{1}{2})\lambda_2 + (-t^2 + \frac{1}{6})\lambda_3 \geq 0.$$

Proof. Let $A = (a_{ij}) = O_{1t}\text{diag}(1,\lambda_2,\lambda_3)O^T_{1t}$ for some $t \in \left[-\frac{\sqrt{6}}{3},0\right]$. A is positive semi definite matrix, then $a_{11} = a_{22} = a_{33} \geq 0$. It suffices to study the sign of a_{12}, a_{13} and a_{23} given by

$$a_{12} = \frac{1}{3} + a_t\lambda_2 + b_t\lambda_3,$$

$$a_{13} = \frac{1}{3} + c_t\lambda_2 + d_t\lambda_3,$$

and,

$$a_{23} = \frac{1}{3} + \lambda_2(t^2 - \frac{1}{2}) + \lambda_3(-t^2 + \frac{1}{6}).$$

A simple check shows that,

- For $t \in \left]-\frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6}\right[$, we have $a_{13} \leqslant a_{12}$ and $a_{13} < a_{23}$.
- For $t = -\frac{\sqrt{6}}{6}$, $a_{13} = a_{23} \leqslant a_{12}$.
- For $t \in \left]-\frac{\sqrt{6}}{6}, 0\right]$, $a_{23} < a_{13} \leqslant a_{12}$.

Therefore, $A \geq 0$ if and only if $a_{13} \geq 0$, for $t \in \left[-\frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6}\right[$, and if and only if $a_{23} \geq 0$, for $t \in \left[-\frac{\sqrt{6}}{6}, 0\right]$.

□

Proposition 3.6.3. *Let* $1 \geq \lambda_2 \geq \lambda_3 \geq 0$ *and* $t \in \left]0, \frac{\sqrt{6}}{3}\right]$, *and* $1 \geq \lambda_2 \geq \lambda_3 \geq 0$. *Then* $O_{2t}diag(1,\lambda_2,\lambda_3)O_{2t}^T \geq 0$ *if and only if*

- *For* $t \in \left]\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}\right]$,
$$\frac{1}{3} + a_t\lambda_2 + b_t\lambda_3 \geq 0.$$
- *For* $t \in \left]0, \frac{\sqrt{6}}{6}\right]$,
$$\frac{1}{3} + \lambda_2(t^2 - \frac{1}{2}) + \lambda_3(-t^2 + \frac{1}{6}) \geq 0.$$

Proof. Let $B = (b_{ij}) = O_{2t}.\text{diag}(1,\lambda_2,\lambda_3).O_{2t}^T$ for some $t \in]0, \frac{\sqrt{6}}{3}]$. B is positive semi definite matrix, then $b_{11} = b_{22} = b_{33} \geq 0$. It suffices to study the sign of b_{12}, b_{13} and b_{23} given by

$$b_{12} = \frac{1}{3} + a_t\lambda_2 + b_t\lambda_3,$$

$$b_{13} = \frac{1}{3} + c_t\lambda_2 + d_t\lambda_3,$$

and,

$$b_{23} = \frac{1}{3} + \lambda_2(t^2 - \frac{1}{2}) + \lambda_3(-t^2 + \frac{1}{6}).$$

By comparing the entries of B, we obtain the following.

- For $t \in \left]-\frac{\sqrt{6}}{6} - \frac{\sqrt{6}}{3}\right]$, we have $b_{12} \leqslant b_{13}$ and $b_{12} < b_{23}$.

- For $t = \frac{\sqrt{6}}{6}$, $b_{12} = b_{23} \leqslant b_{13}$.
- For $t \in \left]0, \frac{\sqrt{6}}{6}\right]$, $b_{23} < b_{12} \leqslant b_{13}$.

Therefore, $B \geq 0$ if and only if $b_{12} \geq 0$, for $t \in \left]\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}\right]$, and $b_{23} \geq 0$, for $t \in \left]0, \frac{\sqrt{6}}{6}\right]$. And the proof is complete. □

As a conclusion we have the following theorem.

Theorem 3.6.4. *Let* $1 \geqslant \lambda_2 \geqslant \lambda_3 \geqslant 0$. *Define* E_{1t}, E_{2t}, E_{3t} *and* E_{4t} *as follows.*

1. $E_{1t} = \left\{ O_{1t} diag(1,\lambda_2,\lambda_3) O_{1t}^T,\ t \in \left]-\frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6}\right[\Big/ \frac{1}{3} + c_t\sqrt{\lambda_2} + d_t\sqrt{\lambda_3} \geq 0 \right\}$,
2. $E_{2t} = \left\{ O_{1t} diag(1,\lambda_2,\lambda_3) O_{1t}^T,\ t \in \left[-\frac{\sqrt{6}}{6}, 0\right] \Big/ \frac{1}{3} + (t^2 - \frac{1}{2})\sqrt{\lambda_2} + (-t^2 + \frac{1}{6})\sqrt{\lambda_3} \geq 0 \right\}$,
3. $E_{3t} = \left\{ O_{2t} diag(1,\lambda_2,\lambda_3) O_{2t}^T,\ t \in \left]\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}\right] \Big/ \frac{1}{3} + a_t\sqrt{\lambda_2} + b_t\sqrt{\lambda_3} \geq 0 \right\}$,

   and,
4. $E_{4t} = \left\{ O_{2t} diag(1,\lambda_2,\lambda_3) O_{2t}^T,\ t \in \left]0, \frac{\sqrt{6}}{6}\right] \Big/ \frac{1}{3} + (t^2 - \frac{1}{2})\sqrt{\lambda_2} + (-t^2 + \frac{1}{6})\sqrt{\lambda_3} \geq 0 \right\}$.

Then $K_3^{1/2} = E_{1t} \cup E_{2t} \cup E_{3t} \cup E_{4t}$

Proof. Let $M \in E_{1t}$. Then there exists $t \in \left[-\frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6}\right[$ such that $M = O_{1t}\text{diag}(1,\lambda_2,\lambda_3)O_{1t}^T$ and $\frac{1}{3} + c_t\sqrt{\lambda_2} + d_t\sqrt{\lambda_3} \geq 0$. Since M is positive semi-definite, then there exists a positive semi-definite matrix N such that $M = N^2$. Therefore, M is given by

$$N = (n_{ij}) = O_{1t}\text{diag}(1, \sqrt{\lambda_2}, \sqrt{\lambda_3})O_{1t}^T.$$

By construction, N is doubly quasi-stochastic. Since, $\frac{1}{3} + c_t\sqrt{\lambda_2} + d_t\sqrt{\lambda_3} \geq 0$ then by Proposition 3.6.2, N is nonnegative; Therefore N is doubly stochastic. Hence, M is positive semi-definite doubly stochastic matrix with doubly stochastic square root. Similar arguments can be employed for $M \in E_{2t}$, E_{3t}, and E_{4t}.

On the other hand, consider a positive semi-definite doubly stochastic matrix with eigenvalues $\lambda_1, \lambda_2, \lambda_3$. Then there exists $t \in \left[-\frac{\sqrt{6}}{3}, 0\right]$ such that $M = O_{1t}\text{diag}(1,\lambda_2,\lambda_3)O_{1t}^T$, or $t \in \left[0, \frac{\sqrt{6}}{3}\right]$ such that $M = O_{2t}\text{diag}(1,\lambda_2,\lambda_3)O_{2t}^T$. Without loss of generality, suppose that $M = O_{1t}\text{diag}(1,\lambda_2,\lambda_3)O_{1t}^T$ for some $t \in \left[-\frac{\sqrt{6}}{3}, 0\right]$. Then

$$M^{1/2} = O_{1t}\text{diag}(1, \sqrt{\lambda_2}, \sqrt{\lambda_3})O_{1t}^T.$$

By Proposition 3.6.2 , $M^{1/2} \geq 0$ if and only if

- $\frac{1}{3} + c_t\sqrt{\lambda_2} + d_t\sqrt{\lambda_3} \geqslant 0$,if $t \in \left[-\frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6}\right[$
- $\frac{1}{3} + (t^2 - \frac{1}{2})\sqrt{\lambda_2} + (-t^2 + \frac{1}{6})\sqrt{\lambda_3} \geq 0$, if $t \in \left[-\frac{\sqrt{6}}{6}, 0\right]$.

Hence, $M \in E_{1t}$ if $t \in \left[-\frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6}\right[$, or $M \in E_{2t}$ if $t \in \left[-\frac{\sqrt{6}}{6}, 0\right]$. Similar argument can be employed for $M = O_{2t}\text{diag}(1, \lambda_2, \lambda_3)O_{2t}^T$ for some $t \in \left[0, \frac{\sqrt{6}}{3}\right]$. Using Proposition 3.6.3, we conclude that $M \in E_{3t}$ for $t \in \left]\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}\right]$, or $M \in E_{4t}$ for $t \in \left[0, \frac{\sqrt{6}}{6}\right]$. And the proof is complete. □

Example 16. The positive semi-definite doubly stochastic matrices with spectrum $(1,1,0)$ and doubly stochastic square root are given by

$$E_{1t} = \left\{ O_{1t}\text{diag}(1,1,0)O_{1t}^T,\, t \in \left]-\frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6}\right[\Big/ \frac{1}{3} + c_t \geq 0 \right\} = \emptyset$$

$$E_{2t} = \left\{ O_{1t}\text{diag}(1,1,0)O_{1t}^T,\, t \in \left[-\frac{\sqrt{6}}{6}, 0\right] \Big/ t^2 - \frac{1}{2} + \frac{1}{3} \geq 0 \right\} = \left\{ O_{1t}\text{diag}(1,1,0)O_{1t}^T, t = -\frac{\sqrt{6}}{6} \right\},$$

$$E_{3t} = \left\{ O_{2t}\text{diag}(1,1,0)O_{2t}^T,\, t \in \left]\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}\right] \Big/ \frac{1}{3} + a_t \geq 0 \right\} = \left\{ O_{2t}\text{diag}(1,1,0)O_{2t}^T, t = \frac{\sqrt{6}}{3} \right\},$$

and,

$$E_{4t} = \left\{ O_{2t}\text{diag}(1,1,0)O_{2t}^T,\, t \in \left]0, \frac{\sqrt{6}}{6}\right] \Big/ t^2 - \frac{1}{2} + \frac{1}{3} \geq 0 \right\} = \left\{ O_{2t}\text{diag}(1,1,0)O_{2t}^T, t = \frac{\sqrt{6}}{6} \right\}.$$

Therefore, we have the three following elements in $K_3^{1/2}$ with spectrum $(1,1,0)$:

$$E_{2(-\frac{1}{6}\sqrt{6})} = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, E_{3(\frac{1}{3}\sqrt{6})} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix}, \text{and} \quad E_{4(\frac{1}{6}\sqrt{6})} = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{pmatrix}.$$

3.6.2 Generalization to order n

In this section, we give a family of positive semi-definite doubly stochastic matrices that belongs to $K_n^{1/p}$ for any n. For this purpose, we consider the particular orthogonal matrix,

$$V_n = \begin{pmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{(n-1)(n-2)}} & \cdots & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{(n-1)(n-2)}} & \cdots & \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{(n-1)(n-2)}} & \cdots & -\frac{2}{\sqrt{6}} & 0 \\ \vdots & \vdots & \vdots & & \vdots & \\ \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{(n-1)(n-2)}} & \cdots & 0 & 0 \\ \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n(n-1)}} & \frac{-(n-2)}{\sqrt{(n-1)(n-2)}} & \cdots & 0 & 0 \\ \frac{1}{\sqrt{n}} & \frac{-(n-1)}{\sqrt{n(n-1)}} & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

which is a Soules matrix of order n [122]. Therefore for every diagonal matrix $\Lambda = \text{diag}(1, \lambda_2, ..., \lambda_n)$ where $1 \geqslant \lambda_2 \geqslant ... \geqslant \lambda_n \geqslant 0$, $A = V_n \Lambda V_n^T \in K_n^{1/2}$. Using $O_{1t} = (e_3|x_2|x_3)$ and $O_{2t} = (e_3|y_2|y_3)$, we construct two families of orthogonal matrices V_{1t} and V_{2t} of order n. For $n > 3$, we construct V_{1t} by considering the first $n-2$ columns of the Soules matrix V_n and the last 2 columns $\begin{pmatrix} x_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ and $\begin{pmatrix} x_3 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ to be the n columns of V_{1t}. Similarly, we construct V_{2t} by considering the first $n-2$ columns of the Soules matrix V_n and the last 2 columns $\begin{pmatrix} y_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ and $\begin{pmatrix} y_3 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ to be the n columns of V_{2t}. A simple matrix multiplication shows that the matrices $A = V_n \Lambda V_n^T$ and $B = V_{1t} \Lambda V_{1t}^T$, with $t \in \left]-\frac{\sqrt{6}}{3}, 0\right]$ differ only by the 3×3 principal submatrix formed from the first 3 columns and the first 3 rows of A and B. Similarly, the matrices A and $C = V_{2t} \Lambda V_{2t}^T$ with $t \in \left]0, \frac{\sqrt{6}}{3}\right]$ differ also only by the 3×3 principal submatrix formed from the first 3 columns and the first 3 rows of A and C.

Hence, the positive semi-definite matrix B (respectively C) has nonnegative off diagonal entries except b_{12}, b_{13} and b_{23} (respectively c_{12}, c_{13} and c_{23}) given by:

$$b_{12} = c_{12} = \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{(4)(3)}\lambda_{n-2} + a_t\lambda_{n-1} + b_t\lambda_n,$$

$$b_{13} = c_{13} = \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{(4)(3)}\lambda_{n-2} + c_t\lambda_{n-1} + d_t\lambda_n,$$

and,

$$b_{23} = c_{23} = \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{(4)(3)}\lambda_{n-2} + (t^2 - \frac{1}{2})\lambda_{n-1} + (-t^2 + \frac{1}{6})\lambda_n.$$

In the following propositions, we investigate when the matrices given by $V_{1t}\Lambda V_{1t}^T$ and $V_{2t}\Lambda V_{2t}^T$ are nonnegative.

Proposition 3.6.5. *Let $n > 3$ be any positive integer and let $1 \geq \lambda_2 \geq .... \geq \lambda_n \geq 0$. Let $t \in \left]-\frac{\sqrt{6}}{3}, 0\right]$. Then $V_{1t}\Lambda V_{1t}^T \geq 0$ if and only if*

$$\frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{(4)(3)}\lambda_{n-2} + c_t\lambda_{n-1} + d_t\lambda_n \geq 0,$$

for $t \in \left[-\frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6}\right[$, and,

$$\frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{(4)(3)}\lambda_{n-2} + (t^2 - \frac{1}{2})\lambda_{n-1} + (-t^2 + \frac{1}{6})\lambda_n \geq 0,$$

for $t \in \left[-\frac{\sqrt{6}}{6}, 0\right]$.

Proof. By a simple check, it suffices to notice the following inequalities.

- For $t \in \left[-\frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6}\right[$, $b_{13} \leqslant b_{12}$and $b_{13} < b_{23}$.
- For $t = -\frac{\sqrt{6}}{6}$, $b_{13} = b_{23} \leqslant b_{12}$.
- For $t \in \left]-\frac{\sqrt{6}}{6}, 0\right]$, $b_{23} < b_{13} \leqslant b_{12}$.

As a result, $B \geq 0$ if and only if $b_{13} \geq 0$, for $t \in \left]-\frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6}\right[$, and if and only if $b_{23} \geq 0$, for $t \in \left[-\frac{\sqrt{6}}{6}, 0\right]$. This completes the proof.

□

Proposition 3.6.6. *Let $n > 3$ be any positive integer and $1 \geq \lambda_2 \geq .... \geq \lambda_n \geq 0$. Let $t \in \left]0, \frac{\sqrt{6}}{3}\right]$. Then, $V_{2t}\Lambda V_{2t}^T \geq 0$ if and only if*

$$\frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{(4)(3)}\lambda_{n-2} + a_t\lambda_{n-1} + b_t\lambda_n \geq 0,$$

for $t \in \left]\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}\right]$, *and,*

$$\frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{(4)(3)}\lambda_{n-2} + (t^2 - \frac{1}{2})\lambda_{n-1} + (-t^2 + \frac{1}{6})\lambda_n \geq 0,$$

for $t \in \left]0, \frac{\sqrt{6}}{6}\right]$.

Proof. A simple check shows that:

- For $t \in \left]\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}\right]$, $c_{12} \leqslant c_{13}$ and $c_{12} < c_{23}$.
- For $t = \frac{\sqrt{6}}{6}$, $c_{12} = c_{23} \leqslant c_{13}$.
- For $t \in \left]0, \frac{\sqrt{6}}{6}\right[$, $c_{23} < c_{12} \leqslant c_{13}$.

As a consequence, $C \geq 0$ if and only if $c_{12} \geq 0$, for $t \in \left]\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}\right]$, and if and only if $c_{23} \geq 0$, for $t \in \left]0, \frac{\sqrt{6}}{6}\right]$. This completes the proof.

□

Using the Propositions 3.6.5 and 3.6.6 and, we can define a family of positive semidefinite doubly stochastic matrices with spectrum $(1, \lambda_2, ..., \lambda_n)$ where $1 \geqslant \lambda_2 \geqslant ... \geqslant \lambda_n \geqslant 0$, having doubly stochastic square root.

Theorem 3.6.7. *Let n be any positive integer and let* $1 = \lambda_1 \geqslant \lambda_2 \geqslant ... \geqslant \lambda_n \geqslant 0$. *Define* F_{1t}, F_{2t}, F_{3t} *and* F_{4t} *as follows.*

$$-F_{1t} = \left\{ V_{1t} diag(1, \lambda_2, ..., \lambda_n) V_{1t}^T, t \in \left]-\frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6}\right[\text{ such that } \right.$$
$$\left. \frac{1}{n} + \frac{1}{n(n-1)}\sqrt{\lambda_2} + ... + \frac{1}{(4)(3)}\sqrt{\lambda_{n-2}} + c_t\sqrt{\lambda_{n-1}} + d_t\sqrt{\lambda_n} \geqslant 0 \right\},$$

$$-F_{2t} = \left\{ V_{1t} diag(1, \lambda_2, ..., \lambda_n) V_{1t}^T, t \in \left[-\frac{\sqrt{6}}{6}, 0\right] \text{ such that } \right.$$
$$\left. \frac{1}{n} + \frac{1}{n(n-1)}\sqrt{\lambda_2} + ... + \frac{1}{(4)(3)}\sqrt{\lambda_{n-2}} + (t^2 - \frac{1}{2})\sqrt{\lambda_{n-1}} + (-t^2 + \frac{1}{6})\sqrt{\lambda_n} \geq 0 \right\},$$

$$-F_{3t} = \left\{ V_{2t} diag(1, \lambda_2, ..., \lambda_n) V_{2t}^T, t \in \left] \frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3} \right] \text{ such that } \right.$$
$$\left. \frac{1}{n} + \frac{1}{n(n-1)} \sqrt{\lambda_2} + ... + \frac{1}{(4)(3)} \sqrt{\lambda_{n-2}} + a_t \sqrt{\lambda_{n-1}} + b_t \sqrt{\lambda_n} \geqslant 0 \right\},$$

and,

$$-F_{4t} = \left\{ V_{2t} diag(1, \lambda_2, .., \lambda_n) V_{2t}^T, t \in \left] 0, \frac{\sqrt{6}}{6} \right] \text{ such that } \right.$$
$$\left. \frac{1}{n} + \frac{1}{n(n-1)} \sqrt{\lambda_2} + ... + \frac{1}{(4)(3)} \sqrt{\lambda_{n-2}} + (t^2 - \frac{1}{2}) \sqrt{\lambda_{n-1}} + (-t^2 + \frac{1}{6}) \sqrt{\lambda_n} \geq 0 \right\}.$$

Then F_{1t}, F_{2t}, F_{3t} or F_{4t} are subsets of $K_n^{1/2}$.

Proof. Let M be a matrix in F_{1t}. Then there exists $t \in \left] -\frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6} \right[$ such that

$$M = V_{1t} \text{diag}(1, \lambda_2, ..., \lambda_n) V_{1t}^T.$$

M is by construction positive semi-definite doubly quasi- stochastic. Then there exists a positive semi-definite doubly quasi-stochastic matrix N such that $M = N^2$. N is then given by

$$N = V_{1t} \text{diag}(1, \sqrt{\lambda_2}, ..., \sqrt{\lambda_n}) V_{1t}^T.$$

Since

$$\frac{1}{n} + \frac{1}{n(n-1)} \sqrt{\lambda_2} + ... + \frac{1}{(4)(3)} \sqrt{\lambda_{n-2}} + c_t \sqrt{\lambda_{n-1}} + d_t \sqrt{\lambda_n} \geqslant 0,$$

Then by Proposition 3.6.5, N is nonnegative. Therefore, M is positive semi-definite doubly stochastic matrix with spectrum $(1, \lambda_2, ..., \lambda_n)$ and having doubly stochastic square root. Similar arguments can be applied if M belongs to F_{2t}, F_{3t} and F_{4t}. □

Chapter 4

The symmetric doubly stochastic inverse eigenvalue problem

4.1 Introduction

We saw in Chapter 3 that there is an obvious intersection between constructing elements in $K_n^{1/p}$ and SDIEP. Let $\sigma = \{1, \lambda_2, ..., \lambda_n\}$ be a list of n real numbers. Recall that, if there exists a symmetric doubly stochastic matrix A with spectrum σ, then we say σ is symmetrically realisable and that A realises σ . The problem of characterising all symmetrically realisable lists is referred to as the "Symmetric Doubly Stochastic Inverse Eigenvalue Problem", or SDIEP. Equivalently, this problem can also be characterised as the problem of finding the region Θ_n^s of $\mathbb{R}^n$such that any point in Θ_n^s is the spectrum of an $n \times n$ symmetric doubly stochastic matrix.

Since doubly stochastic matrices are nonnegative, then obvious necessary conditions for SDIEP are those concerning the nonnegative inverse eigenvalue problem (see Section 1.4):

1. $0 \leq |\lambda_i| \leq 1$ as the Perron-Frobenius 'theorem insures.
2. $s_k(\sigma) = 1 + \lambda_2^k + ... + \lambda_n^k \geq 0, \forall k.$
3. $s_k(\sigma)^m \leq n^{m-1} s_{km}(\sigma)$ for all positive integers k and m, by JLL conditions.

Until now, the SDIEP has only been solved for the case $n = 3$ by Perfect and Mirsky [97] and remains open for the cases $n \geq 4$. Some partial results concerning SDIEP can be found in the literature (see [53, 58, 79, 81, 82, 85, 97, 99, 103, 122, 140] for a collection of most sufficient conditions for the SDIEP and [84] for some necessary conditions for SDIEP). So far, there are two principal methods to solve the SDIEP. The first one consist on taking a diagonal matrix $\Lambda = \text{diag}(1, \lambda_2, ..., \lambda_n)$ and an invertible matrix P, and then exploring the

conditions under which $A = P^{-1}\Lambda P$ is doubly stochastic. However, when A is symmetric, then there exists an orthogonal matrix P_0 with first column e_n such that $A = P_0\Lambda P_0^T$ (see [80, 113]). Therefore one could establish an orthogonal matrix P_0 and explore the conditions under which $P_0\Lambda P_0^T$ is symmetric doubly stochastic. Indeed, in the spirit of finding elements in $K_n^{1/p}$, we used this method in Section 5 and 6 of Chapter 3. The second method falls into the category of constructing new doubly stochastic matrices from smaller size matrices with known spectra. Our work here falls under the second category. Given some earlier results concerning spectral properties of doubly stochastic matrices [29, 82, 83], we will show how to use this results as a tool for deriving recursively new sufficient conditions for the symmetric doubly stochastic inverse eigenvalue problem.

This chapter is organised as follows. In Section 4.2, we collect some preliminary results concerning eigenvalues of doubly stochastic matrices that we will use for our purpose. Moreover, we will recall some earlier results for SDIEP. In section 4.3, we will use the recursive technique for constructing symmetric doubly stochastic matrices to elaborate two new sufficient conditions for the inverse eigenvalue problem of doubly stochastic matrices. We give also a counterexample to Theorem 9 in [1]. In fact, the authors in [1] used similar techniques as in [82, 83] to obtain a recursive method for constructing doubly stochastic matrices. It is asserted of finding a new sufficient condition that improve the Soules condition introduced in [122]. We will show by a counterexample that this refinement is incorrect where n is odd and we present an alternative statement though in this case it will not be an improvement of the results in [122] but rather a new independent sufficient condition.

In the sequel, we shall make the convention that a summation over the empty index set is defined to be zero.

4.2 Preliminairies and overview of some results

The results in this chapter fall into the category of constructing new doubly stochastic matrices from doubly stochastic matrices with known spectra. For this purpose, we present in this section auxiliary results concerning spectral properties of certain block doubly stochastic matrices that we will use to construct the recursive sufficient conditions for the SDIEP. We start with a frequently used lemma appears in Fiedler [29, Lemma 2.2], which has been repeated in many situations in the study of the nonnegative inverse eigenvalue problem.

Lemma 4.2.1. *[29] Let A be an $n \times n$ symmetric matrix with eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ and let u be the unit eigenvector corresponding to λ_1. Let B be an $m \times m$ symmetric matrix with eigenvalues $\mu_1, \mu_2, ..., \mu_m$ and le w be the unit eigenvector corresponding to μ_1 . Then for*

any ρ, the matrix C defined by

$$C = \begin{pmatrix} A & \rho uw^T \\ \rho wu^T & B \end{pmatrix},$$

has eigenvalues $\lambda_2, ..., \lambda_n, \mu_2, ..., \mu_m$ and γ_1, γ_2 where γ_1, γ_2 are the eigenvalues of the matrix $\begin{pmatrix} \lambda_1 & \rho \\ \rho & \mu_1 \end{pmatrix}$.

Then in [16], the authors generalise this Lemma to k diagonal blocks (with $k \geq 2$) instead of two with corresponding k symmetric matrices. A generalisation for all square matrices and not just the symmetric ones can be found in [83, Theorem 2.3]. Its proof relies on a result which is presented in Perfect [96] and is due to R. Rado. As a specialisation applied to doubly stochastic matrices, we have the following theorem from [82, 83].

Theorem 4.2.2. *Let A be an $n \times n$ doubly stochastic matrix whose eigenvalues are given by $1, \lambda_2, ..., \lambda_n$ and let B be an $m \times m$ doubly stochastic matrix with eigenvalues $1, \mu_2, ..., \mu_m$. Then for any $\rho \geq 0$ and for any $\alpha \geq 0$ such that ρ and α do not vanish simultaneously, the $(m+n) \times (m+n)$ matrix C defined by*

i) For $m \geq n$,

$$C = \frac{1}{\alpha + \frac{\rho m}{\sqrt{mn}}} \begin{pmatrix} \alpha A & \rho e_n e_m^T \\ \rho e_m e_n^T & (\alpha + \rho \frac{m-n}{\sqrt{mn}})B \end{pmatrix}$$

is doubly stochastic with eigenvalues

$$1, \frac{\alpha\sqrt{mn} - \rho n}{\alpha\sqrt{mn} + \rho m}, \frac{\alpha}{\alpha + \frac{\rho m}{\sqrt{mn}}}\lambda_2, ..., \frac{\alpha}{\alpha + \frac{\rho m}{\sqrt{mn}}}\lambda_n, \frac{\alpha\sqrt{mn} + \rho(m-n)}{\alpha\sqrt{mn} + \rho m}\mu_2, ..., \frac{\alpha\sqrt{mn} + \rho(m-n)}{\alpha\sqrt{mn} + \rho m}\mu_m.$$

ii) For $n \geq m$,

$$C = \frac{1}{\alpha + \frac{\rho n}{\sqrt{mn}}} \begin{pmatrix} (\alpha + \rho \frac{n-m}{\sqrt{mn}})A & \rho e_n e_m^T \\ \rho e_m e_n^T & \alpha B \end{pmatrix}$$

is doubly stochastic with eigenvalues

$$1, \frac{\alpha\sqrt{mn} - \rho m}{\alpha\sqrt{mn} + \rho n}, \frac{\alpha\sqrt{mn} + \rho(n-m)}{\alpha\sqrt{mn} + \rho n}\lambda_2, ..., \frac{\alpha\sqrt{mn} + \rho(n-m)}{\alpha\sqrt{mn} + \rho n}\lambda_n, \frac{\alpha}{\alpha + \frac{\rho n}{\sqrt{mn}}}\mu_2, ..., \frac{\alpha}{\alpha + \frac{\rho n}{\sqrt{mn}}}\mu_m.$$

Remark 12. An alternative proof can be achieved from Fiedler [29]. Indeed, the proof of Lemma 4.2.1 , is essentially the same as the proof of Theorem 4.2.2. One simply replaces the symmetry assumption in Fiedler's paper by the property that the algebraic and geometric multiplicities of the eigenvalue 1 of any doubly stochastic matrix are the same.

We note that in [82], the author shows how the previous theorem can be a tool for extracting new sufficient conditions for the inverse eigenvalue problem of doubly stochastic matrices (DIEP, RDIEP and SDIEP).

By a similar manner, we have the following two lemmas that can be found in [1] and which constitute the basis for their results.

Lemma 4.2.3. *Let A be an $n \times n$ doubly stochastic matrix with eigenvalues $1, \lambda_2, ..., \lambda_n$. Then, for any $0 \leq r \leq 1$, there exists an $(n+1) \times (n+1)$ doubly stochastic matrix C with eigenvalues given by $1, 1 - \frac{n+1}{n}r, \frac{n-r}{n}\lambda_2, ..., \frac{n-r}{n}\lambda_n$.*

Proof. It suffices to check that the matrix $C = \begin{pmatrix} 1-r & \frac{r}{\sqrt{n}}e_n^T \\ \frac{r}{\sqrt{n}}e_n & (\frac{n-r}{n})A \end{pmatrix}$ is doubly stochastic with eigenvalues $1, 1 - \frac{n+1}{n}r, \frac{n-r}{n}\lambda_2, ..., \frac{n-r}{n}\lambda_n$. □

Lemma 4.2.4. *Let A be an $n \times n$ doubly stochastic matrix with eigenvalues $1, \lambda_2, ..., \lambda_n$. In addition, let B be an $m \times m$ doubly stochastic matrix with eigenvalues $1, \mu_2, ..., \mu_m$. Then for any $0 \leq r \leq 1$ and $n \geq m$, there exists an $(n+m) \times (n+m)$ doubly stochastic matrix C with eigenvalues given by $1, \alpha + r - 1, \alpha\lambda_2, ..., \alpha\lambda_n, r\mu_2, ..., r\mu_m$ where $\frac{1-\alpha}{m} = \frac{1-r}{n}$.*

Proof. Take in Theorem 4.2.2 Part i), $\rho = (1-\alpha)\frac{\sqrt{n}}{\sqrt{m}}$, then $\rho \geq 0$ for any $1 \geq \alpha \geq 0$. Then the matrix $C = \begin{pmatrix} \alpha A & \frac{(1-r)\sqrt{m}}{\sqrt{n}}e_n e_m^T \\ \frac{(1-r)\sqrt{m}}{\sqrt{n}}e_m e_n^T & rB \end{pmatrix}$ is doubly stochastic with eigenvalues $1, \alpha + r - 1, \alpha\lambda_2, ..., \alpha\lambda_n, r\mu_2, ..., r\mu_m$. □

Next, we give some earlier results in the theory of symmetric doubly stochastic inverse eigenvalue problem. We recall the first known sufficient condition for the SDIEP, which is found in [97] and known as Perfect and Mirsky's condition.

Theorem 4.2.5. *[97]If $1 \geq \lambda_2 \geq .... \geq \lambda_n \geq -1$ and*

$$\frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + \frac{1}{(n-1)(n-2)}\lambda_3 + ... + \frac{1}{2.1}\lambda_n \geq 0 \tag{4.1}$$

then there exists a symmetric doubly stochastic matrix D such that D has eigenvalues $1, \lambda_2, ..., \lambda_n$.

Perfect and Mirsky [97] used the constructing idempotent system method to obtain the sufficient condition (4.1). They proved also that for $n \leq 3$, the condition (4.1) is also necessary (see also [80] for a geometric solution to this problem). The inverse eigenvalue problem for 4×4 symmetric doubly stochastic matrices of trace zero is completely solved in [97].

Theorem 4.2.6. *[97] Let* $1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$. *Then* $(1, \lambda_2, \lambda_3, \lambda_4)$ *is the spectrum of* 4×4 *symmetric doubly stochastic matrix of trace 0 if and only if* $1 + \lambda_2 + \lambda_3 + \lambda_4 = 0$.

Mourad also gave a geometric solution to the above case [81].

In 1983, Soules [122] uses the particular orthogonal matrix V_n (see Chapter 4) to get Perfect and Mirsky's condition. In fact, it is proved there(see [122]) that the symmetric matrix $M = V_n \Lambda V_n^T$ has nonnegative off diagonal entries while the ith diagonal entry of M is given by

$$m_{ii} = \frac{1}{n} + \sum_{k=i+1}^{n} \left(\frac{1}{(k-1)k} \right) \lambda_{n-k+2} + \left(\frac{i-1}{i} \right) \lambda_{n-i+2},$$

for $i = 1, ..., n$. Moreover the m_{ii} are increasing so the smallest one is m_{11}. Hence, if $m_{11} > 0$ then, M is doubly stochastic and σ is realisable by M.

Soules refined Perfect and Mirsky's condition and obtained the following theorem that we saw in the previous chapter. We will repeated here for convenience and completeness.

Theorem 4.2.7. *[122][Soules'condition]*
If $1 \geq \lambda_2 \geq ... \geq \lambda_n \geq -1$ *and*

$$\frac{1}{n} + \frac{n-m-1}{n(m+1)} \lambda_2 + \sum_{k=1}^{m} \frac{\lambda_{n-2k+2}}{(k+1)k} \geq 0, \tag{4.2}$$

holds with $n = 2m+2$ *if n even and* $n = 2m+1$ *if n odd, then there exists an* $n \times n$ *symmetric doubly stochastic matrix D such that D has eigenvalues* $1, \lambda_2, ..., \lambda_n$.

Remark 13. The proof of the preceding theorem is done by constructing from V_n another $n \times n$ matrix V_s, then constructing a symmetric matrix $N = V_s \Lambda V_s^T$ whose off-diagonal entries are all nonnegative and its smallest diagonal entry is given by the left-hand side of (4.2). So if (4.2) is valid then we obtain the symmetric doubly stochastic matrix D mentioned in the preceding theorem and so its trace is nonnegative. Consequently, if $1 \geq \lambda_2 \geq ... \geq \lambda_n \geq -1$ and (4.2) is true, then $1 + \lambda_2 + ... + \lambda_n \geq 0$.

Observation 4.2.8. *Perfect and Mirsky's condition implies Soules condition and the inclusion is strict.*

Proof. If $\sigma = \{1, \lambda_2, ..., \lambda_n\}$ with $1 \geq \lambda_2 \geq .... \geq \lambda_n \geq -1$ satisfies Perfect and Mirsky's condition, then we have

$$\frac{1}{n} + \frac{1}{n(n-1)} \lambda_2 + \frac{1}{(n-1)(n-2)} \lambda_3 + ... + \frac{1}{2.1} \lambda_n \geq 0.$$

Since $\lambda_2 \geq \lambda_j$ for $j = 3, ..., m$ and $\lambda_{n-2i} \geq \lambda_{n-i}$ for $i = 1, ..., m-1$, then

$$\begin{aligned}
&\frac{1}{n} + \frac{n-m-1}{n(m+1)}\lambda_2 + \sum_{k=1}^{m} \frac{\lambda_{n-2k+2}}{(k+1)k} \\
&= \frac{1}{n} + \left(\frac{1}{n(n-1)} + \frac{1}{(n-1)(n-2)} + ... + \frac{1}{(m+2)(m+1)} \right) \lambda_2 + \sum_{k=1}^{m} \frac{\lambda_{n-2(k-1)}}{(k+1)k} \\
&\geq \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + \frac{1}{(n-1)(n-2)}\lambda_3 + ... + \frac{1}{(m+2)(m+1)}\lambda_{n-m} + \sum_{k=1}^{m} \frac{\lambda_{n-(k-1)}}{(k+1)k} \\
&\geq 0
\end{aligned}$$

Therefore, σ verifies Soules condition. However, $(1, \frac{1}{2}, 0, \frac{3}{4})$ verifies Soules condition but does not satisfy Perfect and Mirsky's condition . Hence, the inclusion is strict. □

Many sufficient conditions appeared for SDIEP after Soules condition but no one is defined to be a refinement of Soules condition (see [58] for inclusion relations and independence between many sufficient conditions).

4.3 Recursive sufficient conditions for SDIEP

In this section, we will construct recursive sufficient conditions for the SDIEP of a special importance by using the preceding techniques of combining two doubly stochastic matrices to obtain a doubly stochastic matrix with a large spectra. But first, we give the following simple lemma which is very useful for us and will be used later.

Lemma 4.3.1. *Let $n > 1$ be a positive integer. Then, the following statements hold.*

i) $\frac{1}{n} + \frac{1}{n(n-1)} + ... + \frac{1}{3\times 2} + \frac{1}{2\times 1} = 1.$

ii) Let p be any positive integer which is greater than 1. If $n > 2^p$, then

$$\frac{1}{n} + \frac{1}{n(n-1)} + ... + \frac{1}{2^p(2^p+1)} + \frac{2^p-1}{2^p} = 1.$$

iii) For any positive integer $m \geq 1$, it holds that $\sum_{k=1}^{m} \frac{1}{(k+1)k} = 1 - \frac{1}{m+1} = \frac{m}{m+1}$. Consequently, for any positive integer $m \geq 2$, we have $\sum_{k=2}^{m} \frac{1}{(k+1)k} = \frac{m}{m+1} - \frac{1}{2}$.

iv) For any positive even integer $m \geq 2$, it holds that $\sum_{k=\frac{m}{2}}^{m-1} \frac{1}{(k+1)k} = \frac{1}{m}$. Hence, for any positive even integer $m \geq 4$, we have $\sum_{k=\frac{m}{2}}^{m-2} \frac{1}{(k+1)k} = \frac{1}{m} - \frac{1}{m(m-1)}$.

Proof. i) The left hand side in Part i) can be rewritten as

$$\begin{aligned}
&\frac{1}{n} + \sum_{k=1}^{n-1} \frac{1}{k(k+1)} = \frac{1}{n} + \sum_{k=1}^{n-1} \left(\frac{1}{k} - \frac{1}{k+1} \right) \\
&= \frac{1}{n} + \left(\frac{1}{n-1} - \frac{1}{n} \right) + \ldots + \left(\frac{1}{2} - \frac{1}{3} \right) + \left(1 - \frac{1}{2} \right) \\
&= 1.
\end{aligned}$$

ii) From Part i), we can write

$$\frac{1}{2^p(2^p-1)} + \frac{1}{(2^p-1)(2^p-2)} + \ldots + \frac{1}{3 \times 2} + \frac{1}{2 \times 1} \qquad = 1 - \frac{1}{2^p} = \frac{2^p-1}{2^p}.$$

It follows that

$$\begin{aligned}
&\frac{1}{n} + \frac{1}{n(n-1)} + \ldots + \frac{1}{2^p(2^p+1)} + \frac{2^p-1}{2^p} \\
&= \frac{1}{n} + \frac{1}{n(n-1)} + \ldots + \frac{1}{2^p(2^p+1)} + \frac{1}{2^p(2^p-1)} + \ldots + \frac{1}{2 \times 1} = 1.
\end{aligned}$$

iii) It follows easily from the first part.

iv) For $m = 2$, the assertion is true. Suppose $m > 2$, then from Part i), we certainly know that $\sum_{k=1}^{m-1} \frac{1}{(k+1)k} = 1 - \frac{1}{m} = \frac{m-1}{m}$. Therefore, $\sum_{k=1}^{\frac{m}{2}-1} \frac{1}{(k+1)k} + \sum_{k=\frac{m}{2}}^{m-1} \frac{1}{(k+1)k} = \frac{m-1}{m}$, and thus we obtain

$$\sum_{k=\frac{m}{2}}^{m-1} \frac{1}{(k+1)k} = \frac{m-1}{m} - \sum_{k=1}^{\frac{m}{2}-1} \frac{1}{(k+1)k} = \frac{m-1}{m} - \left(1 - \frac{1}{\frac{m}{2}} \right) = \frac{1}{m}.$$

□

4.3.1 Inaccurate refinement of Soules

In a recent paper [1], the authors used Lemma 4.2.3 and Lemma 4.2.4 to obtain a recursive method for constructing doubly stochastic matrices for the inverse eigenvalue problem by distinguishing between the cases n even and n odd. In addition, it is claimed that these new sufficient conditions improve Soules condition in both cases.

Theorem 4.3.2. *[1] Let $n \geq 5$. If $1 \geq \lambda_2 \geq .... \geq \lambda_n \geq -1$ and*

$$\frac{1}{n}+\frac{1}{n}\lambda_2+\frac{\frac{n}{2}-\left[\frac{n+2}{4}\right]}{\left(\frac{n}{2}\right)\left[\frac{n+2}{4}\right]}\lambda_4+\sum_{k=1}^{\left[\frac{n+2}{4}\right]-1}\frac{\lambda_{n-4k+4}}{(k+1)k}\geq 0, \tag{4.3}$$

for n even ($[.]$ stands for the integer part) and

$$\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_3+\sum_{k=1}^{\left[\frac{n+3}{4}\right]-1}\frac{\lambda_{n-4k+4}}{(k+1)k}\geq 0, \tag{4.4}$$

for n odd, holds. Then there exists an $n\times n$ symmetric doubly stochastic matrix C whose eigenvalues are $1,\lambda_2,\lambda_3,...,\lambda_n$.

Next we give a counterexample to the preceding theorem in the case when n is odd. For this purpose, let us recall the refinement of the JLL condition in a special case of n odd and $Tr(A)=0$, due to Laffey and Meehan [56] stated as follows.

Lemma 4.3.3. *[56] Let n be an odd integer and A be an $n\times n$ nonnegative matrix with $Tr(A)=0$ and let its spectrum be $\sigma(A)=(\lambda_1,...,\lambda_n)$. Then*

$$(n-1)s_4\geq s_2^2.$$

Consider now the following list of real numbers $\sigma=(1,1,1,-\frac{2}{3},-\frac{2}{3},-\frac{2}{3},-1)$. It is easy to see that σ satisfies (4.4). Hence, by Theorem 4.3.2, there exists a 7×7 symmetric doubly stochastic matrix A with spectrum σ. On the other hand, σ can not be the spectrum of any 7×7 nonnegative matrix, since otherwise, by Lemma 4.3.3, we have

$$6s_4\geq s_2^2,$$

that is,

$$\frac{248}{9}\geq\frac{256}{9}.$$

This leads to a contradiction.

More explicitly, the authors in [1] claimed (without justification) that when n is odd, and (4.4) is verified, then we have

$$\frac{1}{\frac{n+1}{2}}+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\frac{1}{\alpha}\lambda_3+\sum_{k=1}^{\left[\frac{n+3}{4}\right]-1}\frac{1}{\alpha}\frac{\lambda_{n-4k+4}}{(k+1)k}\geq 0, \tag{4.5}$$

and

$$\frac{1}{\frac{n-1}{2}}+\frac{\frac{n-1}{2}-\left[\frac{n+1}{4}\right]}{\left(\frac{n-1}{2}\right)\left[\frac{n+1}{4}\right]}\frac{1}{r}\lambda_4+\sum_{k=1}^{\left[\frac{n+1}{4}\right]-1}\frac{1}{r}\frac{\lambda_{n-4k+3}}{(k+1)k}\geq 0. \tag{4.6}$$

where $\alpha=\frac{(n-1)\lambda_2+(n+1)}{2n}$ and $r=\frac{(n+1)\lambda_2+(n-1)}{2n}$. In fact, if (4.4) is valid then it holds that

$$\begin{aligned}
0&\leq\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_3+\sum_{k=1}^{\left[\frac{n+3}{4}\right]-1}\frac{\lambda_{n-4k+4}}{(k+1)k}\\
&=\frac{(n-1)\lambda_2+(n+1)}{2n}\left(\frac{1}{\frac{n+1}{2}}+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\frac{1}{\alpha}\lambda_3+\sum_{k=1}^{\left[\frac{n+3}{4}\right]-1}\frac{1}{\alpha}\frac{\lambda_{n-4k+4}}{(k+1)k}\right)\\
&=\alpha\left(\frac{1}{\frac{n+1}{2}}+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\frac{1}{\alpha}\lambda_3+\sum_{k=1}^{\left[\frac{n+3}{4}\right]-1}\frac{1}{\alpha}\frac{\lambda_{n-4k+4}}{(k+1)k}\right).
\end{aligned}$$

As α is positive, then this shows that if (4.4) is valid then inequality (4.5) is always true. However, inequality (4.6) may not necessarily be true when (4.4) is valid, as can be easily checked for the list $\sigma=(1,1,1,-\frac{2}{3},-\frac{2}{3},-\frac{2}{3},-1)$ that satisfies (4.4) and does not verify (4.6).

4.3.2 An alternative sufficient condition when n is odd

In this section, we will present an alternative sufficient condition for the odd case by using the same recursive method for constructing doubly stochastic matrices. However, this will not be an improvement of Soules condition, but rather an independent condition as we shall prove. Before doing so, we need the following auxiliary lemma.

Lemma 4.3.4. *Let $n \geq 5$ be odd and let $1 \geq \lambda_2 \geq .... \geq \lambda_n \geq -1$. Then*

$$\frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{\frac{n+1}{2} - [\frac{n+3}{4}]}{(\frac{n+1}{2})[\frac{n+3}{4}]}\lambda_4 + \sum_{k=1}^{[\frac{n+3}{4}]-1} \frac{\lambda_{n-4k+4}}{(k+1)k}$$
$$\leq \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{4}{(n-1)(n+1)}\lambda_2 + \frac{\frac{n-1}{2} - [\frac{n+1}{4}]}{(\frac{n-1}{2})[\frac{n+1}{4}]}\lambda_4 + \sum_{k=1}^{[\frac{n+1}{4}]-1} \frac{\lambda_{n-4k+3}}{(k+1)k}.$$

Proof. We distinguish between the cases $n = 4m+1$ and $n = 4m+3$ (since n here takes only odd values).
For $n = 4m+1$, we have

$$\frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{m}{(2m+1)(m+1)}\lambda_4 + \sum_{k=1}^{m} \frac{\lambda_{n-4k+4}}{(k+1)k}$$
$$\leq \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{m}{(2m+1)(m+1)}\lambda_4 + \sum_{k=1}^{m} \frac{\lambda_{n-4k+3}}{(k+1)k} \qquad (\text{since } \lambda_{n-4k+4} \leq \lambda_{n-4k+3})$$
$$= \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{m}{(2m+1)(m+1)}\lambda_4 + \frac{1}{m(m+1)}\lambda_4 + \sum_{k=1}^{m-1} \frac{\lambda_{n-4k+3}}{(k+1)k}$$
$$= \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{m+1}{m(2m+1)}\lambda_4 + \sum_{k=1}^{m-1} \frac{\lambda_{n-4k+3}}{(k+1)k}$$
$$= \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \left(\frac{1}{2m(2m+1)} + \frac{1}{2m}\right)\lambda_4 + \sum_{k=1}^{m-1} \frac{\lambda_{n-4k+3}}{(k+1)k}$$
$$\leq \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{1}{2m(2m+1)}\lambda_2 + \frac{1}{2m}\lambda_4 + \sum_{k=1}^{m-1} \frac{\lambda_{n-4k+3}}{(k+1)k} \qquad (\text{since } \lambda_4 \leq \lambda_2)$$
$$= \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{4}{(n-1)(n+1)}\lambda_2 + \frac{\frac{n-1}{2} - [\frac{n+1}{4}]}{(\frac{n-1}{2})[\frac{n+1}{4}]}\lambda_4 + \sum_{k=1}^{[\frac{n+1}{4}]-1} \frac{\lambda_{n-4k+3}}{(k+1)k}.$$

For $n = 4m+3$, we have

$$\begin{aligned}
&\frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{1}{2m+2}\lambda_4 + \sum_{k=1}^{m} \frac{\lambda_{n-4k+4}}{(k+1)k} \\
&= \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \left(\frac{1}{(2m+2)(2m+1)} + \frac{m}{(m+1)(2m+1)}\right)\lambda_4 + \sum_{k=1}^{m} \frac{\lambda_{n-4k+4}}{(k+1)k} \\
&\leq \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{1}{(2m+2)(2m+1)}\lambda_2 + \frac{m}{(m+1)(2m+1)}\lambda_4 + \sum_{k=1}^{m} \frac{\lambda_{n-4k+3}}{(k+1)k} \\
&= \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{4}{(n-1)(n+1)}\lambda_2 + \frac{\frac{n-1}{2} - \left[\frac{n+1}{4}\right]}{\left(\frac{n-1}{2}\right)\left[\frac{n+1}{4}\right]}\lambda_4 + \sum_{k=1}^{[\frac{n+1}{4}]-1} \frac{\lambda_{n-4k+3}}{(k+1)k}.
\end{aligned}$$

Thus the proof is complete.

□

Theorem 4.3.5. *Let $n \geq 5$ be odd and let $1 \geq \lambda_2 \geq .... \geq \lambda_n \geq -1$. If*

$$\frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{\frac{n+1}{2} - \left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_4 + \sum_{k=1}^{[\frac{n+3}{4}]-1} \frac{\lambda_{n-4k+4}}{(k+1)k} \geq 0, \tag{4.7}$$

then there exists an $n \times n$ symmetric doubly stochastic matrix C with eigenvalues $1, \lambda_2, \lambda_3, ..., \lambda_n$.

Proof. As n odd, then by (4.7), we know that

$$\begin{aligned}
0 &\leq \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{\frac{n+1}{2} - \left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_4 + \sum_{k=1}^{[\frac{n+3}{4}]-1} \frac{\lambda_{n-4k+4}}{(k+1)k} \\
&\leq \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{\frac{n+1}{2} - \left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_3 + \sum_{k=1}^{[\frac{n+3}{4}]-1} \frac{\lambda_{n-4k+4}}{(k+1)k} \quad (\text{since } \lambda_4 \leq \lambda_3 \text{ and its coefficient is } > 0) \\
&= \left(\frac{(n-1)\lambda_2 + (n+1)}{2n}\right)\frac{1}{\frac{n+1}{2}} + \frac{\frac{n+1}{2} - \left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_3 + \sum_{k=1}^{[\frac{n+3}{4}]-1} \frac{\lambda_{n-4k+4}}{(k+1)k} \\
&= \alpha\left(\frac{1}{\frac{n+1}{2}} + \frac{\frac{n+1}{2} - \left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\frac{1}{\alpha}\lambda_3 + \sum_{k=1}^{[\frac{n+3}{4}]-1} \frac{1}{\alpha}\frac{\lambda_{n-4k+4}}{(k+1)k}\right),
\end{aligned}$$

where $\alpha = \frac{(n-1)\lambda_2+(n+1)}{2n} > 0$. In order to apply Theorem 4.2.7, we first show by contradiction that $\frac{\lambda_n}{\alpha} \geq -1$ and $\frac{\lambda_3}{\alpha} \leq 1$ (since the λ_i are in the decreasing order). Suppose that $\frac{\lambda_n}{\alpha} < -1$,

that is, $2n\lambda_n + (n-1)\lambda_2 + (n+1) < 0$. It then follows that, $\frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{2}{n+1}\lambda_n < 0$, and hence (4.7) implies that

$$\frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{\frac{n+1}{2} - \left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_4 + \sum_{k=1}^{\left[\frac{n+3}{4}\right]-1} \frac{\lambda_{n-4k+4}}{(k+1)k} \geq 0 > \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{2}{n+1}\lambda_n,$$

that is,

$$\frac{\frac{n+1}{2} - \left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_4 + \sum_{k=2}^{\left[\frac{n+3}{4}\right]-1} \frac{\lambda_{n-4k+4}}{(k+1)k} + \frac{1}{2}\lambda_n - \frac{2}{n+1}\lambda_n > 0.$$

Since $1 \geq \lambda_2 \geq ... \geq \lambda_n \geq -1$, then we can write

$$\begin{aligned}
0 &< \frac{\frac{n+1}{2} - \left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_2 + \sum_{k=2}^{\left[\frac{n+3}{4}\right]-1} \frac{\lambda_2}{(k+1)k} + \frac{1}{2}\lambda_n - \frac{2}{n+1}\lambda_n \\
&= \left(\frac{\frac{n+1}{2} - \left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]} + \sum_{k=2}^{\left[\frac{n+3}{4}\right]-1} \frac{1}{(k+1)k} \right)\lambda_2 + \left(\frac{1}{2} - \frac{2}{n+1}\right)\lambda_n \\
&\qquad \text{(using Lemma 4.3.1, Part iii) with } m = \left[\tfrac{n+3}{4}\right] - 1) \\
&= \left(\frac{\frac{n+1}{2} - \left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]} + \frac{\left[\frac{n+3}{4}\right]-1}{\left[\frac{n+3}{4}\right]} - \frac{1}{2} \right)\lambda_2 + \left(\frac{1}{2} - \frac{2}{n+1}\right)\lambda_n \\
&= \frac{n-3}{2(n+1)}(\lambda_2 + \lambda_n).
\end{aligned}$$

So $\lambda_2 + \lambda_n > 0$, and this implies that $0 > 2n\lambda_n + (n-1)\lambda_2 + (n+1) = n(\lambda_2 + \lambda_n) + n(1 + \lambda_n) + 1 - \lambda_2 > 0$, and we get a contradiction. Therefore, $\frac{\lambda_n}{\alpha} \geq -1$. Suppose now that $\frac{\lambda_3}{\alpha} > 1$, that is, $2n\lambda_3 - (n-1)\lambda_2 - (n+1) > 0$. But this implies that $2n\lambda_3 - (n-1)\lambda_2 - (n+1) = (n-1)(\lambda_3 - \lambda_2) + (n+1)(\lambda_3 - 1) > 0$, and we also get a contradiction. Therefore, $\frac{\lambda_3}{\alpha} \leq 1$ and hence by Soules condition, there exists an $(\frac{n+1}{2}) \times (\frac{n+1}{2})$ symmetric doubly stochastic matrix $\widehat{A}$ with eigenvalues $1, \frac{1}{\alpha}\lambda_3, \frac{1}{\alpha}\lambda_5, ..., \frac{1}{\alpha}\lambda_n$. On the other hand, by (4.7) and Lemma

4.3.4, we have

$$
\begin{aligned}
0 &\le \frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_4+\sum_{k=1}^{\left[\frac{n+3}{4}\right]-1}\frac{\lambda_{n-4k+4}}{(k+1)k}\\
&\le \frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{4}{(n-1)(n+1)}\lambda_2+\frac{\frac{n-1}{2}-\left[\frac{n+1}{4}\right]}{\left(\frac{n-1}{2}\right)\left[\frac{n+1}{4}\right]}\lambda_4+\sum_{k=1}^{\left[\frac{n+1}{4}\right]-1}\frac{\lambda_{n-4k+3}}{(k+1)k}\\
&= \left(\frac{(n+1)\lambda_2+(n-1)}{2n}\right)\frac{1}{\frac{n-1}{2}}+\frac{\frac{n-1}{2}-\left[\frac{n+1}{4}\right]}{\left(\frac{n-1}{2}\right)\left[\frac{n+1}{4}\right]}\lambda_4+\sum_{k=1}^{\left[\frac{n+1}{4}\right]-1}\frac{\lambda_{n-4k+3}}{(k+1)k}\\
&= r\left(\frac{1}{\frac{n-1}{2}}+\frac{\frac{n-1}{2}-\left[\frac{n+1}{4}\right]}{\left(\frac{n-1}{2}\right)\left[\frac{n+1}{4}\right]}\frac{1}{r}\lambda_4+\sum_{k=1}^{\left[\frac{n+1}{4}\right]-1}\frac{1}{r}\frac{\lambda_{n-4k+3}}{(k+1)k}\right),
\end{aligned}
$$

where $r=\frac{(n+1)\lambda_2+(n-1)}{2n}$. Next, we prove that r is positive. Indeed, by (4.7), we know that

$$
\begin{aligned}
0 &\le \frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_4+\sum_{k=1}^{\left[\frac{n+3}{4}\right]-1}\frac{\lambda_{n-4k+4}}{(k+1)k}\\
&\le \frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_2+\sum_{k=1}^{\left[\frac{n+3}{4}\right]-1}\frac{\lambda_2}{(k+1)k}\\
&\qquad\qquad (\text{since } \lambda_2\ge\lambda_4\ge\ldots, \text{ and their coefficients are } >0)\\
&= \frac{1}{n}+\left(\frac{n-1}{n(n+1)}+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}+\sum_{k=1}^{\left[\frac{n+3}{4}\right]-1}\frac{1}{(k+1)k}\right)\lambda_2\\
&= \frac{1}{n}+\left(\frac{n-1}{n(n+1)}+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}+1-\frac{1}{\left[\frac{n+3}{4}\right]}\right)\lambda_2\\
&= \frac{1}{n}+\frac{n-1}{n}\lambda_2.
\end{aligned}
$$

So obviously $\lambda_2\neq -1$, and it follows that $(n+1)\lambda_2+(n-1)=1+(n-1)\lambda_2+(n-2)+2\lambda_2>0$ for $n>4$. This implies that $r>0$. Also, as $(n+1)\lambda_2+(n-1)\le n+1+n-1=2n$, then $r\le 1$. Next, in order to again apply Theorem 4.2.7, we shall show that $\frac{\lambda_n}{r}\ge -1$ and $\frac{\lambda_3}{r}\le 1$ by contradiction. Suppose that $\frac{\lambda_n}{r}<-1$, that is, $2n\lambda_n+(n+1)\lambda_2+(n-1)<0$. It

follows that, $\frac{1}{n}+\frac{n+1}{n(n-1)}\lambda_2+\frac{2}{n-1}\lambda_n<0$. From (4.7), we get

$$\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_4+\sum_{k=1}^{\left[\frac{n+3}{4}\right]-1}\frac{\lambda_{n-4k+4}}{(k+1)k}\geq 0>\frac{1}{n}+\frac{n+1}{n(n-1)}\lambda_2+\frac{2}{n-1}\lambda_n,$$

that is,

$$\left(\frac{n-1}{n(n+1)}-\frac{n+1}{n(n-1)}\right)\lambda_2+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_4+\sum_{k=2}^{\left[\frac{n+3}{4}\right]-1}\frac{\lambda_{n-4k+4}}{(k+1)k}+\frac{1}{2}\lambda_n-\frac{2}{n-1}\lambda_n>0.$$

Since $1\geq\lambda_2\geq\ldots\geq\lambda_n\geq-1$, then we can write

$$\begin{aligned}
0&<\left(\frac{n-1}{n(n+1)}-\frac{n+1}{n(n-1)}\right)\lambda_2+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_2+\sum_{k=2}^{\left[\frac{n+3}{4}\right]-1}\frac{\lambda_2}{(k+1)k}+\frac{1}{2}\lambda_n-\frac{2}{n-1}\lambda_n\\
&=\left(\frac{n-1}{n(n+1)}-\frac{n+1}{n(n-1)}+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}+\sum_{k=2}^{\left[\frac{n+3}{4}\right]-1}\frac{1}{(k+1)k}\right)\lambda_2+\left(\frac{1}{2}-\frac{2}{n-1}\right)\lambda_n\\
&=\left(\frac{n-1}{n(n+1)}-\frac{n+1}{n(n-1)}+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}+\frac{\left[\frac{n+3}{4}\right]-1}{\left[\frac{n+3}{4}\right]}-\frac{1}{2}\right)\lambda_2+\left(\frac{1}{2}-\frac{2}{n-1}\right)\lambda_n\\
&\qquad\qquad\text{(by Lemma 4.3.1, Part iii))}\\
&=\left(\frac{n-1}{n(n+1)}-\frac{n+1}{n(n-1)}+\frac{1}{\left[\frac{n+3}{4}\right]}-\frac{1}{\frac{n+1}{2}}+1-\frac{1}{\left[\frac{n+3}{4}\right]}-\frac{1}{2}\right)\lambda_2+\left(\frac{1}{2}-\frac{2}{n-1}\right)\lambda_n\\
&=\frac{n-5}{2(n-1)}(\lambda_2+\lambda_n).
\end{aligned}$$

For $n=5$, we have $0<0$ a contradiction. If $n\neq 5$, we have $\lambda_2+\lambda_n>0$, and it then follows that $0>2n\lambda_n+(n+1)\lambda_2+(n-1)=n(\lambda_2+\lambda_n)+(n-1)(1+\lambda_n)+(\lambda_2+\lambda_n)>0$, and we have a contradiction. Suppose now that $\frac{\lambda_3}{r}>1$. This implies that $2n\lambda_3-(n+1)\lambda_2-(n-1)=(n+1)(\lambda_3-\lambda_2)+(n-1)(\lambda_3-1)>0$, and we again get a contradiction. Therefore, by Soules condition (1), there exists an $(\frac{n-1}{2})\times(\frac{n-1}{2})$ symmetric doubly stochastic matrix $\widehat{B}$ with eigenvalues $1,\frac{1}{r}\lambda_4,\frac{1}{r}\lambda_6,\ldots,\frac{1}{r}\lambda_{n-1}$.

Finally, applying Lemma 4.2.4 with $A=\widehat{A}$ and $B=\widehat{B}$, we obtain an $n\times n$ symmetric doubly stochastic matrix C whose eigenvalues are $1,\lambda_2,\ldots,\lambda_n$.

□

Notation. We will say that New condition 1 holds if (4.3) is satisfied when n is even, while (4.7) is satisfied if n is odd.

Observation 4.3.6. *For n even, Soules implies New condition 1. For n odd, Soules and New condition 1 are independent.*

Proof. Let n be even. Suppose now that σ satisfies Soules condition. Since

$$\lambda_{n-4i} \geq \lambda_{n-2i}, \qquad i = 1,2,...,\left[\tfrac{n+2}{4}\right]-2$$

and

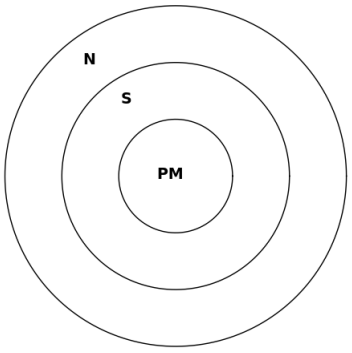
$$\lambda_4 \geq \lambda_i, \qquad i = 6,8,...,n-\left[\tfrac{n+2}{2}\right]+2,$$

then σ obviously satisfies New condition 1. For the second part, it suffices to check that for $n = 7$, the list $(1,0,0,0,-\frac{2}{7},-\frac{2}{7},-\frac{2}{7})$ satisfies New condition 1 and does not satisfy Soules condition, while the list $(1,0,0,-\frac{3}{14},-\frac{3}{14},-\frac{3}{14},-\frac{3}{14})$ does the opposite i.e. it satisfies Soules condition and does not satisfy New condition 1. □

Notation. For convenience, we shall denote by PM, S, N the sets of all n-tuples σ satisfying Perfect and Mirsky's condition, Soules condition and New condition 1, respectively.

Observation 4.3.7. *The Venn diagrams of the point sets PM, S, and N are the following:*

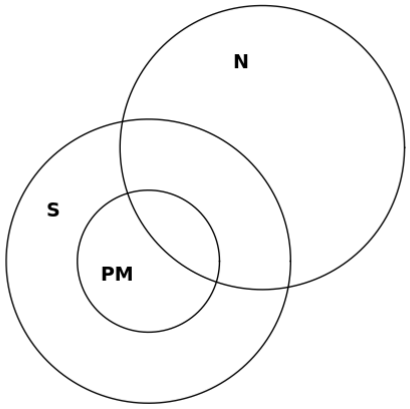
- *for n even:*



We illustrate the above diagram by the following simple examples:

1. $(1,0,0,0,0,0,0,-\frac{1}{4}) \in PM \cap S \cap N$.
2. $(1,0,0,0,0,0,-\frac{1}{4},-\frac{1}{4}) \in (S \cap N) \setminus PM$.
3. $(1,0,0,0,0,-\frac{1}{4},-\frac{1}{4},-\frac{1}{4}) \in N \setminus (PM \cup S)$.

- *for n odd:*



We illustrate the above diagram by the following examples:

1. $(1,0,0,-\frac{1}{14},-\frac{1}{14},-\frac{1}{14},-\frac{3}{14}) \in PM \cap S \cap N$.
2. $(1,0,0,-\frac{1}{14},-\frac{1}{14},-\frac{1}{14},-\frac{24}{100}) \in (PM \cap S) \setminus N$.
3. $(1,0,0,-\frac{3}{14},-\frac{3}{14},-\frac{3}{14},-\frac{3}{14}) \in S \setminus (PM \cup N)$.
4. $(1,0,0,0,-\frac{2}{7},-\frac{2}{7},-\frac{2}{7}) \in N \setminus (PM \cup S)$.
5. $(1,0,0,0,0,-\frac{1}{7},-\frac{2}{7}) \in (S \cap N) \setminus PM$.

4.3.3 The recursive approach to SDIEP

In this section, we will go a step further by doing another recursive phase in constructing doubly stochastic elements using Lemma 4.2.4 to conclude by a important conjecture that yields to infinite independent sufficient conditions for SDIEP. It is worthy to point out here that, the technique used in each case of the following proofs is similar to that of Theorem 4.3.5 with only minor difference. First, we present the following auxiliary lemma.

Lemma 4.3.8. *Let $n = 4m+1 \geq 9$ where m is a positive integer and $1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq -1$. Then*

$$\frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{2(n-1)}{(n+1)(n+3)}\lambda_4 + \frac{\frac{n+3}{4} - [\frac{n+7}{8}]}{(\frac{n+3}{4})[\frac{n+7}{8}]}\lambda_8 + \sum_{k=1}^{[\frac{n+7}{8}]-1} \frac{\lambda_{n-8k+8}}{(k+1)k}$$
$$\leq \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{4}{(n-1)(n+1)}\lambda_2 + \frac{2}{n-1}\lambda_4 + \frac{\frac{n-1}{4} - [\frac{n+3}{8}]}{(\frac{n-1}{4})[\frac{n+3}{8}]}\lambda_8 + \sum_{k=1}^{[\frac{n+3}{8}]-1} \frac{\lambda_{n-8k+7}}{(k+1)k}.$$

Proof. As $n = 4m+1$, we have

$$\begin{aligned}
&\frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{2(n-1)}{(n+1)(n+3)}\lambda_4 + \frac{\frac{n+3}{4} - \left[\frac{n+7}{8}\right]}{\left(\frac{n+3}{4}\right)\left[\frac{n+7}{8}\right]}\lambda_8 + \sum_{k=1}^{\left[\frac{n+7}{8}\right]-1} \frac{\lambda_{n-8k+8}}{(k+1)k} \\
&= \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \left(\frac{1}{2m(2m+1)} + \frac{m-1}{2m(m+1)}\right)\lambda_4 + \frac{\frac{n+3}{4} - \left[\frac{n+7}{8}\right]}{\left(\frac{n+3}{4}\right)\left[\frac{n+7}{8}\right]}\lambda_8 + \sum_{k=1}^{\left[\frac{n+7}{8}\right]-1} \frac{\lambda_{n-8k+8}}{(k+1)k} \\
&\leq \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{1}{2m(2m+1)}\lambda_2 + \frac{m-1}{2m(m+1)}\lambda_4 + \frac{\frac{n+3}{4} - \left[\frac{n+7}{8}\right]}{\left(\frac{n+3}{4}\right)\left[\frac{n+7}{8}\right]}\lambda_8 + \sum_{k=1}^{\left[\frac{n+7}{8}\right]-1} \frac{\lambda_{n-8k+7}}{(k+1)k}.
\end{aligned}$$

Next, we distinguish between the cases m odd and m even.
Case 1: For $m = 2p+1$, then $n = 8p+5$ and hence

$$\left[\frac{n+7}{8}\right] = \left[\frac{n+3}{8}\right] = p+1.$$

Therefore,

$$\begin{aligned}
&\frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{1}{2m(2m+1)}\lambda_2 + \frac{m-1}{2m(m+1)}\lambda_4 + \frac{\frac{n+3}{4} - \left[\frac{n+7}{8}\right]}{\left(\frac{n+3}{4}\right)\left[\frac{n+7}{8}\right]}\lambda_8 + \sum_{k=1}^{\left[\frac{n+7}{8}\right]-1} \frac{\lambda_{n-8k+7}}{(k+1)k} \\
&= \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{1}{2m(2m+1)}\lambda_2 + \frac{m-1}{2m(m+1)}\lambda_4 + \frac{1}{2p+2}\lambda_8 + \sum_{k=1}^{p} \frac{\lambda_{n-8k+7}}{(k+1)k} \\
&= \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{1}{2m(2m+1)}\lambda_2 + \frac{m-1}{2m(m+1)}\lambda_4 + \left(\frac{1}{(2p+1)(2p+2)} + \frac{p}{(2p+1)(p+1)}\right)\lambda_8 \\
&\quad + \sum_{k=1}^{p} \frac{\lambda_{n-8k+7}}{(k+1)k} \\
&\leq \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{1}{2m(2m+1)}\lambda_2 + \frac{m-1}{2m(m+1)}\lambda_4 + \frac{1}{(2p+1)(2p+2)}\lambda_4 + \frac{p}{(2p+1)(p+1)}\lambda_8 \\
&\quad + \sum_{k=1}^{p} \frac{\lambda_{n-8k+7}}{(k+1)k} \\
&= \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{1}{2m(2m+1)}\lambda_2 + \frac{1}{2m}\lambda_4 + \frac{p}{(2p+1)(p+1)}\lambda_8 + \sum_{k=1}^{p} \frac{\lambda_{n-8k+7}}{(k+1)k} \\
&= \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{4}{(n-1)(n+1)}\lambda_2 + \frac{2}{n-1}\lambda_4 + \frac{\frac{n-1}{4} - \left[\frac{n+3}{8}\right]}{\left(\frac{n-1}{4}\right)\left[\frac{n+3}{8}\right]}\lambda_8 + \sum_{k=1}^{\left[\frac{n+3}{8}\right]-1} \frac{\lambda_{n-8k+7}}{(k+1)k}.
\end{aligned}$$

Case 2: For $m = 2p$, then $n = 8p+1$ and in this case, we have

$$\left[\frac{n+7}{8}\right] = \left[\frac{n+3}{8}\right] + 1 = p+1.$$

Therefore,

$$\begin{aligned}
&\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{1}{2m(2m+1)}\lambda_2+\frac{m-1}{2m(m+1)}\lambda_4+\frac{\frac{n+3}{4}-\left[\frac{n+7}{8}\right]}{\left(\frac{n+3}{4}\right)\left[\frac{n+7}{8}\right]}\lambda_8+\sum_{k=1}^{\left[\frac{n+7}{8}\right]-1}\frac{\lambda_{n-8k+7}}{(k+1)k}\\
&=\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{1}{2m(2m+1)}\lambda_2+\frac{m-1}{2m(m+1)}\lambda_4+\frac{p}{(2p+1)(p+1)}\lambda_8+\sum_{k=1}^{p}\frac{\lambda_{n-8k+7}}{(k+1)k}\\
&=\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{1}{2m(2m+1)}\lambda_2+\frac{m-1}{2m(m+1)}\lambda_4+\frac{p}{(2p+1)(p+1)}\lambda_8+\frac{1}{p(p+1)}\lambda_8\\
&\quad+\sum_{k=1}^{p-1}\frac{\lambda_{n-8k+7}}{(k+1)k}\\
&=\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{1}{2m(2m+1)}\lambda_2+\frac{m-1}{2m(m+1)}\lambda_4+\frac{p+1}{p(2p+1)}\lambda_8+\sum_{k=1}^{p-1}\frac{\lambda_{n-8k+7}}{(k+1)k}\\
&=\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{1}{2m(2m+1)}\lambda_2+\frac{m-1}{2m(m+1)}\lambda_4+\left(\frac{1}{2p(2p+1)}+\frac{1}{2p}\right)\lambda_8+\sum_{k=1}^{p-1}\frac{\lambda_{n-8k+7}}{(k+1)k}\\
&\leq\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{1}{2m(2m+1)}\lambda_2+\frac{m-1}{2m(m+1)}\lambda_4+\frac{1}{2p(2p+1)}\lambda_4+\frac{1}{2p}\lambda_8+\sum_{k=1}^{p-1}\frac{\lambda_{n-8k+7}}{(k+1)k}\\
&=\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{1}{2m(2m+1)}\lambda_2+\frac{1}{2m}\lambda_4+\frac{1}{2p}\lambda_8+\sum_{k=1}^{p-1}\frac{\lambda_{n-8k+7}}{(k+1)k}\\
&=\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{4}{(n-1)(n+1)}\lambda_2+\frac{2}{n-1}\lambda_4+\frac{\frac{n-1}{4}-\left[\frac{n+3}{8}\right]}{\left(\frac{n-1}{4}\right)\left[\frac{n+3}{8}\right]}\lambda_8+\sum_{k=1}^{\left[\frac{n+3}{8}\right]-1}\frac{\lambda_{n-8k+7}}{(k+1)k}.
\end{aligned}$$

Thus the proof is complete. □

By substituting in Lemma 4.2.4, the doubly stochastic matrices obtained when New condition 1 holds, we have the following theorem.

Theorem 4.3.9. *Let $n \geq 8$ and let $1 \geq \lambda_2 \geq .... \geq \lambda_n \geq -1$. If*

$$\frac{1}{n}+\frac{1}{n}\lambda_2+\frac{2}{n}\lambda_4+\frac{\frac{n}{4}-\left[\frac{n+4}{8}\right]}{\left(\frac{n}{4}\right)\left[\frac{n+4}{8}\right]}\lambda_8+\sum_{k=1}^{\left[\frac{n+4}{8}\right]-1}\frac{\lambda_{n-8k+8}}{(k+1)k}\geq 0, \tag{4.8}$$

for $n = 4m$,

$$\frac{1}{n}+\frac{1}{n}\lambda_2+\frac{2(n-2)}{n(n+2)}\lambda_4+\frac{\frac{n+2}{4}-\left[\frac{n+6}{8}\right]}{\left(\frac{n+2}{4}\right)\left[\frac{n+6}{8}\right]}\lambda_8+\sum_{k=1}^{\left[\frac{n+6}{8}\right]-1}\frac{\lambda_{n-8k+8}}{(k+1)k}\geq 0, \tag{4.9}$$

for $n = 4m+2$,

$$\frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{2}{n+1}\lambda_4 + \frac{\frac{n+1}{4} - \left[\frac{n+5}{8}\right]}{\left(\frac{n+1}{4}\right)\left[\frac{n+5}{8}\right]}\lambda_8 + \sum_{k=1}^{\left[\frac{n+5}{8}\right]-1} \frac{\lambda_{n-8k+8}}{(k+1)k} \geq 0, \tag{4.10}$$

for $n = 4m+3$, *and*

$$\frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{2(n-1)}{(n+1)(n+3)}\lambda_4 + \frac{\frac{n+3}{4} - \left[\frac{n+7}{8}\right]}{\left(\frac{n+3}{4}\right)\left[\frac{n+7}{8}\right]}\lambda_8 + \sum_{k=1}^{\left[\frac{n+7}{8}\right]-1} \frac{\lambda_{n-8k+8}}{(k+1)k} \geq 0, \tag{4.11}$$

for $n = 4m+1$ *hold, then there exists an* $n \times n$ *symmetric doubly stochastic matrix with eigenvalues* $1, \lambda_2, ..., \lambda_n$.

Proof. The proof is similar to that of Theorem 4.3.5. First suppose that $n = 4m$. By (4.8), we have

$$\alpha\left(\frac{2}{n} + \frac{2}{n}\frac{1}{\alpha}\lambda_4 + \frac{\frac{n}{4} - \left[\frac{n+4}{8}\right]}{\left(\frac{n}{4}\right)\left[\frac{n+4}{8}\right]}\frac{1}{\alpha}\lambda_8 + \sum_{k=1}^{\left[\frac{n+4}{8}\right]-1} \frac{1}{\alpha}\frac{\lambda_{n-8k+8}}{(k+1)k}\right) \geq 0,$$

where $\alpha = \frac{1+\lambda_2}{2} \geq 0$. We show by contradiction that $\alpha > 0$. Suppose that $\alpha = 0$. Then, $\lambda_2 = -1$, which implies that $\lambda_3 = ... = \lambda_n = -1$. It follows that

$$\frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{2}{n}\lambda_4 + \frac{\frac{n}{4} - \left[\frac{n+4}{8}\right]}{\left(\frac{n}{4}\right)\left[\frac{n+4}{8}\right]}\lambda_8 + \sum_{k=1}^{\left[\frac{n+4}{8}\right]-1} \frac{\lambda_{n-8k+8}}{(k+1)k} < 0,$$

and this contradicts (4.8). Equivalently, we have

$$\frac{1}{\frac{n}{2}} + \frac{1}{\frac{n}{2}}\frac{1}{\alpha}\lambda_4 + \frac{\frac{n}{4} - \left[\frac{n+4}{8}\right]}{\left(\frac{n}{4}\right)\left[\frac{n+4}{8}\right]}\frac{1}{\alpha}\lambda_8 + \sum_{k=1}^{\left[\frac{n+4}{8}\right]-1} \frac{1}{\alpha}\frac{\lambda_{n-8k+8}}{(k+1)k} \geq 0,$$

and hence

$$\frac{1}{\frac{n}{2}} + \frac{1}{\frac{n}{2}}\frac{1}{\alpha}\lambda_3 + \frac{\frac{n}{4} - \left[\frac{n+4}{8}\right]}{\left(\frac{n}{4}\right)\left[\frac{n+4}{8}\right]}\frac{1}{\alpha}\lambda_7 + \sum_{k=1}^{\left[\frac{n+4}{8}\right]-1} \frac{1}{\alpha}\frac{\lambda_{n-8k+7}}{(k+1)k} \geq 0.$$

We first show by contradiction that $\frac{\lambda_n}{\alpha} \geq -1$. The proof of this, is virtually the same as earlier with only minor difference. Indeed, suppose that $\frac{\lambda_n}{\alpha} < -1$, that is, $2\lambda_n + \lambda_2 + 1 < 0$. It follows that, $\frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{2}{n}\lambda_n < 0$. From condition (4.8), we conclude that

$$\frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{2}{n}\lambda_4 + \frac{\frac{n}{4} - \left[\frac{n+4}{8}\right]}{\left(\frac{n}{4}\right)\left[\frac{n+4}{8}\right]}\lambda_8 + \sum_{k=1}^{\left[\frac{n+4}{8}\right]-1} \frac{\lambda_{n-8k+8}}{(k+1)k} \geq 0 > \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{2}{n}\lambda_n,$$

that is,

$$\frac{2}{n}\lambda_4+\frac{\frac{n}{4}-[\frac{n+4}{8}]}{(\frac{n}{4})[\frac{n+4}{8}]}\lambda_8+\sum_{k=2}^{[\frac{n+4}{8}]-1}\frac{\lambda_{n-8k+8}}{(k+1)k}+\frac{1}{2}\lambda_n-\frac{2}{n}\lambda_n>0.$$

Since $1\geq\lambda_2\geq...\geq\lambda_n\geq-1$, then we clearly have

$$\begin{aligned}0&<\left(\frac{2}{n}+\frac{\frac{n}{4}-[\frac{n+4}{8}]}{(\frac{n}{4})[\frac{n+4}{8}]}+\sum_{k=2}^{[\frac{n+4}{8}]-1}\frac{1}{(k+1)k}\right)\lambda_2+\frac{1}{2}\lambda_n-\frac{2}{n}\lambda_n\\&=\left(\frac{2}{n}+\frac{\frac{n}{4}-[\frac{n+4}{8}]}{(\frac{n}{4})[\frac{n+4}{8}]}+\frac{[\frac{n+4}{8}]-1}{[\frac{n+4}{8}]}-\frac{1}{2}\right)\lambda_2+\frac{1}{2}\lambda_n-\frac{2}{n}\lambda_n\\&\qquad\qquad\text{(using Lemma 4.3.1, Part iii) with } m=[\tfrac{n+4}{8}]-1)\\&=\frac{n-4}{2n}(\lambda_2+\lambda_n).\end{aligned}$$

Therefore $\lambda_2+\lambda_n>0$. It follows that $0>2\lambda_n+\lambda_2+1=(\lambda_2+\lambda_n)+(1+\lambda_n)>0$, and we have a contradiction. Next, we show that $\frac{\lambda_3}{\alpha}\leq 1$ by contradiction. If $\frac{\lambda_3}{\alpha}>1$, then clearly $2\lambda_3-\lambda_2-1>0$, and we get a contradiction. Thus, by New condition 1, there are $(\frac{n}{2})\times(\frac{n}{2})$ symmetric doubly stochastic matrices $\widehat{A}$ and $\widehat{B}$ with eigenvalues $1,\frac{1}{\alpha}\lambda_4,\frac{1}{\alpha}\lambda_6,...,\frac{1}{\alpha}\lambda_n$ and $1,\frac{1}{\alpha}\lambda_3,\frac{1}{\alpha}\lambda_5,...,\frac{1}{\alpha}\lambda_{n-1}$ respectively. Now applying Lemma 4.2.4 with $A=\widehat{A}$, $B=\widehat{B}$ and $r=\alpha=\frac{1+\lambda_2}{2}$, we obtain an $n\times n$ symmetric doubly stochastic matrix C whose eigenvalues are $1,\lambda_2,...,\lambda_n$.

For $n=4m+2$, using (4.9), we conclude that

$$\frac{1}{\frac{n}{2}}+\frac{\frac{n}{2}-1}{\frac{n}{2}(\frac{n}{2}+1)}\frac{1}{\alpha}\lambda_4+\frac{\frac{n+2}{4}-[\frac{n+6}{8}]}{(\frac{n+2}{4})[\frac{n+6}{8}]}\frac{1}{\alpha}\lambda_8+\sum_{k=1}^{[\frac{n+6}{8}]-1}\frac{1}{\alpha}\frac{\lambda_{n-8k+8}}{(k+1)k}\geq 0,$$

and

$$\frac{1}{\frac{n}{2}}+\frac{\frac{n}{2}-1}{\frac{n}{2}(\frac{n}{2}+1)}\frac{1}{\alpha}\lambda_3+\frac{\frac{n+2}{4}-[\frac{n+6}{8}]}{(\frac{n+2}{4})[\frac{n+6}{8}]}\frac{1}{\alpha}\lambda_7+\sum_{k=1}^{[\frac{n+6}{8}]-1}\frac{1}{\alpha}\frac{\lambda_{n-8k+7}}{(k+1)k}\geq 0,$$

where $\alpha=\frac{1+\lambda_2}{2}>0$. We show that $\frac{\lambda_n}{\alpha}\geq-1$ by a similar argument as in the previous case. Suppose that $\frac{\lambda_n}{\alpha}<-1$, that is, $2\lambda_n+\lambda_2+1<0$. It follows that, $\frac{1}{n}+\frac{1}{n}\lambda_2+\frac{2}{n}\lambda_n<0$. Then by (4.9), it holds that

$$\frac{1}{n}+\frac{1}{n}\lambda_2+\frac{2(n-2)}{n(n+2)}\lambda_4+\frac{\frac{n+2}{4}-[\frac{n+6}{8}]}{(\frac{n+2}{4})[\frac{n+6}{8}]}\lambda_8+\sum_{k=1}^{[\frac{n+6}{8}]-1}\frac{\lambda_{n-8k+8}}{(k+1)k}\geq 0>\frac{1}{n}+\frac{1}{n}\lambda_2+\frac{2}{n}\lambda_n,$$

that is,

$$\frac{2(n-2)}{n(n+2)}\lambda_4+\frac{\frac{n+2}{4}-\left[\frac{n+6}{8}\right]}{\left(\frac{n+2}{4}\right)\left[\frac{n+6}{8}\right]}\lambda_8+\sum_{k=2}^{\left[\frac{n+6}{8}\right]-1}\frac{\lambda_{n-8k+8}}{(k+1)k}+\frac{1}{2}\lambda_n-\frac{2}{n}\lambda_n>0.$$

Since $1\geq\lambda_2\geq\ldots\geq\lambda_n\geq-1$, we have

$$\begin{aligned}0&<\left(\frac{2(n-2)}{n(n+2)}+\frac{\frac{n+2}{4}-\left[\frac{n+6}{8}\right]}{\left(\frac{n+2}{4}\right)\left[\frac{n+6}{8}\right]}+\sum_{k=2}^{\left[\frac{n+6}{8}\right]-1}\frac{1}{(k+1)k}\right)\lambda_2+\frac{1}{2}\lambda_n-\frac{2}{n}\lambda_n\\&=\left(\frac{2(n-2)}{n(n+2)}+\frac{\frac{n+2}{4}-\left[\frac{n+6}{8}\right]}{\left(\frac{n+2}{4}\right)\left[\frac{n+6}{8}\right]}+\frac{\left[\frac{n+6}{8}\right]-1}{\left[\frac{n+6}{8}\right]}-\frac{1}{2}\right)\lambda_2+\frac{1}{2}\lambda_n-\frac{2}{n}\lambda_n\\&\qquad\qquad\text{(using Lemma 4.3.1, Part iii) with } m=\left[\tfrac{n+6}{8}\right]-1)\\&=\frac{n-4}{2n}(\lambda_2+\lambda_n).\end{aligned}$$

Hence $\lambda_2+\lambda_n>0$, and it then follows that $0>2\lambda_n+\lambda_2+1=(\lambda_2+\lambda_n)+(1+\lambda_n)>0$, and we have a contradiction. Note that $\frac{\lambda_3}{\alpha}\leq 1$ by a similar argument as in the previous case. Thus, by New condition 1, there are $(\frac{n}{2})\times(\frac{n}{2})$ symmetric doubly stochastic matrices $\widehat{A}$ and $\widehat{B}$ with eigenvalues $1,\frac{1}{\alpha}\lambda_4,\frac{1}{\alpha}\lambda_6,\ldots,\frac{1}{\alpha}\lambda_n$ and $1,\frac{1}{\alpha}\lambda_3,\frac{1}{\alpha}\lambda_5,\ldots,\frac{1}{\alpha}\lambda_{n-1}$ respectively. Now by applying Lemma 4.2.4 with $A=\widehat{A}$, $B=\widehat{B}$ and $r=\alpha=\frac{1+\lambda_2}{2}$, the proof of the second part can be achieved.

For the case $n=4m+3$, using (4.10) we can write

$$\begin{aligned}0&\leq\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{2}{n+1}\lambda_4+\frac{\frac{n+1}{4}-\left[\frac{n+5}{8}\right]}{\left(\frac{n+1}{4}\right)\left[\frac{n+5}{8}\right]}\lambda_8+\sum_{k=1}^{\left[\frac{n+5}{8}\right]-1}\frac{\lambda_{n-8k+8}}{(k+1)k}\\&\leq\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{2}{n+1}\lambda_3+\frac{\frac{n+1}{4}-\left[\frac{n+5}{8}\right]}{\left(\frac{n+1}{4}\right)\left[\frac{n+5}{8}\right]}\lambda_7+\sum_{k=1}^{\left[\frac{n+5}{8}\right]-1}\frac{\lambda_{n-8k+8}}{(k+1)k}\\&=\alpha\left(\frac{1}{\frac{n+1}{2}}+\frac{1}{\frac{n+1}{2}}\frac{1}{\alpha}\lambda_3+\frac{\frac{n+1}{4}-\left[\frac{n+5}{8}\right]}{\left(\frac{n+1}{4}\right)\left[\frac{n+5}{8}\right]}\frac{1}{\alpha}\lambda_7+\sum_{k=1}^{\left[\frac{n+5}{8}\right]-1}\frac{1}{\alpha}\frac{\lambda_{n-8k+8}}{(k+1)k}\right),\end{aligned}$$

where $\alpha=\frac{(n-1)\lambda_2+(n+1)}{2n}>0$. As earlier, we shall show now that $\frac{\lambda_n}{\alpha}\geq-1$ by contradiction. Suppose that $\frac{\lambda_n}{\alpha}<-1$, that is, $2n\lambda_n+(n-1)\lambda_2+(n+1)<0$. It follows that $\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+$

$\frac{2}{n+1}\lambda_n < 0$. From (4.10), we conclude that

$$\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{2}{n+1}\lambda_4+\frac{\frac{n+1}{4}-\left[\frac{n+5}{8}\right]}{\left(\frac{n+1}{4}\right)\left[\frac{n+5}{8}\right]}\lambda_8+\sum_{k=1}^{\left[\frac{n+5}{8}\right]-1}\frac{\lambda_{n-8k+8}}{(k+1)k}\geq 0>\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{2}{n+1}\lambda_n,$$

that is,

$$\frac{2}{n+1}\lambda_4+\frac{\frac{n+1}{4}-\left[\frac{n+5}{8}\right]}{\left(\frac{n+1}{4}\right)\left[\frac{n+5}{8}\right]}\lambda_8+\sum_{k=2}^{\left[\frac{n+5}{8}\right]-1}\frac{\lambda_{n-8k+8}}{(k+1)k}+\frac{1}{2}\lambda_n-\frac{2}{n+1}\lambda_n>0.$$

Since $1\geq\lambda_2\geq\ldots\geq\lambda_n\geq -1$, we have

$$\begin{aligned}0&<\left(\frac{2}{n+1}+\frac{\frac{n+1}{4}-\left[\frac{n+5}{8}\right]}{\left(\frac{n+1}{4}\right)\left[\frac{n+5}{8}\right]}+\sum_{k=2}^{\left[\frac{n+5}{8}\right]-1}\frac{1}{(k+1)k}\right)\lambda_2+\frac{1}{2}\lambda_n-\frac{2}{n+1}\lambda_n\\&=\left(\frac{2}{n+1}+\frac{\frac{n+1}{4}-\left[\frac{n+5}{8}\right]}{\left(\frac{n+1}{4}\right)\left[\frac{n+5}{8}\right]}+\frac{\left[\frac{n+5}{8}\right]-1}{\left[\frac{n+5}{8}\right]}-\frac{1}{2}\right)\lambda_2+\frac{1}{2}\lambda_n-\frac{2}{n+1}\lambda_n\\&\qquad\qquad\text{(using Lemma 3, Part iii) with } m=\left[\tfrac{n+5}{8}\right]-1)\\&=\frac{n-3}{2(n+1)}(\lambda_2+\lambda_n).\end{aligned}$$

Hence $\lambda_2+\lambda_n>0$, and it then follows that $0>2n\lambda_n+(n-1)\lambda_2+(n+1)=n(\lambda_2+\lambda_n)+n(1+\lambda_n)+1-\lambda_2>0$, and again we get a contradiction. In addition, $\frac{\lambda_3}{\alpha}\leq 1$ by a similar argument as in the proof of Theorem 4.3.5. Therefore by New condition 1, there exists an $(\frac{n+1}{2})\times(\frac{n+1}{2})$ symmetric doubly stochastic matrix $\widehat{A}$ with eigenvalues $1,\frac{1}{\alpha}\lambda_3,\frac{1}{\alpha}\lambda_5,\ldots,\frac{1}{\alpha}\lambda_n$.

Now from (4.10), we can write

$$
\begin{aligned}
0 &\leq \frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{2}{n+1}\lambda_4+\frac{\frac{n+1}{4}-\left[\frac{n+5}{8}\right]}{\left(\frac{n+1}{4}\right)\left[\frac{n+5}{8}\right]}\lambda_8+\sum_{k=1}^{\left[\frac{n+5}{8}\right]-1}\frac{\lambda_{n-8k+8}}{(k+1)k}\\
&=\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\left(\frac{1}{\left(\frac{n+1}{2}\right)\frac{n-1}{2}}+\frac{\frac{n-3}{2}}{\left(\frac{n+1}{2}\right)\frac{n-1}{2}}\right)\lambda_4+\frac{\frac{n+1}{4}-\left[\frac{n+5}{8}\right]}{\left(\frac{n+1}{4}\right)\left[\frac{n+5}{8}\right]}\lambda_8+\sum_{k=1}^{\left[\frac{n+5}{8}\right]-1}\frac{\lambda_{n-8k+8}}{(k+1)k}\\
&\leq\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{1}{\left(\frac{n+1}{2}\right)\frac{n-1}{2}}\lambda_2+\frac{\frac{n-3}{2}}{\left(\frac{n+1}{2}\right)\frac{n-1}{2}}\lambda_4+\frac{\frac{n+1}{4}-\left[\frac{n+5}{8}\right]}{\left(\frac{n+1}{4}\right)\left[\frac{n+5}{8}\right]}\lambda_8+\sum_{k=1}^{\left[\frac{n+5}{8}\right]-1}\frac{\lambda_{n-8k+7}}{(k+1)k}\\
&=r\left(\frac{1}{\frac{n-1}{2}}+\frac{\frac{n-3}{2}}{\left(\frac{n+1}{2}\right)\frac{n-1}{2}}\frac{1}{r}\lambda_4+\frac{\frac{n+1}{4}-\left[\frac{n+5}{8}\right]}{\left(\frac{n+1}{4}\right)\left[\frac{n+5}{8}\right]}\frac{1}{r}\lambda_8+\sum_{k=1}^{\left[\frac{n+5}{8}\right]-1}\frac{1}{r}\frac{\lambda_{n-8k+7}}{(k+1)k}\right),
\end{aligned}
$$

where $r=\frac{(n+1)\lambda_2+(n-1)}{2n}$. A virtually identical proof to that used in Theorem 4.3.5 for the same claim, shows that $0<r\leq 1$. Next, in order to apply Lemma 4.2.4, we shall show as earlier that $\frac{\lambda_n}{r}\geq -1$ by contradiction. Suppose that $\frac{\lambda_n}{r}<-1$, that is, $2n\lambda_n+(n+1)\lambda_2+(n-1)<0$. It follows that $\frac{1}{n}+\frac{n+1}{n(n-1)}\lambda_2+\frac{2}{n-1}\lambda_n<0$, and hence using (4.10), we obtain

$$
\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{2}{n+1}\lambda_4+\frac{\frac{n+1}{4}-\left[\frac{n+5}{8}\right]}{\left(\frac{n+1}{4}\right)\left[\frac{n+5}{8}\right]}\lambda_8+\sum_{k=1}^{\left[\frac{n+5}{8}\right]-1}\frac{\lambda_{n-8k+8}}{(k+1)k}\geq 0>\frac{1}{n}+\frac{n+1}{n(n-1)}\lambda_2+\frac{2}{n-1}\lambda_n,
$$

that is,

$$
\left(\frac{n-1}{n(n+1)}-\frac{n+1}{n(n-1)}\right)\lambda_2+\frac{2}{n+1}\lambda_4+\frac{\frac{n+1}{4}-\left[\frac{n+5}{8}\right]}{\left(\frac{n+1}{4}\right)\left[\frac{n+5}{8}\right]}\lambda_8+\sum_{k=2}^{\left[\frac{n+5}{8}\right]-1}\frac{\lambda_{n-8k+8}}{(k+1)k}+\frac{1}{2}\lambda_n-\frac{2}{n-1}\lambda_n>0.
$$

Since $1\geq\lambda_2\geq\ldots\geq\lambda_n\geq -1$, then we have

$$
\begin{aligned}
0&<\left(\frac{n-1}{n(n+1)}-\frac{n+1}{n(n-1)}+\frac{2}{n+1}+\frac{\frac{n+1}{4}-\left[\frac{n+5}{8}\right]}{\left(\frac{n+1}{4}\right)\left[\frac{n+5}{8}\right]}+\sum_{k=2}^{\left[\frac{n+5}{8}\right]-1}\frac{1}{(k+1)k}\right)\lambda_2+\frac{1}{2}\lambda_n-\frac{2}{n-1}\lambda_n\\
&=\left(\frac{n-1}{n(n+1)}-\frac{n+1}{n(n-1)}+\frac{2}{n+1}+\frac{\frac{n+1}{4}-\left[\frac{n+5}{8}\right]}{\left(\frac{n+1}{4}\right)\left[\frac{n+5}{8}\right]}+\frac{\left[\frac{n+5}{8}\right]-1}{\left[\frac{n+5}{8}\right]}-\frac{1}{2}\right)\lambda_2+\left(\frac{1}{2}-\frac{2}{n-1}\right)\lambda_n\\
&=\frac{n-5}{2(n-1)}(\lambda_2+\lambda_n).
\end{aligned}
$$

It then follows that $0>2n\lambda_n+(n+1)\lambda_2+(n-1)=n(\lambda_2+\lambda_n)+(n-1)(1+\lambda_n)+(\lambda_2+\lambda_n)>0$, and we have a contradiction. In addition, the proof that $\frac{\lambda_3}{r}\leq 1$ can be easily done

as that of Theorem 4.3.5. Therefore, by New condition 1, there exists an $(\frac{n-1}{2}) \times (\frac{n-1}{2})$ symmetric doubly stochastic matrix $\widehat{B}$ with eigenvalues $1, \frac{1}{r}\lambda_4, \frac{1}{r}\lambda_6, ..., \frac{1}{r}\lambda_{n-1}$. Now applying Lemma 4.2.4 with $A = \widehat{A}$ and $B = \widehat{B}$, we obtain an $n \times n$ symmetric doubly stochastic matrix C with eigenvalues $1, \lambda_2, ..., \lambda_n$.

Finally, let $n = 4m+1$. From (4.11), we conclude that

$$
\begin{aligned}
0 &\leq \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{2(n-1)}{(n+1)(n+3)}\lambda_4 + \frac{\frac{n+3}{4} - [\frac{n+7}{8}]}{(\frac{n+3}{4})\,[\frac{n+7}{8}]}\lambda_8 + \sum_{k=1}^{[\frac{n+7}{8}]-1} \frac{\lambda_{n-8k+8}}{(k+1)k}, \\
&\leq \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{2(n-1)}{(n+1)(n+3)}\lambda_3 + \frac{\frac{n+3}{4} - [\frac{n+7}{8}]}{(\frac{n+3}{4})\,[\frac{n+7}{8}]}\lambda_7 + \sum_{k=1}^{[\frac{n+7}{8}]-1} \frac{\lambda_{n-8k+8}}{(k+1)k}, \\
&= \alpha \left(\frac{1}{\frac{n+1}{2}} + \frac{\frac{n+1}{2} - 1}{\frac{n+1}{2}(\frac{n+1}{2}+1)} \frac{1}{\alpha}\lambda_3 + \frac{\frac{n+3}{4} - [\frac{n+7}{8}]}{(\frac{n+3}{4})\,[\frac{n+7}{8}]} \frac{1}{\alpha}\lambda_7 + \sum_{k=1}^{[\frac{n+7}{8}]-1} \frac{1}{\alpha} \frac{\lambda_{n-8k+8}}{(k+1)k} \right),
\end{aligned}
$$

where $\alpha = \frac{(n-1)\lambda_2 + (n+1)}{2n} > 0$. By a similar argument as the one used in the previous case, we can show that $\frac{\lambda_n}{\alpha} \geq -1$ and $\frac{\lambda_3}{\alpha} \leq 1$. On the other hand, by (4.11) and Lemma 4.3.8 we also have

$$
\begin{aligned}
0 &\leq \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{2(n-1)}{(n+1)(n+3)}\lambda_4 + \frac{\frac{n+3}{4} - [\frac{n+7}{8}]}{(\frac{n+3}{4})\,[\frac{n+7}{8}]}\lambda_8 + \sum_{k=1}^{[\frac{n+7}{8}]-1} \frac{\lambda_{n-8k+8}}{(k+1)k}, \\
&\leq \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{4}{(n-1)(n+1)}\lambda_2 + \frac{2}{n-1}\lambda_4 + \frac{\frac{n-1}{4} - [\frac{n+3}{8}]}{(\frac{n-1}{4})\,[\frac{n+3}{8}]}\lambda_8 + \sum_{k=1}^{[\frac{n+3}{8}]-1} \frac{\lambda_{n-8k+7}}{(k+1)k}, \\
&= r \left(\frac{1}{\frac{n-1}{2}} + \frac{1}{\frac{n-1}{2}} \frac{1}{r}\lambda_4 + \frac{\frac{n-1}{4} - [\frac{n+3}{8}]}{(\frac{n-1}{4})\,[\frac{n+3}{8}]} \frac{1}{r}\lambda_8 + \sum_{k=1}^{[\frac{n+3}{8}]-1} \frac{1}{r} \frac{\lambda_{n-8k+7}}{(k+1)k} \right),
\end{aligned}
$$

where $r = \frac{(n+1)\lambda_2 + (n-1)}{2n}$ with $\frac{\lambda_n}{r} \geq -1$ and $\frac{\lambda_3}{r} \leq 1$ (again for this, a similar proof as above can be employed). Thus by New condition 1, there are $(\frac{n+1}{2}) \times (\frac{n+1}{2})$ and $(\frac{n-1}{2}) \times (\frac{n-1}{2})$ symmetric doubly stochastic matrices $\widehat{A}$ and $\widehat{B}$ with eigenvalues $1, \frac{1}{\alpha}\lambda_3, \frac{1}{\alpha}\lambda_5, ..., \frac{1}{\alpha}\lambda_n$ and $1, \frac{1}{r}\lambda_4, \frac{1}{r}\lambda_6, ..., \frac{1}{r}\lambda_{n-1}$ respectively and such that $0 < r \leq 1$ as earlier. Finally, applying Lemma 4.2.4 with $A = \widehat{A}$ and $B = \widehat{B}$, we obtain an $n \times n$ symmetric doubly stochastic matrix C whose eigenvalues are $1, \lambda_2, ..., \lambda_n$. □

Notation. The conditions of Theorem 4.3.9 are referred to as New condition 2.

Observation 4.3.10. *New condition 2 and New condition 1 are independent. For $n = 14$, the list given by*

$$(1,1,1,1,1,1,0,0,0,0,0,0,0,-\frac{2}{3})$$

satisfies New condition 1 and does not satisfy neither New condition 2. However the list

$$(1,0,0,0,0,-\frac{3}{50},-\frac{3}{50},-\frac{3}{35},-\frac{3}{35},-\frac{1}{10},-\frac{1}{10},-\frac{1}{10},-\frac{1}{10},-\frac{1}{10})$$

satisfies New condition 2 and does not satisfy New condition 1

For $n = 13$, the list given by

$$(1,1,1,1,1,0,0,0,0,0,0,0,-\frac{33}{50})$$

satisfies New condition 1 and does not satisfy New condition 2. However the list

$$(1,0,0,0,-\frac{2}{25},-\frac{2}{25},-\frac{2}{25},-\frac{2}{25},-\frac{1}{9},-\frac{1}{9},-\frac{1}{9},-\frac{1}{9},-\frac{1}{9})$$

satisfies New condition 2 and does not satisfy New condition 1.

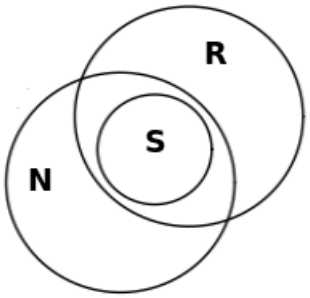
Notation. For convenience, we shall denote by R the sets of all n-tuples σ satisfying New condition 2.

Observation 4.3.11. *For n even, Soules implies New condition 2.*

The proof of the above observation is given in Appendix A.

Observation 4.3.12. *The Venn diagrams of the point sets S, N, and R are the following:*

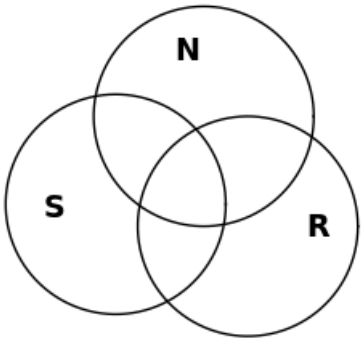
- *for n even:*



We illustrate the above diagram by the following examples:

1. $(1,1,1,1,0,0,0,-\frac{1}{10},-\frac{1}{10},-\frac{1}{10},-\frac{1}{10},-\frac{1}{10},-\frac{1}{10},-\frac{1}{4}) \in S \cap N \cap R$.
2. $(1,1,1,1,0,0,0,0,-\frac{1}{10},-\frac{1}{10},-\frac{1}{10},-\frac{1}{10},-\frac{1}{10},-\frac{23}{50}) \in (N \cap R) \setminus S$.
3. $(1,1,1,1,1,1,0,0,0,0,-\frac{1}{3},-\frac{1}{3},-\frac{1}{2},-\frac{2}{3}) \in N \setminus R$.
4. $(1,0,0,0,-\frac{1}{20},-\frac{3}{50},-\frac{3}{40},-\frac{3}{35},-\frac{3}{35},-\frac{2}{35},-\frac{1}{10},-\frac{1}{10},-\frac{1}{10},-\frac{1}{10}) \in R \setminus N$.

- *for n odd:*



We illustrate the above diagram by the following examples:

1. $(1,1,1,1,1,1,1,\frac{1}{2},\frac{1}{2},\frac{1}{1},\frac{1}{2},-\frac{3}{4},-\frac{3}{4}) \in S \cap N \cap R$.
2. $(1,1,1,0,0,0,0,-\frac{1}{5},-\frac{1}{5},--\frac{1}{5},-\frac{1}{5},-\frac{1}{5},-\frac{1}{5}) \in (S \cap N) \setminus R$.
3. $(1,1,1,0,-\frac{2}{25},-\frac{2}{25},-\frac{2}{25},-\frac{1}{5},-\frac{1}{5},-\frac{1}{5},-\frac{11}{50},-\frac{11}{50}) \in (S \cap R) \setminus N$.
4. $(1,1,1,1,1,1,1,\frac{1}{2},\frac{1}{2},-\frac{3}{4},-\frac{3}{4},-\frac{3}{4},-\frac{3}{4}) \in (N \cap R) \setminus S$.
5. $(1,1,1,0,0,0,0,0,0,0,0,0,-\frac{33}{100}) \in S \setminus (N \cup R)$.
6. $(1,1,1,1,1,0,0,0,0,0,0,0,-\frac{33}{50}) \in N \setminus (S \cup R)$.
7. $(1,0,0,0,-\frac{2}{25},-\frac{2}{25},-\frac{2}{25},-\frac{2}{25},-\frac{1}{9},-\frac{1}{9},-\frac{1}{9},-\frac{1}{9},-\frac{1}{9}) \in R \setminus (S \cup N)$.

We can recursively apply the same process to obtain more general sufficient conditions for the SDIEP based on the Euclidean division of n by 2^k with k positive integer, by substituting constructed doubly stochastic matrices for the case of 2^{k-1} in Lemma 4.2.4. Although, it seems that there is no systematic way of doing this for general n, however we put forward the following conjecture.

Conjecture 4.3.13. *Let k and n be any positive integers such that $n \geq 2^k$, and let $1 \geq \lambda_2 \geq ... \geq \lambda_n \geq -1$. For each $i = 1, ..., k$, denote by α_{i-1} to be the remainder of the Euclidean division of n by 2^{i-1} and define*

$$r_i^{(n)} = \begin{cases} 2^{i-1} - \alpha_{i-1} & \text{if } \alpha_{i-1} \neq 0 \\ 0 & \text{if } \alpha_{i-1} = 0. \end{cases}$$

If

$$\frac{1}{n} + \frac{n - \left[\frac{n+1}{2}\right]}{n\left[\frac{n+1}{2}\right]}\lambda_2 + \frac{\frac{n+r_2^{(n)}}{2} - \left[\frac{n+2+r_2^{(n)}}{4}\right]}{\left(\frac{n+r_2^{(n)}}{2}\right)\left[\frac{n+2+r_2^{(n)}}{4}\right]}\lambda_4 + \frac{\frac{n+r_3^{(n)}}{4} - \left[\frac{n+4+r_3^{(n)}}{8}\right]}{\left(\frac{n+r_3^{(n)}}{4}\right)\left[\frac{n+4+r_3^{(n)}}{8}\right]}\lambda_8 + ...$$

$$... + \frac{\frac{n+r_k^{(n)}}{2^{k-1}} - \left[\frac{n+2^{k-1}+r_k^{(n)}}{2^k}\right]}{\left(\frac{n+r_k^{(n)}}{2^{k-1}}\right)\left[\frac{n+2^{k-1}+r_k^{(n)}}{2^k}\right]}\lambda_{2^k} + \sum_{i=1}^{\left[\frac{n+2^{k-1}+r_k^{(n)}}{2^k}\right]-1} \frac{\lambda_{n-2^k i+2^k}}{(i+1)i} \geq 0,$$

then there exists a $n \times n$ symmetric doubly stochastic matrix with eigenvalues $1, \lambda_2, ..., \lambda_n$.

Note that for $k = 1$, $k = 2$ and $k = 3$, we obtain respectively Soules condition, New condition 1 and New condition 2. In order to support this conjecture, we shall verify it for the cases $n = 26$ and $n = 52$.

Example 17. Let $\sigma = \{1, \lambda_2, ..., \lambda_{26}\}$ be a list of 26 real numbers such that $1 \geq \lambda_2 \geq ... \geq \lambda_{26} \geq -1$. Since $2^5 > 26 > 2^4$, we have in view of the previous conjecture one additional condition to Soules, New condition 1 and New condition 2. We refer to it in this case as *New condition 3*. As for $n = 26$, $r_2 = 0$, $r_3 = 2$, and $r_4 = 6$, New condition 3 states

$$0 \leq \frac{1}{26} + \frac{1}{26}\lambda_2 + \frac{6}{(13)(7)}\lambda_4 + \frac{3}{28}\lambda_8 + \frac{1}{4}\lambda_{16} + \frac{1}{2}\lambda_{26} \tag{4.12}$$

$$= \alpha\left(\frac{1}{13} + \frac{6}{(13)(7)}\frac{1}{\alpha}\lambda_4 + \frac{3}{28}\frac{1}{\alpha}\lambda_8 + \frac{1}{4}\frac{1}{\alpha}\lambda_{16} + \frac{1}{2}\frac{1}{\alpha}\lambda_{26}\right)$$

with $\alpha = \frac{1+\lambda_2}{2}$. By (4.11) of Theorem 4.3.9, there exists a 13×13 symmetric doubly stochastic matrix $\widehat{A}$ with eigenvalues $1, \frac{\lambda_4}{\alpha}, \frac{\lambda_6}{\alpha}, ..., \frac{\lambda_{26}}{\alpha}$. On the other hand, since $1 \geq \lambda_2 \geq ... \geq$

$\lambda_n \geq -1$, we conclude that

$$\frac{1}{13} + \frac{6}{(13)(7)}\frac{1}{\alpha}\lambda_3 + \frac{3}{28}\frac{1}{\alpha}\lambda_7 + \frac{1}{4}\frac{1}{\alpha}\lambda_{15} + \frac{1}{2}\frac{1}{\alpha}\lambda_{25} \geq 0.$$

Hence, there exists a 13×13 symmetric doubly stochastic matrix $\widehat{B}$ with eigenvalues $1, \frac{\lambda_3}{\alpha}, \frac{\lambda_5}{\alpha}, ..., \frac{\lambda_{25}}{\alpha}$. Again applying Lemma 4.2.4 with $A = \widehat{A}$ and $B = \widehat{B}$ and $r = \alpha = \frac{1+\lambda_2}{2}$, we obtain a 26×26 symmetric doubly stochastic matrix C with eigenvalues $1, \lambda_2, ..., \lambda_{26}$.

Note that the list given by

$$(1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,-\frac{1}{13},-\frac{1}{13},-\frac{1}{13},-\frac{1}{13},-\frac{1}{13},-\frac{1}{13},-\frac{1}{13},-\frac{1}{13},-\frac{1}{13})$$

satisfies New condition 3 and does not satisfy New condition 1 nor New condition 2.

With the same manner, we can shall give an example for the case $n = 52$.

Example 18. Let $\sigma = \{1, \lambda_2, ..., \lambda_{52}\}$ be a list of 52 real numbers such that $1 \geq \lambda_2 \geq ... \geq \lambda_{52} \geq -1$. Since $2^6 > 52 > 2^5$, we have two additional conditions to Soules, New condition 1 and New condition 2. We will refer to them in this case as *New condition 3* and *New condition 4*. New condition 3 states

$$\begin{aligned} 0 &\leq \frac{1}{52} + \frac{1}{52}\lambda_2 + \frac{1}{26}\lambda_4 + \frac{6}{(13)(7)}\lambda_8 + \frac{3}{28}\lambda_{16} + \frac{1}{12}\lambda_{20} + \frac{1}{6}\lambda_{36} + \frac{1}{2}\lambda_{52} \\ &= \alpha\left(\frac{1}{26} + \frac{1}{26}\frac{1}{\alpha}\lambda_4 + \frac{6}{(13)(7)}\frac{1}{\alpha}\lambda_8 + \frac{3}{28}\frac{1}{\alpha}\lambda_{16} + \frac{1}{12}\frac{1}{\alpha}\lambda_{20} + \frac{1}{6}\frac{1}{\alpha}\lambda_{36} + \frac{1}{2}\frac{1}{\alpha}\lambda_{52}\right) \end{aligned}$$

with $\alpha = \frac{1+\lambda_2}{2}$. By (4.9) of Theorem 4.3.9, there exists a 26×26 symmetric doubly stochastic matrix $\widehat{A}$ with eigenvalues $1, \frac{\lambda_4}{\alpha}, \frac{\lambda_6}{\alpha}, ..., \frac{\lambda_{52}}{\alpha}$. On the other hand, since $1 \geq \lambda_2 \geq ... \geq \lambda_{52} \geq -1$, we conclude that

$$\frac{1}{26} + \frac{1}{26}\frac{1}{\alpha}\lambda_3 + \frac{6}{(13)(7)}\frac{1}{\alpha}\lambda_7 + \frac{3}{28}\frac{1}{\alpha}\lambda_{15} + \frac{1}{12}\frac{1}{\alpha}\lambda_{19} + \frac{1}{6}\frac{1}{\alpha}\lambda_{35} + \frac{1}{2}\frac{1}{\alpha}\lambda_{51} \geq 0.$$

Hence, there exists a 26×26 symmetric doubly stochastic matrix $\widehat{B}$ with eigenvalues $1, \frac{\lambda_3}{\alpha}, \frac{\lambda_5}{\alpha}, ..., \frac{\lambda_{51}}{\alpha}$. Again applying Lemma 4.2.4 with $A = \widehat{A}$ and $B = \widehat{B}$ and $r = \alpha = \frac{1+\lambda_2}{2}$, we obtain a 52×52 symmetric doubly stochastic matrix C with eigenvalues $1, \lambda_2, ..., \lambda_{52}$.

On the other hand, New condition 4 states

$$\begin{aligned} 0 &\leq \frac{1}{52}+\frac{1}{52}\lambda_2+\frac{1}{26}\lambda_4+\frac{6}{(13)(7)}\lambda_8+\frac{3}{28}\lambda_{16}+\frac{1}{4}\lambda_{32}+\frac{1}{2}\lambda_{52} \\ &= \alpha\left(\frac{1}{26}+\frac{1}{26}\frac{1}{\alpha}\lambda_4+\frac{6}{(13)(7)}\frac{1}{\alpha}\lambda_8+\frac{3}{28}\frac{1}{\alpha}\lambda_{16}+\frac{1}{4}\frac{1}{\alpha}\lambda_{32}+\frac{1}{2}\frac{1}{\alpha}\lambda_{52}\right), \end{aligned}$$

with $\alpha=\frac{1+\lambda_2}{2}$. By (4.12) of the previous example, there exist a 26×26 symmetric doubly stochastic matrix $\widehat{A}$ with eigenvalues $1,\frac{\lambda_4}{\alpha},\frac{\lambda_6}{\alpha},...,\frac{\lambda_{52}}{\alpha}$ and a 26×26 symmetric doubly stochastic matrix $\widehat{B}$ with eigenvalues $1,\frac{\lambda_3}{\alpha},\frac{\lambda_5}{\alpha},...,\frac{\lambda_{51}}{\alpha}$. Again applying Lemma 4.2.4 with $A=\widehat{A}$ and $B=\widehat{B}$ and $r=\alpha=\frac{1+\lambda_2}{2}$, we obtain a 52×52 symmetric doubly stochastic matrix C with eigenvalues $1,\lambda_2,...,\lambda_{52}$.

Chapter 5

On the realizability of normalized Suleimanova spectra

5.1 Introduction

In this chapter, we will study SDIEP for a particular type of spectra. We begin by the following definition.

Definition 5.1.1. We call $\sigma = \{\lambda_1, \lambda_2, ..., \lambda_n\} \subset \mathbb{R}$ a Suleimanova spectrum if $s_1(\sigma) \geq 0$ and σ contains exatly one positive value.

In [123], Suleimanova stated and loosely proved that every such spectrum is realizable. Fiedler [29] proved that every Suleimanova spectrum is symmetrically realisable. In [52], Johnson and Paparella provide a constructive version of Fiedler's result for Hadamard orders. In [92], Paparella provide a constructive version of Suleimanova's result via permutative matrices.

A natural variation of Definition 5.1.1 is the following.

Definition 5.1.2. We call $\sigma = \{1, \lambda_2, ..., \lambda_n\} \subset \mathbb{R}$ a *normalized Suleimanova spectrum* if $1 + \lambda_2 + ... + \lambda_n \geq 0$ and $0 \geq \lambda_i \geq -1$ for all $i = 2, ..., n$.

In [52], the following question was posed.

Problem 3. *If σ is a normalized Suleimanova spectrum, is σ realisable by a doubly stochastic matrix?*

The authors in [52] proved that the answer is yes for all Hadamard orders (i.e. $n = 2^k$) and in this case the realizing matrix is symmetric.

We will denote by NS-SDIEP to be the problem which asks what sets of normalized Suleimanova spectrum occur as the spectrum of a symmetric doubly stochastic matrix.

This chapter is organised as follows. In Section 5.2 we will give a negative answer to the preceding problem when n is odd. In Section 5.3, we will use the recursive method introduced in Chapter 5 to give families of sufficient conditions for NS-SDIEP that improves Soules condition for the realizability of normalized Suleimanova spectra. Finally, in Section 5.4, we study the inclusion relations between the elements of each sequence and show how these new sufficient conditions can be of a big importance.

It is worthy to mention here that previous theorems and lemmas mentioned in Chapter 5, in particular Theorem 4.2.2, Lemmas 4.2.3 and 4.2.4, are needed for our purposes. Moreover, we shall denote as previously mentioned, by PM, S, N the sets of all n-tuples σ satisfying Perfect and Mirsky's condition, Soules condition and New condition 1, respectively.

5.2 NS-SDIEP for n odd

In this section, we will prove that not all normalized Suleimanova spectrum are realizable by doubly stochastic matrices. We begin with a lemma whose proof can be found in [77].

Lemma 5.2.1. *Let A be an $n \times n$ indecomposable doubly stochastic matrix. If A has exactly r roots of unit modulus, then these are the r-roots of unity. If $r > 1$, then r is a divisor of n. Moreover, A is cogredient to a matrix of the form*

$$\begin{pmatrix} 0 & A_1 & 0 & \dots & 0 \\ 0 & 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & A_{r-1} \\ A_r & 0 & 0 & \dots & 0 \end{pmatrix}, \tag{5.1}$$

where the A_i are doubly stochastic of order $\frac{n}{r} \times \frac{n}{r}$, $i = 1, \dots, r$.

Making use of the above lemma we have the following:

Proposition 5.2.2. *If n is odd and $\lambda = (1, \lambda_2, ..., \lambda_{n-1}, -1)$ with $|\lambda_i| < 1$ for all $i = 2, ..., n-1$, then λ cannot be the spectrum of any $n \times n$ doubly stochastic matrix.*

Proof. Suppose that $\lambda = (1, \lambda_2, \dots, \lambda_{n-1}, -1)$ is the spectrum of an $n \times n$ doubly stochastic matrix A. Now as A has 2 eigenvalues of unit modulus which are 1 and -1, then by virtue of the preceding lemma, 2 must be a divisor of n which is a contradiction as n is odd. □

A direct conclusion is the following corollary that gives a negative answer for Problem 3 in the case when n is odd.

Corollary 5.2.3. *If n is odd, then $\lambda = (1,0,...,0,-1)$ cannot be the spectrum of any $n \times n$ doubly stochastic matrix.*

It should be noted here that for n odd, $\lambda = (1,0,...,0,-1)$ is not the only normalized Suleimanova spectrum that is not realizable by an $n \times n$ doubly stochastic matrix. Indeed, for $n = 3$ there exists an infinite number of such points. To see this, we first recall the following result from Perfect and Mirsky[97].

Theorem 5.2.4. *[97] Let $1 \geq \lambda_2 \geq \lambda_3 \geq -1$, and $1+\lambda_2+\lambda_3 \geq 0$. There exists a 3×3 doubly stochastic matrix with spectrum $(1,\lambda_2,\lambda_3)$ if and only if $2+\lambda_2+3\lambda_3 \geq 0$.*

As a conclusion, we have the following.

Corollary 5.2.5. *Let $1 \geq \lambda_2 \geq \lambda_3 \geq -1$, and $1+\lambda_2+\lambda_3 \geq 0$. There exists a 3×3 doubly stochastic matrix with spectrum $(1,\lambda_2,\lambda_3)$ if and only if $(1,\lambda_2,\lambda_3)$ is in*

$$Convhull[(1,1,1),(1,1,-1),(1,-1/2,-1/2)],$$

where Convhull stands for convex hull.

Thus we have the following conclusion.

Corollary 5.2.6. *The region of $\mathbb{R}^3$ that contains all decreasingly ordered normalized Suleimanova spectra that are realizable by doubly stochastic matrices, is*

$$Convhull[(1,-1/2,-1/2),(1,0,0),(1,0,-2/3)],$$

and the region of all decreasingly ordered normalized Suleimanova spectra that are not realizable by doubly stochastic matrices, is

$$Convhull[(1,-1/2,-1/2),(1,0,-1),(1,0,-2/3)] \setminus [(1,-1/2,-1/2),(1,0,-2/3)],$$

where $[(1,-1/2,-1/2),(1,0,-2/3)]$ is the line-segment joining $(1,-1/2,-1/2)$ to $(1,0,-2/3)$.

The following figure present the decreasingly ordered normalized Suleimanova spectra of dimension 3. We presented by a 2D plot since 1 is always an eigenvalue.

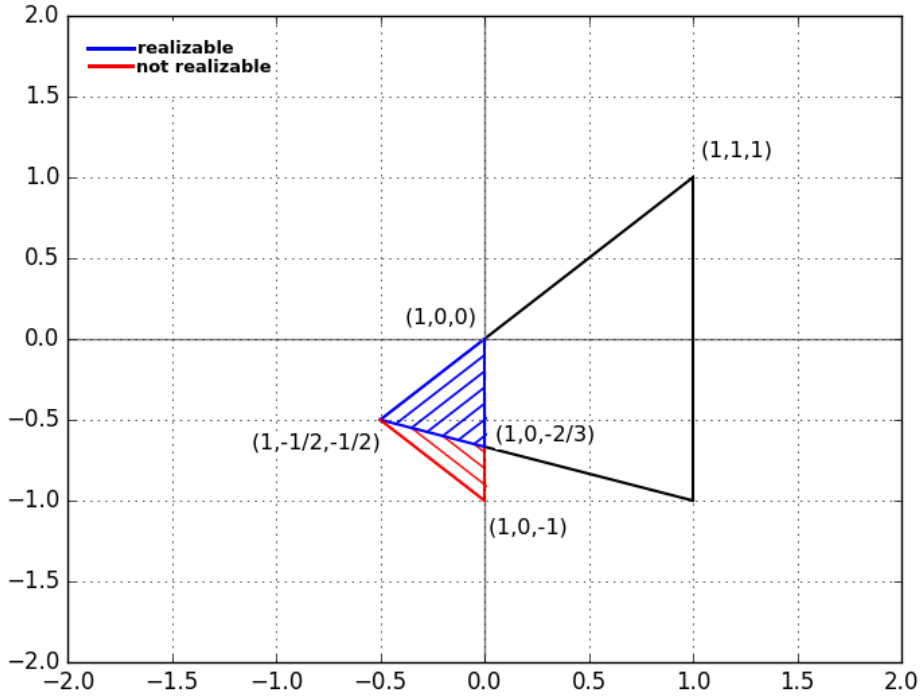


Fig. 5.1 The decreasingly ordered normalized Suleimanova spectra

5.3 Sufficient conditions

Here we will derive new sufficient conditions for a normalized Suleimanova spectrum to be the spectrum of a symmetric doubly stochastic matrix. We shall start with the following definition that can be found in [139, Chapter 5]

Definition 5.3.1. An $n \times n$ matrix $H = (h_{ij})$ is a Hadamard matrix if $h_{ij} \in \{\pm 1\}$ and $HH^T = nI_n$. If n is a positive integer such that an $n \times n$ Hadamard matrix exists, then n is said to be a Hadamard order.

Let $H_0 = (1)$, $H_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, and for $n \in \mathbb{N}\backslash\{0,1\}$ (where $\mathbb{N}$ is the set of natural numbers), define

$$H_n := H_1 \otimes H_{n-1} = \begin{pmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{pmatrix} \in M_{2^n}(\mathbb{R}).$$

It is well-known that H_n is a Hadamard matrix for every $n \in \mathbb{N}$, and the matrix H_n obtained from the previous construction is known as the Walsh matrix of order 2^n. Note that Walsh matrices satisfy the following additional well-known properties (see [139, Chapter 5]):

- $H_n^T = H_n$;
- $H_n^{-1} = \frac{1}{2^n} H_n$.

Next, we recall the following result which is due to Johnson and Paparella [52].

Theorem 5.3.2. *[52] For any positive integer k and for any $n = 2^k$, let $\lambda = (1, \lambda_2, \lambda_3, ..., \lambda_n)$ be a list of n real numbers with*

$$1 \geqslant \lambda_2 \geqslant \lambda_3 \geqslant ... \geqslant \lambda_n \geqslant -1,$$

$$\lambda_i \leq 0 \quad for \quad i = 2, ..., n,$$

and

$$1 + \sum_{i=2}^{n} \lambda_i \geq 0.$$

Then there exists an $n \times n$ symmetric doubly stochastic matrix with spectrum λ.

It is worth mentioning here that the proof of preceding theorem is constructive. Indeed, with the same notation as above, the authors prove that $H_k \mathrm{diag}(1, \lambda_2, \lambda_3, ..., \lambda_n) H_k^{-1}$ is actually an $n \times n$ symmetric doubly stochastic matrix with spectrum $(1, \lambda_2, \lambda_3, ..., \lambda_n)$.

Our next objective is to exploit the recursive method described earlier in order to find new families of sufficient conditions for NS-SDIEP that improve Soules condition, and New condition 1 for the case of normalized Suleimanova spectra.

Theorem 5.3.3. *Let p and n be two positive integers such that $p \geq 2$ and $n \geq 2^p$, and let $0 \geq \lambda_2 \geq ... \geq \lambda_n \geq -1$. If*

$$\frac{1}{n} + \frac{1}{n(n-1)} \lambda_2 + ... + \frac{1}{2^p(2^p+1)} \lambda_{n-2^p+1} + \frac{1}{2^p} (\lambda_{n-2^p+2} + ... + \lambda_{n-1} + \lambda_n) \geq 0, \quad (5.2)$$

then there is an $n \times n$ symmetric doubly stochastic matrix with normalized Suleimanova spectrum $1, \lambda_2, ..., \lambda_n$.

Proof. We will proceed by induction. For $n = 2^p$, this is true by Theorem 5.3.2. Let $n > 2^p$ and suppose that the assertion is true for $n-1$. Since

$$\begin{aligned}
0 &\leq \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + \ldots + \frac{1}{2^p(2^p+1)}\lambda_{n-2^p+1} + \frac{1}{2^p}\left(\lambda_{n-2^p+2} + \ldots + \lambda_{n-1} + \lambda_n\right) \\
&= \frac{n-1+\lambda_2}{n(n-1)} + \frac{1}{(n-1)(n-2)}\lambda_3 + \ldots + \frac{1}{2^p(2^p+1)}\lambda_{n-2^p+1} + \frac{1}{2^p}\left(\lambda_{n-2^p+2} + \ldots + \lambda_{n-1} + \lambda_n\right) \\
&= \theta\left(\frac{1}{n-1} + \frac{1}{(n-1)(n-2)}\frac{\lambda_3}{\theta} + \ldots + \frac{1}{2^p(2^p+1)}\frac{\lambda_{n-2^p+1}}{\theta} + \frac{1}{2^p}\left(\frac{\lambda_{n-2^p+2}}{\theta} + \ldots + \frac{\lambda_{n-1}}{\theta} + \frac{\lambda_n}{\theta}\right)\right)
\end{aligned}$$

with $\theta = \frac{n-1+\lambda_2}{n} > 0$, and consequently $\frac{\lambda_i}{\theta} \leq 0$ for $i = 3, \ldots, n$. Next, we show that $\frac{\lambda_i}{\theta} \geq -1$ by contradiction. Suppose that $\frac{\lambda_n}{\theta} < -1$, then $\frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + \frac{1}{n-1}\lambda_n < 0$. In view of (5.2), we conclude that

$$\begin{aligned}
&\frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + \ldots + \frac{1}{2^p(2^p+1)}\lambda_{n-2^p+1} + \frac{1}{2^p}(\lambda_{n-2^p+2} + \ldots + \lambda_{n-1} + \lambda_n) \geq 0 \\
&> \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + \frac{1}{n-1}\lambda_n,
\end{aligned}$$

that is,

$$\frac{1}{(n-1)(n-2)}\lambda_3 + \ldots + \frac{1}{2^p(2^p+1)}\lambda_{n-2^p+1} + \frac{1}{2^p}(\lambda_{n-2^p+2} + \ldots + \lambda_{n-1}) + \left(\frac{1}{2^p} - \frac{1}{n-1}\right)\lambda_n > 0.$$

As $n > 2^p$ and all the $\lambda_i \leq 0$, we get a contradiction. It follows that,

$$\frac{1}{n-1} + \frac{1}{(n-1)(n-2)}\frac{\lambda_3}{\theta} + \ldots + \frac{1}{2^p(2^p+1)}\frac{\lambda_{n-2^p+1}}{\theta} + \frac{1}{2^p}\left(\frac{\lambda_{n-2^p+2}}{\theta} + \ldots + \frac{\lambda_{n-1}}{\theta} + \frac{\lambda_n}{\theta}\right) \geq 0.$$

Therefore, by the induction hypothesis, there exists an $(n-1)\times(n-1)$ symmetric doubly stochastic matrix $\widehat{A}$ with eigenvalues $(1, \frac{\lambda_3}{\theta}, \ldots, \frac{\lambda_n}{\theta})$. Next, let $r = \frac{n-1}{n}(1-\lambda_2) = 1 - \frac{1+(n-1)\lambda_2}{n}$. In order to apply Lemma 4.2.3, we shall prove that $0 \leq r \leq 1$. First observe that $n \geq 1 + (n-1)\lambda_2$. Moreover, using (5.2), we can write

$$\begin{aligned}
0 &\leq \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + \ldots + \frac{1}{2^p(2^p+1)}\lambda_{n-2^p+1} + \frac{1}{2^p}(\lambda_{n-2^p+2} + \ldots + \lambda_{n-1} + \lambda_n) \\
&\leq \frac{1}{n} + \left(\frac{1}{n(n-1)} + \ldots + \frac{1}{2^p(2^p+1)} + \frac{2^p-1}{2^p}\right)\lambda_2 \\
&= \frac{1}{n} + \frac{n-1}{n}\lambda_2.
\end{aligned}$$

Hence $1+(n-1)\lambda_2 \geq 0$ and therefore $0 \leq r \leq 1$. Now applying Lemma 4.2.3 with this r and with $A = \widehat{A}$, we obtain an $n \times n$ symmetric doubly stochastic matrix C with eigenvalues $1, \lambda_2, ..., \lambda_n$. □

Remark 14. For $p = 2$, and $n > 4$, Theorem 5.3.3 states that if $0 \geq \lambda_2 \geq ... \geq \lambda_n \geq -1$ and

$$\frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{(5)(4)}\lambda_{n-3} + \frac{1}{4}(\lambda_{n-2} + \lambda_{n-1} + \lambda_n) \geq 0,$$

then there is an $n \times n$ symmetric doubly stochastic matrix with eigenvalues $1, \lambda_2, ..., \lambda_n$. It is worthy to mention here that we can obtain this sufficient condition by taking the matrix $\frac{1}{\sqrt{4}}H_2 = [\frac{1}{\sqrt{4}}e_4|x_2|x_3|x_4]$, then constructing an orthogonal matrix W_n of any order $n \geq 4$. We construct W_n by considering the first $n-3$ columns of the Soules matrix V_n and the last 3 columns $\begin{pmatrix} x_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $\begin{pmatrix} x_3 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ and $\begin{pmatrix} x_4 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ to be the n columns of W_n.

A simple matrix multiplication shows that the matrices $A = V_n \Lambda V_n^T$ and $B = W_n \Lambda W_n^T$, differ only by the 4×4 principal submatrix C formed from the first 4 columns and the first 4 rows of A and B. So that a virtual identical proof to that used in [122] shows that all the entries of the symmetric matrix $B = W_n \Lambda W_n^T$ are nonnegative except for the diagonal entries and for the entries of the matrix C which are:

$$c_{11} = c_{22} = c_{33} = c_{44} = \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{20}\lambda_{n-3} + \frac{1}{4}\lambda_{n-2} + \frac{1}{4}\lambda_{n-1} + \frac{1}{4}\lambda_{n-1},$$

$$c_{12} = c_{21} = c_{34} = c_{43} = \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{20}\lambda_{n-3} - \frac{1}{4}\lambda_{n-2} - \frac{1}{4}\lambda_{n-1} + \frac{1}{4}\lambda_{n-1},$$

$$c_{13} = c_{31} = c_{24} = c_{42} = \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{20}\lambda_{n-3} - \frac{1}{4}\lambda_{n-2} + \frac{1}{4}\lambda_{n-1} - \frac{1}{4}\lambda_{n-1},$$

and,

$$c_{14} = c_{41} = c_{32} = c_{23} = \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{20}\lambda_{n-3} + \frac{1}{4}\lambda_{n-2} - \frac{1}{4}\lambda_{n-1} - \frac{1}{4}\lambda_{n-1}.$$

The remaining diagonal entries $b_{55},...,b_{nn}$ are increasing and the smallest one is

$$\begin{aligned} b_{55} &= \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + \frac{1}{(n-1)(n-2)}\lambda_3 + ... + \frac{1}{30}\lambda_{n-4} + \frac{16}{20}\lambda_{n-3} \\ &= \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + \frac{1}{(n-1)(n-2)}\lambda_3 + ... + \frac{1}{30}\lambda_{n-4} + \frac{1}{4}\lambda_{n-3} + \frac{1}{4}\lambda_{n-3} + \frac{1}{4}\lambda_{n-3}. \end{aligned}$$

As $1 \geq \lambda_2 \geq .... \geq \lambda_n \geq -1$,, then clearly $c_{11} \geq b_{55}$. Moreover since $0 \geq \lambda_2 \geq .... \geq \lambda_n \geq -1$, then it is easy to see that $c_{11} \leq c_{12}$, $c_{11} \leq c_{13}$ and $c_{11} \leq c_{14}$.

We note that using the recursive method is more easy then to verify the positiveness of the entries especially when we consider bigger values of p and n.

Next, we give a simple example that shows that the method used in Theorem 5.3.3 and Lemma 4.2.3 is also constructive, i e. we can construct a realizable matrix.

Example 19. Let $\sigma = \{1, 0, -\frac{1}{10}, -\frac{1}{5}, -\frac{2}{5}\}$ be a list of 5 real numbers. It is easy to check that σ verifies (5.2). It follows that it is the spectrum of 5×5 symmetric doubly stochastic matrix C. By (5.2), we have

$$\begin{aligned} 0 &\leq \frac{1}{5} + \frac{1}{20}\lambda_2 + \frac{1}{4}(\lambda_3 + \lambda_4 + \lambda_5) \\ &= \theta\left(\frac{1}{4} + \frac{1}{4}(\frac{1}{\theta}\lambda_3 + \frac{1}{\theta}\lambda_4 + \frac{1}{\theta}\lambda_5)\right) \end{aligned}$$

with $\theta = \frac{4}{5}$. Therefore, by Theorem 5.3.2, there exists a symmetric doubly stochastic matrix A with eigenvalues $1, \frac{\lambda_3}{\theta} = -\frac{1}{8}, \frac{\lambda_4}{\theta} = -\frac{1}{4}, \frac{\lambda_5}{\theta} - \frac{1}{2}$. By Theorem 5.3.2, A can be obtained by

$$A = H_2 \text{diag}\left(1, -\frac{1}{8}, -\frac{1}{4}, -\frac{1}{2}\right) H_2^{-1},$$

that is,

$$\begin{pmatrix} \frac{1}{32} & \frac{11}{32} & \frac{13}{32} & \frac{7}{32} \\ \frac{11}{32} & \frac{1}{32} & \frac{7}{32} & \frac{13}{32} \\ \frac{13}{32} & \frac{7}{32} & \frac{1}{32} & \frac{11}{32} \\ \frac{7}{32} & \frac{13}{32} & \frac{11}{32} & \frac{1}{32} \end{pmatrix}.$$

Therefore, by Lemma 4.2.3, $C = \begin{pmatrix} 1-r & \frac{r}{2}e_4^T \\ \frac{r}{2}e_4^T & \frac{4-r}{4}A \end{pmatrix}$, with $r = \frac{4}{5}$. It follows that

$$C = \begin{pmatrix} \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{40} & \frac{11}{40} & \frac{7}{40} & \frac{13}{40} \\ \frac{1}{5} & \frac{11}{40} & \frac{1}{40} & \frac{13}{40} & \frac{7}{40} \\ \frac{1}{5} & \frac{7}{40} & \frac{13}{40} & \frac{1}{40} & \frac{11}{40} \\ \frac{1}{5} & \frac{13}{40} & \frac{7}{40} & \frac{11}{40} & \frac{1}{40} \end{pmatrix}$$

is doubly stochastic with spectrum $(1, 0, -\frac{1}{10}, -\frac{1}{5}, -\frac{2}{5})$.

Now substituting in Lemma 4.2.4, the doubly stochastic matrices obtained from Theorem 5.3.3, we arrive at the following theorem.

Theorem 5.3.4. *Let p and n be two positive integers such that $p \geq 2$ and $n \geq 2^{p+1}$, and let $0 \geq \lambda_2 \geq ... \geq \lambda_n \geq -1$. If*

$$\frac{1}{n} + \frac{n-k-1}{n(k+1)}\lambda_2 + \sum_{i=2^p}^{k} \frac{\lambda_{n-2i+2}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1} \lambda_{n-2i+2} \geq 0 \tag{5.3}$$

holds with $n = 2k+2$ for n even (with the convention that $\sum_{i=2^p}^{k} \frac{\lambda_{n-2i+2}}{(i+1)i} = 0$ for $n = 2^{p+1}$), and $n = 2k+1$ for n odd, then there exists an $n \times n$ symmetric doubly stochastic matrix with eigenvalues $1, \lambda_2, ..., \lambda_n$.

Proof. Case 1: For $n = 2k+2$, Inequality (5.3) can be rewritten as

$$\begin{aligned} 0 &\leq \frac{1}{n} + \frac{1}{n}\lambda_2 + \sum_{i=2^p}^{k} \frac{\lambda_{n-2i+2}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1} \lambda_{n-2i+2} \\ &= \alpha\left(\frac{1}{k+1} + \sum_{i=2^p}^{k} \frac{1}{\alpha}\frac{\lambda_{n-2i+2}}{(i+1)i} + \frac{1}{2^p}\frac{1}{\alpha}\sum_{i=1}^{2^p-1} \lambda_{n-2i+2}\right) \end{aligned}$$

with $\alpha = \frac{1+\lambda_2}{2} \geq 0$. We show by contradiction that $\alpha > 0$. Suppose that $\alpha = 0$. Then, $\lambda_2 = -1$, which implies that $\lambda_3 = ... = \lambda_n = -1$. It follows that

$$\frac{1}{n} + \frac{1}{n}\lambda_2 + \sum_{i=2^p}^{k} \frac{\lambda_{n-2i+2}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1} \lambda_{n-2i+2} < 0,$$

and this contradicts (5.3). As before, we show next that $\frac{\lambda_n}{\alpha} \geq -1$ by contradiction. Suppose that $\frac{\lambda_n}{\alpha} < -1$, that is, $2\lambda_n + \lambda_2 + 1 < 0$ so that $\frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{2}{n}\lambda_n < 0$. From condition (5.3), we conclude that

$$\frac{1}{n} + \frac{1}{n}\lambda_2 + \sum_{i=2^p}^{k} \frac{\lambda_{n-2i+2}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1} \lambda_{n-2i+2} \geq 0 > \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{2}{n}\lambda_n,$$

that is,

$$\sum_{i=2^p}^{k} \frac{\lambda_{n-2i+2}}{(i+1)i} + \frac{1}{2^p}\sum_{i=2}^{2^p-1} \lambda_{n-2i+2} + \left(\frac{1}{2^p} - \frac{2}{n}\right)\lambda_n > 0.$$

Again as $n \geq 2^{p+1}$ and all $\lambda_i \leq 0$, we have a contradiction. Hence by Theorem 5.3.3, there exists a $(k+1)\times(k+1)$ symmetric doubly stochastic matrix $\widehat{A}$ with eigenvalues $1, \frac{1}{\alpha}\lambda_4, ..., \frac{1}{\alpha}\lambda_{2k+2}$. Moreover, since $1 \geq \lambda_2 \geq ... \geq \lambda_n \geq -1$, then

$$\frac{1}{k+1} + \sum_{i=2^p}^{k} \frac{1}{\alpha}\frac{\lambda_{n-2i+1}}{(i+1)i} + \frac{1}{2^p}\frac{1}{\alpha}\sum_{i=1}^{2^p-1} \lambda_{n-2i+1} \geq 0,$$

and therefore, there exists a $(k+1)\times(k+1)$ symmetric doubly stochastic matrix $\widehat{B}$ with eigenvalues $1, \frac{1}{\alpha}\lambda_3, ..., \frac{1}{\alpha}\lambda_{2k+1}$. Applying now Lemma 4.2.4, with $A = \widehat{A}$, $B = \widehat{B}$ and $\alpha = r$, we obtain an $n \times n$ symmetric doubly stochastic matrix C with eigenvalues $1, \lambda_2, ..., \lambda_n$.
Case 2: For $n = 2k+1$, we have

$$\begin{aligned} 0 &\leq \frac{1}{2k+1} + \frac{k}{(2k+1)(k+1)}\lambda_2 + \sum_{i=2^p}^{k} \frac{\lambda_{n-2i+2}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1} \lambda_{n-2i+2} \\ &= \alpha\left(\frac{1}{k+1} + \sum_{i=2^p}^{k} \frac{1}{\alpha}\frac{\lambda_{n-2i+2}}{(i+1)i} + \frac{1}{2^p}\frac{1}{\alpha}\sum_{i=1}^{2^p-1} \lambda_{n-2i+2}\right) \end{aligned}$$

with $\alpha = \frac{(k+1)+k\lambda_2}{2k+1} = \frac{(n+1)+(n-1)\lambda_2}{2n} = \frac{2+(n-1)(1+\lambda_2)}{2n} > 0$. Again, at this point we shall show by contradiction that $\frac{\lambda_n}{\alpha} \geq -1$. Suppose that $\frac{\lambda_n}{\alpha} < -1$, that is, $2n\lambda_n + (n-1)\lambda_2 + (n+1) < 0$. It follows that, $\frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{2}{n+1}\lambda_n < 0$, and hence by (5.3), we conclude that

$$\frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \sum_{i=2^p}^{k} \frac{\lambda_{n-2i+2}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1} \lambda_{n-2i+2} \geq 0 > \frac{1}{n} + \frac{n-1}{n}\lambda_2 + \frac{2}{n+1}\lambda_n,$$

that is,

$$\sum_{i=2^p}^{k} \frac{\lambda_{n-2i+2}}{(i+1)i} + \frac{1}{2^p}\sum_{i=2}^{2^p-1} \lambda_{n-2i+2} + \left(\frac{1}{2^p} - \frac{2}{n+1}\right)\lambda_n > 0.$$

As $n \geq 2^{p+1}$ and all the $\lambda_i \leq 0$, this leads to a contradiction. Hence, by Theorem 5.3.3, there exists a $(k+1)\times(k+1)$ symmetric doubly stochastic matrix with eigenvalues $1, \frac{1}{\alpha}\lambda_3, \ldots, \frac{1}{\alpha}\lambda_{2k+1}$. On the other hand, we know that

$$\begin{aligned}
0 &\leq \frac{1}{2k+1} + \frac{k}{(2k+1)(k+1)}\lambda_2 + \sum_{i=2^p}^{k} \frac{\lambda_{n-2i+2}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1} \lambda_{n-2i+2} \\
&\leq \frac{1}{2k+1} + \frac{k}{(2k+1)(k+1)}\lambda_2 + \frac{1}{k(k+1)}\lambda_2 + \sum_{i=2^p}^{k-1} \frac{\lambda_{n-2i+1}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1} \lambda_{n-2i+1} \\
&= r\left(\frac{1}{k} + \sum_{i=2^p}^{k-1} \frac{1}{r}\frac{\lambda_{n-2i+1}}{(i+1)i} + \frac{1}{2^p}\frac{1}{r}\sum_{i=1}^{2^p-1} \lambda_{n-2i+1}\right)
\end{aligned}$$

where $r = \frac{(k+1)\lambda_2+k}{2k+1}$. In order to apply Lemma 4.2.4, we need to show that $0 < r \leq 1$. Now since r can be rewritten as $r = \frac{(k+1)\lambda_2+k}{2k+1} = \frac{(n+1)\lambda_2+(n-1)}{2n} = \frac{(n-1)\lambda_2+1+2\lambda_2+n-2}{2n}$, then it is enough to show that $1+(n-1)\lambda_2 \geq 0$ (as $n > 4$). From (5.3), we conclude that

$$\begin{aligned}
0 &\leq \frac{1}{2k+1} + \frac{k}{(2k+1)(k+1)}\lambda_2 + \sum_{i=2^p}^{k} \frac{\lambda_{n-2i+2}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1} \lambda_{n-2i+2} \\
&\leq \frac{1}{2k+1} + \left(\frac{k}{(2k+1)(k+1)} + \sum_{i=2^p}^{k} \frac{1}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1} 1\right)\lambda_2 \\
&= \frac{1}{2k+1} + \left(\frac{k}{(2k+1)(k+1)} + \frac{k}{k+1} - \frac{2^p-1}{2^p} + \frac{2^p-1}{2^p}\right)\lambda_2 \\
&= \frac{1}{2k+1} + \frac{2k}{2k+1}\lambda_2 \\
&= \frac{1}{n} + \frac{n-1}{n}\lambda_2.
\end{aligned}$$

Hence $1+(n-1)\lambda_2 \geq 0$ and so $1+(n-1)\lambda_2+2\lambda_2+n-2 \geq 1+(n-1)\lambda_2-2+n-2 > 0$. Hence, $r > 0$. Also, as $(n+1)\lambda_2+(n-1) \leq n+1+n-1 = 2n$, then $r \leq 1$.

As earlier, we next show by contradiction that $\frac{\lambda_n}{r} \geq -1$. Suppose that $\frac{\lambda_n}{r} < -1$, that is, $2n\lambda_n + (n+1)\lambda_2 + (n-1) < 0$. It follows that, $\frac{1}{n} + \frac{n+1}{n(n-1)}\lambda_2 + \frac{2}{n-1}\lambda_n < 0$, and therefore by

(5.3), we conclude that

$$\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\sum_{i=2^p}^{k}\frac{\lambda_{n-2i+2}}{(i+1)i}+\frac{1}{2^p}\sum_{i=1}^{2^p-1}\lambda_{n-2i+2}\geq 0>\frac{1}{n}+\frac{n+1}{n(n-1)}\lambda_2+\frac{2}{n-1}\lambda_n,$$

that is,

$$\left(\frac{n-1}{n(n+1)}-\frac{n+1}{n(n-1)}\right)\lambda_2+\sum_{i=2^p}^{k}\frac{\lambda_{n-2i+2}}{(i+1)i}+\frac{1}{2^p}\sum_{i=2}^{2^p-1}\lambda_{n-2i+2}+\left(\frac{1}{2^p}-\frac{2}{n-1}\right)\lambda_n>0.$$

Since $1\geq\lambda_2\geq ...\geq\lambda_n\geq -1$, then we can write

$$\begin{aligned}
0&<\left(\frac{n-1}{n(n+1)}-\frac{n+1}{n(n-1)}+\sum_{i=2^p}^{k}\frac{1}{(i+1)i}+\frac{1}{2^p}\sum_{i=2}^{2^p-1}1\right)\lambda_2+\left(\frac{1}{2^p}-\frac{2}{n-1}\right)\lambda_n\\
&=\left(\frac{n-1}{n(n+1)}-\frac{n+1}{n(n-1)}+\frac{k}{k+1}-\frac{2^p-1}{2^p}+\frac{2^p-2}{2^p}\right)\lambda_2+\left(\frac{1}{2^p}-\frac{2}{n-1}\right)\lambda_n\\
&=\left(\frac{n-1}{n(n+1)}-\frac{n+1}{n(n-1)}+\frac{n-1}{n+1}-1+\frac{1}{2^p}+1-\frac{2}{2^p}\right)\lambda_2+\left(\frac{1}{2^p}-\frac{2}{n-1}\right)\lambda_n\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\left(\text{since } k=\frac{n-1}{2}\right)\\
&=\left(\frac{n-5}{n-1}+\frac{2}{n-1}-\frac{1}{2^p}\right)\lambda_2+\left(\frac{1}{2^p}-\frac{2}{n-1}\right)\lambda_n\\
&=\frac{n-5}{n-1}\lambda_2+\left(\frac{1}{2^p}-\frac{2}{n-1}\right)(\lambda_n-\lambda_2).
\end{aligned}$$

Since $n=2k+1$ and $n\geq 2^{p+1}$, then obviously $n\geq 2^{p+1}+1$. However, this leads to a contradiction as $\lambda_2\leq 0$ and $(\lambda_n-\lambda_2)\leq 0$. Therefore, by Theorem 5.3.3, there exists a $k\times k$ symmetric doubly stochastic matrix with eigenvalues $1,\frac{1}{r}\lambda_4,...,\frac{1}{r}\lambda_{2k}$. Now the proof can be easily completed by applying Lemma 4.2.4, with $A=\widehat{A}$ and $B=\widehat{B}$. □

Again, substituting in Lemma 4.2.4 the doubly stochastic matrices obtained from Theorem 5.3.4, we obtain the following theorem for which its proof is virtually identical to that of Theorem 5.3.4. So, we will give only a sketch of its proof.

Theorem 5.3.5. *Let p and n be two positive integers such that $p\geq 2$ and $n\geq 2^{p+2}$, and let $0\geq\lambda_2\geq ...\geq\lambda_n\geq -1$. Suppose that*

$$\frac{1}{n}+\frac{1}{n}\lambda_2+\frac{\frac{n}{2}-\left[\frac{n+2}{4}\right]}{\left(\frac{n}{2}\right)\left[\frac{n+2}{4}\right]}\lambda_4+\sum_{i=2^p}^{\left[\frac{n+2}{4}\right]-1}\frac{\lambda_{n-4i+4}}{(i+1)i}+\frac{1}{2^p}\sum_{i=1}^{2^p-1}\lambda_{n-4i+4}\geq 0, \tag{5.4}$$

for n even (with the convention that $\sum_{i=2^p}^{[\frac{n+2}{4}]-1} \frac{\lambda_{n-4i+4}}{(i+1)i} = 0$ *for* $n = 2^{p+2}$*) and*

$$\frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{\frac{n+1}{2} - [\frac{n+3}{4}]}{(\frac{n+1}{2})[\frac{n+3}{4}]}\lambda_4 + \sum_{i=2^p}^{[\frac{n+3}{4}]-1} \frac{\lambda_{n-4i+4}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1} \lambda_{n-4i+4} \geq 0, \tag{5.5}$$

for n odd, hold. Then there exists an $n \times n$ *symmetric doubly stochastic matrix D whose eigenvalues are* $1, \lambda_2, ..., \lambda_n$.

Proof. If n is even, then by (5.4) and Theorem 5.3.4, there are $(\frac{n}{2}) \times (\frac{n}{2})$ symmetric doubly stochastic matrices $\widehat{A}$ and $\widehat{B}$ with eigenvalues $1, \frac{1}{\alpha}\lambda_4, \frac{1}{\alpha}\lambda_6, ..., \frac{1}{\alpha}\lambda_n$ and $1, \frac{1}{\alpha}\lambda_3, \frac{1}{\alpha}\lambda_5, ..., \frac{1}{\alpha}\lambda_{n-1}$ respectively, where $\alpha = \frac{1+\lambda_2}{2}$. Apply now Lemma 2, with $A = \widehat{A}$ and $B = \widehat{B}$, then we obtain an $n \times n$ symmetric doubly stochastic C with eigenvalues $1, \lambda_2, ..., \lambda_n$.
If n is odd, then by (5.5) and Theorem 5.3.4, there are $(\frac{n+1}{2}) \times (\frac{n+1}{2})$ and $(\frac{n-1}{2}) \times (\frac{n-1}{2})$ symmetric doubly stochastic matrices $\widehat{A}$ and $\widehat{B}$ with eigenvalues $1, \frac{1}{\alpha}\lambda_3, \frac{1}{\alpha}\lambda_5, ..., \frac{1}{\alpha}\lambda_n$ and $1, \frac{1}{r}\lambda_4, \frac{1}{r}\lambda_6, ..., \frac{1}{r}\lambda_{n-1}$ respectively with $\alpha = \frac{(n-1)\lambda_2+(n+1)}{2n}$ and $r = \frac{(n+1)\lambda_2+(n-1)}{2n}$. Applying Lemma 4.2.4 with $A = \widehat{A}$ and $B = \widehat{B}$, we obtain an $n \times n$ symmetric doubly stochastic matrix C with eigenvalues $1, \lambda_2, .., \lambda_n$. □

Notation. For convenience, we denote by M_p the condition of Theorem 5.3.3, S_p the condition of Theorem 5.3.4 and N_p the condition of Theorem 5.3.5.

5.4 Sequences of sufficient conditions for NS-SDIEP

In this section, we will prove that the family S_p for $p \geq 2$ is an improvement of Soules condition S for normalized Suleimanova spectra. In addition, we will give the inclusion relations between the elements of M_p, S_p and N_p and we will show by examples how these conditions can be of a great importance for NS-SDIEP.

With the previous notations, we have the following observations.

Observation 5.4.1. *Let* $\sigma = \{1, \lambda_2, ..., \lambda_n\}$ *be a normalized Suleimanova spectrum. If* σ *satisfies PM then* σ *satisfies* M_2*. Conversely, it is not always true. Consequently, with a slight abuse of notation, we conclude that* $PM \subset M_2$.

Proof. Suppose that σ satisfies *PM*, then

$$\begin{aligned}
0 &\leq \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{20}\lambda_{n-3} + \frac{1}{12}\lambda_{n-2} + \frac{1}{6}\lambda_{n-1} + \frac{1}{2.1}\lambda_n \\
&\leq \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{20}\lambda_{n-3} + \frac{1}{12}\lambda_{n-2} + \frac{1}{6}\lambda_{n-2} + \frac{1}{4}\lambda_{n-1} + \frac{1}{4}\lambda_n \\
&= \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{20}\lambda_{n-3} + \frac{1}{4}\lambda_{n-2} + \frac{1}{4}\lambda_{n-1} + \frac{1}{4}\lambda_n
\end{aligned}$$

Hence σ satisfies M_2. Finally, the list $\sigma = \{1, -\frac{1}{20}, -\frac{1}{20}, -\frac{1}{10}, -\frac{1}{10}, -\frac{2}{5}\}$ satisfies M_2, but does not satisfy *PM*. Thus the proof is completed.

□

Observation 5.4.2. *Let $\sigma = \{1, \lambda_2, ..., \lambda_n\}$ be a normalized Suleimanova spectrum with $n \geq 8$. If σ satisfies S then σ satisfies S_2. Conversely, it is not always true. Consequently, with a slight abuse of notation, we conclude that $S \subset S_2$.*

Proof. Suppose that σ satisfies *S*. If $n = 2k+1$ for n odd and $n = 2k+2$ for n even, then clearly we have

$$\begin{aligned}
0 &\leq \frac{1}{n} + \frac{n-k-1}{n(k+1)}\lambda_2 + \sum_{i=4}^{k} \frac{\lambda_{n-2i+2}}{(i+1)i} + \frac{1}{12}\lambda_{n-4} + \frac{1}{6}\lambda_{n-2} + \frac{1}{2}\lambda_n \\
&\leq \frac{1}{n} + \frac{n-k-1}{n(k+1)}\lambda_2 + \sum_{i=4}^{k} \frac{\lambda_{n-2i+2}}{(i+1)i} + \frac{1}{12}\lambda_{n-4} + \frac{1}{6}\lambda_{n-4} + \frac{1}{4}\lambda_{n-2} + \frac{1}{4}\lambda_n \\
&= \frac{1}{n} + \frac{n-k-1}{n(k+1)}\lambda_2 + \sum_{i=4}^{k} \frac{\lambda_{n-2i+2}}{(i+1)i} + \frac{1}{4}\lambda_{n-4} + \frac{1}{4}\lambda_{n-2} + \frac{1}{4}\lambda_n.
\end{aligned}$$

Therefore, σ satisfies S_2. Finally, the list $\sigma = \{1, 0, 0, 0, 0, 0, 0, 0, 0, 0, -\frac{1}{2}\}$ satisfies S_2, but does not satisfy *S*. Thus the proof is completed. □

Observation 5.4.3. *Let $\sigma = \{1, \lambda_2, ..., \lambda_n\}$ be a normalized Suleimanova spectrum with $n \geq 16$. If σ satisfies N then σ satisfies N_2. Conversely, it is not always true. Thus, with a slight abuse of notation, we have $N \subset N_2$.*

Proof. We distinguish between the cases n odd and n even.
Case 1: If n is even and σ satisfies N, then

$$\begin{aligned}
0 &\leq \frac{1}{n}+\frac{1}{n}\lambda_2+\frac{\frac{n}{2}-\left[\frac{n+2}{4}\right]}{\left(\frac{n}{2}\right)\left[\frac{n+2}{4}\right]}\lambda_4+\sum_{i=4}^{\left[\frac{n+2}{4}\right]-1}\frac{\lambda_{n-4i+4}}{(i+1)i}+\frac{1}{12}\lambda_{n-8}+\frac{1}{6}\lambda_{n-4}+\frac{1}{2}\lambda_n\\
&\leq \frac{1}{n}+\frac{1}{n}\lambda_2+\frac{\frac{n}{2}-\left[\frac{n+2}{4}\right]}{\left(\frac{n}{2}\right)\left[\frac{n+2}{4}\right]}\lambda_4+\sum_{i=4}^{\left[\frac{n+2}{4}\right]-1}\frac{\lambda_{n-4i+4}}{(i+1)i}+\frac{1}{12}\lambda_{n-8}+\frac{1}{6}\lambda_{n-8}+\frac{1}{4}\lambda_{n-4}+\frac{1}{4}\lambda_n\\
&= \frac{1}{n}+\frac{1}{n}\lambda_2+\frac{\frac{n}{2}-\left[\frac{n+2}{4}\right]}{\left(\frac{n}{2}\right)\left[\frac{n+2}{4}\right]}\lambda_4+\sum_{i=4}^{\left[\frac{n+2}{4}\right]-1}\frac{\lambda_{n-4i+4}}{(i+1)i}+\frac{1}{4}\lambda_{n-8}+\frac{1}{4}\lambda_{n-4}+\frac{1}{4}\lambda_n.
\end{aligned}$$

Therefore, σ satisfies N_2. However, the list

$$\left(1,0,0,0,0,0,0,0,0,0,0,-\frac{1}{20},\frac{1}{20},\frac{1}{10},\frac{1}{10},\frac{1}{10},\frac{1}{10},\frac{1}{10}\right),$$

satisfies N_2 and does not satisfy N.

Case 2: If n is odd and σ satisfies N, then

$$\begin{aligned}
0 &\leq \frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_4+\sum_{i=4}^{\left[\frac{n+3}{4}\right]-1}\frac{\lambda_{n-4i+4}}{(i+1)i}+\frac{1}{12}\lambda_{n-8}+\frac{1}{6}\lambda_{n-4}+\frac{1}{2}\lambda_n\\
&\leq \frac{1}{n}+\frac{n-1}{n}\lambda_2+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_4+\sum_{i=4}^{\left[\frac{n+3}{4}\right]-1}\frac{\lambda_{n-4i+4}}{(i+1)i}+\frac{1}{12}\lambda_{n-8}+\frac{1}{6}\lambda_{n-8}+\frac{1}{4}\lambda_{n-4}+\frac{1}{4}\lambda_n\\
&= \frac{1}{n}+\frac{n-1}{n}\lambda_2+\frac{\frac{n+1}{2}-\left[\frac{n+2}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_4+\sum_{i=4}^{\left[\frac{n+3}{4}\right]-1}\frac{\lambda_{n-4i+4}}{(i+1)i}+\frac{1}{4}\lambda_{n-8}+\frac{1}{4}\lambda_{n-4}+\frac{1}{4}\lambda_n.
\end{aligned}$$

Therefore, σ satisfies N_2. Finally, the list

$$\left(1,0,0,0,0,0,0,0,-\frac{3}{100},-\frac{3}{100},-\frac{3}{100},-\frac{5}{100},-\frac{5}{100},-\frac{1}{10},-\frac{1}{10},-\frac{1}{10},-\frac{1}{10},-\frac{1}{10}\right),$$

satisfies N_2 and does not satisfy N. □

In the next theorem, we study the inclusion relations between the elements of each sequence of the three sequences defined by M_p, S_p and N_p for $p \geq 2$.

Theorem 5.4.4. *Let $n \geq 4$ be an integer, and let $0 \geq \lambda_2 \geq ... \geq \lambda_n \geq -1$. Define w_p, v_p, z_p and y_p as follows.*

- *For any positive integer $p \geq 2$ with $2^p \leq n$, let*

$$w_p := \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{2^p(2^p+1)}\lambda_{n-2^p+1} + \frac{1}{2^p}(\lambda_{n-2^p+2} + ... + \lambda_{n-1} + \lambda_n).$$

- *For $n = \begin{cases} 2k+1 & \text{if } n \text{ is odd} \\ 2k+2 & \text{if } n \text{ is even} \end{cases}$, and for all positive integers p with $2^{p+1} \leq n$, let*

$$v_p := \frac{1}{n} + \frac{n-k-1}{n(k+1)}\lambda_2 + \sum_{i=2^p}^{k} \frac{\lambda_{n-2i+2}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1} \lambda_{n-2i+2}.$$

- *For n even, and for all positive integers $p \geq 2$ with $2^{p+2} \leq n$, let*

$$z_p := \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{\frac{n}{2} - [\frac{n+2}{4}]}{(\frac{n}{2})[\frac{n+2}{4}]}\lambda_4 + \sum_{i=2^p}^{[\frac{n+2}{4}]-1} \frac{\lambda_{n-4i+4}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1} \lambda_{n-4i+4}.$$

- *For n odd, and for any positive integer $p \geq 2$ with $2^{p+2} \leq n$, define*

$$y_p := \frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{\frac{n+1}{2} - [\frac{n+3}{4}]}{(\frac{n+1}{2})[\frac{n+3}{4}]}\lambda_4 + \sum_{i=2^p}^{[\frac{n+3}{4}]-1} \frac{\lambda_{n-4i+4}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1} \lambda_{n-4i+4}.$$

Then, for p>2, we have

1) $w_p - w_{p-1} \geq 0$,

2) $v_p - v_{p-1} \geq 0$,

3) $z_p - z_{p-1} \geq 0$,

4) $y_p - y_{p-1} \geq 0$.

Proof.

1) For simplicity, let us denote $m = 2^p$. Then,

$$w_p = \frac{1}{n} + \sum_{i=m}^{n-1} \frac{\lambda_{n-i+1}}{i(i+1)} + \frac{1}{m} \sum_{i=1}^{m-1} \lambda_{n-i+1},$$

and

$$w_{p-1} = \frac{1}{n} + \sum_{i=\frac{m}{2}}^{n-1} \frac{\lambda_{n-i+1}}{i(i+1)} + \frac{1}{\frac{m}{2}} \sum_{i=1}^{\frac{m}{2}-1} \lambda_{n-i+1}.$$

It follows that for $p > 2$,

$$\begin{aligned}
&w_p - w_{p-1} \\
&= -\sum_{i=\frac{m}{2}}^{m-1} \frac{\lambda_{n-i+1}}{i(i+1)} + \frac{1}{m} \sum_{i=1}^{m-1} \lambda_{n-i+1} - \frac{1}{\frac{m}{2}} \sum_{i=1}^{\frac{m}{2}-1} \lambda_{n-i+1} \\
&= \sum_{i=\frac{m}{2}}^{m-1} \left(\frac{1}{m} - \frac{1}{i(i+1)} \right) \lambda_{n-i+1} + \left(\frac{1}{m} - \frac{1}{\frac{m}{2}} \right) \sum_{i=1}^{\frac{m}{2}-1} \lambda_{n-i+1} \\
&= \sum_{i=\frac{m}{2}}^{m-2} \left(\frac{1}{m} - \frac{1}{i(i+1)} \right) \lambda_{n-i+1} + \left(\frac{1}{m} - \frac{1}{m(m-1)} \right) \lambda_{n-m+2} - \frac{1}{m} \sum_{i=1}^{\frac{m}{2}-1} \lambda_{n-i+1} \\
&= \frac{1}{m} \left(\sum_{i=\frac{m}{2}}^{m-2} \lambda_{n-i+1} - \sum_{i=1}^{\frac{m}{2}-1} \lambda_{n-i+1} \right) + \left(\frac{1}{m} - \frac{1}{m(m-1)} \right) \lambda_{n-m+2} - \sum_{i=\frac{m}{2}}^{m-2} \frac{1}{i(i+1)} \lambda_{n-i+1} \\
&= \frac{1}{m} \left(\sum_{i=\frac{m}{2}}^{m-2} \lambda_{n-i+1} - \sum_{i=1}^{\frac{m}{2}-1} \lambda_{n-i+1} \right) + \sum_{i=\frac{m}{2}}^{m-2} \frac{1}{i(i+1)} \lambda_{n-m+2} - \sum_{i=\frac{m}{2}}^{m-2} \frac{1}{i(i+1)} \lambda_{n-i+1} \\
&= \frac{1}{m} \left(\sum_{i=\frac{m}{2}}^{m-2} \lambda_{n-i+1} - \sum_{i=1}^{\frac{m}{2}-1} \lambda_{n-i+1} \right) + \sum_{i=\frac{m}{2}}^{m-2} \frac{1}{i(i+1)} (\lambda_{n-m+2} - \lambda_{n-i+1}) \geq 0,
\end{aligned}$$

since $0 \geq \lambda_2 \geq ... \geq \lambda_n \geq -1$.

2) By an identical proof of Part 1), we have

$$v_p - v_{p-1} \quad = \frac{1}{m} \left(\sum_{i=\frac{m}{2}}^{m-2} \lambda_{n-2i+2} - \sum_{i=1}^{\frac{m}{2}-1} \lambda_{n-2i+2} \right) + \sum_{i=\frac{m}{2}}^{m-2} \frac{1}{i(i+1)} (\lambda_{n-2m+4} - \lambda_{n-2i+2}) \geq 0,$$

since the eigenvalues are in the decreasing order.

3) and 4) By a similar proof of Part 1), we can check that

$$\begin{aligned}
&z_p - z_{p-1}\\
&= y_p - y_{p-1}\\
&= \frac{1}{m}\left(\sum_{i=\frac{m}{2}}^{m-2}\lambda_{n-4i+4} - \sum_{i=1}^{\frac{m}{2}-1}\lambda_{n-4i+4}\right) + \sum_{i=\frac{m}{2}}^{m-2}\frac{1}{i(i+1)}\left(\lambda_{n-4m+8} - \lambda_{n-4i+4}\right) \geq 0.
\end{aligned}$$

□

As a result, we have the following observation.

Observation 5.4.5. *For a normalized Suleimanova spectrum, we have*

- $M_2 \subset M_3 \subset ... \subset M_p \subset M_{p+1} \subset ....$
- $S_2 \subset S_3 \subset ... \subset S_p \subset S_{p+1} \subset ....$
- $N_2 \subset N_3 \subset ... \subset N_p \subset N_{p+1} \subset ....$

We know that $PM \subset S$. So a natural question arises here; if for a fixed p, $M_p \subset S_p$. That we will see in the next observation.

Observation 5.4.6. *Let $p \geq 2$ be a fixed integer and $n \geq 2^{p+1}$. Then, for a normalized Suleimanova spectrum $\sigma = \{1, \lambda_2, ..., \lambda_n\}$ with $0 \geq \lambda_2 \geq .... \geq \lambda_n \geq -1$, $M_p \subset S_p$.*

Proof. Suppose that σ satisfies M_p. If $n = 2k+1$ for n odd or $n = 2k+2$ for n even, then we have

$$\begin{aligned}
0 &\leq \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{2^p(2^p+1)}\lambda_{n-2^p+1} + \frac{1}{2^p}(\lambda_{n-2^p+2} + ... + \lambda_{n-1} + \lambda_n)\\
&= \frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{(k+2)(k+1)}\lambda_{n-k} + \sum_{i=2^p}^{k}\frac{\lambda_{n-i+1}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1}\lambda_{n-i+1}\\
&\leq \frac{1}{n} + \left(\frac{1}{n(n-1)} + ... + \frac{1}{(k+2)(k+1)}\right)\lambda_2 + \sum_{i=2^p}^{k}\frac{\lambda_{n-2(i+1)}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1}\lambda_{n-2(i+1)}\\
&\qquad (\text{since } \lambda_2 \geq \lambda_j \text{ for } j = 3, ..., k \text{ and } \lambda_{n-2i} \geq \lambda_{n-i} \text{ for } i = 1, ..., k-1)\\
&= \frac{1}{n} + \frac{n-k-1}{n(k+1)}\lambda_2 + \sum_{i=2^p}^{k}\frac{\lambda_{n-2i+2}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1}\lambda_{n-2i+2}
\end{aligned}$$

It follows that σ satisfies S_p.

□

At this point, it is worthy to mention that for a fixed integer p, the condition S_p is a refinement of M_p. However, we can make use of M_p in the cases where S_p cannot be applied as we see in the following examples.

Example 20. Consider the list $\sigma = (1,0,0,0,0,-\frac{2}{3})$. It is easy to see that σ does not verify Soules condition but it verifies condition M_2 (S_2 cannot be applied as $n < 8$). Therefore, it is the spectrum of a 6×6 symmetric doubly stochastic matrix.

Example 21. For $n = 9$, consider the list $\sigma = (1,0,0,0,-\frac{1}{14},-\frac{1}{5},-\frac{1}{5},-\frac{1}{5},-\frac{1}{5})$. A simple check shows that σ does not verify Soules condition, nor New condition 1. Moreover, σ does not verify the conditions M_2 and S_2. However σ verifies M_3 (note that S_3 and N_2 cannot be applied), and therefore it is the spectrum of a 9×9 symmetric doubly stochastic matrix.

Finally, by making use of Theorem 4.2.2, we end our chapter with the following sufficient condition concerning realizable Suleimanova spectra of even order.

Theorem 5.4.7. *For any two nonnegative powers of 2, n and m with $m \geq n$, let $\sigma = \{1, \lambda_2, ..., \lambda_{n+m}\} \subset \mathbb{R}$, such that $0 \geq \lambda_2 \geq ... \geq \lambda_{n+m} \geq -1$. If there exists a permutation π on $\{\lambda_2, ..., \lambda_{n+m}\}$ such that $\pi(\lambda_2, ..., \lambda_{n+m}) = (\mu_2, ..., \mu_{n+m})$ and*

$$\mu_2 > -\frac{n}{m},$$

$$\mu_3 + ... + \mu_{n+1} \geq -\frac{m\mu_2 + n}{m+n},$$

and

$$\mu_{n+2} + ... + \mu_{n+m} \geq -\frac{n\mu_2 + m}{m+n},$$

then σ is realizable by an $(n+m) \times (n+m)$ symmetric doubly stochastic matrix.

Proof. Since $0 \geq \frac{m+n}{m\mu_2+n}\mu_3 + ... + \frac{m+n}{m\mu_2+n}\mu_{n+1} \geq -1$, and n is a nonnegative power of 2, then in view of Theorem 5.3.2, there exists an $n \times n$ symmetric doubly stochastic matrix $\widehat{A}$ with eigenvalues $(1, \frac{m+n}{m\mu_2+n}\mu_3, ..., \frac{m+n}{m\mu_2+n}\mu_{n+1})$. Similarly, since $0 \geq \frac{m+n}{n\mu_2+m}\lambda_{n+2} + ... + \frac{m+n}{n\mu_2+m}\mu_{n+m} \geq -1$ and m is a nonnegative power of 2, then by Theorem 5.3.2, there exists an $m \times m$ symmetric doubly stochastic matrix $\widehat{B}$ with spectrum $(1, \frac{m+n}{n\mu_2+m}\mu_{n+2}, ..., \frac{m+n}{n\mu_2+m}\mu_{n+m})$. In Theorem 4.2.2, taking $\alpha = 1$, $\rho = \frac{\sqrt{mn}(1-\mu_2)}{m\mu_2+n}$ (as $\mu_2 > -\frac{n}{m}$, then $\rho > 0$), $A = \widehat{A}$ and $B = \widehat{B}$, then we get an $(n+m) \times (n+m)$ symmetric doubly stochastic matrix D with eigenvalues $1, \lambda_2, \lambda_3, ..., \lambda_{n+m}$. □

By making use of the previous theorem, we will show that conditions S, N, M_p, S_p and N_p are not necessary conditions for NS-SDIEP, in the following example.

Example 22. Consider the list of 6 real numbers $\sigma = \{1,0,0,-\frac{1}{6},-\frac{1}{6},-\frac{11}{18}\}$. It is easy to check that σ does not verify S nor N. Moreover, σ does not verify M_2 (note that S_2 and N_2 cannot be applied). However, taking $\mu_2 = -\frac{1}{6}$, $\mu_3 = -\frac{1}{6}$, $\mu_4 = 0$, $\mu_5 = 0$, and $\mu_6 = -\frac{11}{18}$, then σ verifies Theorem 5.4.7 for the case of order $6 = m+n$ with $m = 4$ and $n = 2$. Therefore, it is realizable by a 6×6 symmetric doubly stochastic matrix C given by

$$C = \frac{2}{9}\begin{pmatrix} A & \frac{7\sqrt{2}}{4}e_2e_4^T \\ \frac{7\sqrt{2}}{4}e_4e_2^T & \frac{11}{4}B \end{pmatrix},$$

where A is a 2×2 symmetric doubly stochastic matrix with eigenvalues $1, -\frac{3}{4}$ and B is 4×4 symmetric doubly stochastic matrix with eigenvalues 1,0,0,-1. Using Theorem 5.3.2, we can construct A and B as follows:

$$A = H_1 \mathrm{diag}\left(1, -\frac{3}{4}\right) H_1^{-1} = \begin{pmatrix} \frac{1}{8} & \frac{7}{8} \\ \frac{7}{8} & \frac{1}{8} \end{pmatrix},$$

and,

$$B = H_2 \mathrm{diag}(1,0,0,-1) H_2^{-1} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}.$$

It follows that

$$C = \begin{pmatrix} \frac{1}{36} & \frac{7}{36} & \frac{7}{36} & \frac{7}{36} & \frac{7}{36} & \frac{7}{36} \\ \frac{7}{36} & \frac{1}{36} & \frac{7}{36} & \frac{7}{36} & \frac{7}{36} & \frac{7}{36} \\ \frac{7}{36} & \frac{7}{36} & 0 & \frac{11}{36} & \frac{11}{36} & 0 \\ \frac{7}{36} & \frac{7}{36} & \frac{11}{36} & 0 & 0 & \frac{11}{36} \\ \frac{7}{36} & \frac{7}{36} & \frac{11}{36} & 0 & 0 & \frac{11}{36} \\ \frac{7}{36} & \frac{7}{36} & 0 & \frac{11}{36} & \frac{11}{36} & 0 \end{pmatrix}.$$

Chapter 6

Conclusions and Futur works

The final chapter summarises the work presented in this thesis and our findings, and describes a few further research lines.

6.1 Summary

In Chapter 2, we aimed to prove that general class of matrices are positive semi-definite by constructing Gram matrices between characteristic functions of subsets of a finite set E (see Theorem 2.4.5 and Theorem 2.4.7). Our approach for that is explained by the fact that binary similarity matrices can be written as cardinals of theoretical operations between subsets of a finite set (see Section 2.5.1). As a result, our class of PSD matrices paved the way to prove that many binary similarity matrices, given in Section 2.5, are PSD. In fact, we have proven that two family of binary similarity measures S_θ and T_θ, introduced in [36], are positive semi-definite for $\theta \geq 1$. Moreover, we have shown that from 30 popular similarity matrices defined in the literature, 21 are PSD and 9 are not PSD. The motivation for this work was the growing number of applications of PSD similarity matrices in deriving distance functions. One simple method of defining an Euclidean distance from a PSD similarity s defined on $[0,1]$, is the transformation $d = \sqrt{1-s}$ where the positive semi-definite property is a sufficient condition to assure that d is Euclidean. Moreover, the positive semi-definite property ensures that generalised Euclidean distance d_M defined by $d_M(x,y) = \sqrt{(x-y)^\top M(x-y)}$, where M is any $n \times n$ PSD matrix and x and $y \in \mathbb{R}^n$, satisfies the properties of a pseudo-metric. All these nice metric or pseudo metric constructions between data points plays an important role in many applications of machine learning, pattern recognition and data mining techniques such as K-means and K-Nearest Neighbour classifier, that we discussed in details in Section 2.10. Another contribution in Chapter 2, is proving the strictly positive definite property of Tversky

family of similarity matrices (see Theorem 2.6.5), as a generalisation to the Jaccard index, the only similarity matrix to be shown that is PD [11].

Our second direction on studying positive semi-definite matrices was to explore the conditions under which a positive semi-definite doubly stochastic matrix has nonnegative pth roots for a given p. The motivation for that can be explained by the many interesting applications of the computation of stochastic roots in different areas such as finance and healthcare. As the problem of considering roots of doubly stochastic matrices that are positive semi-definite, was only discussed by Marcus and Minc [77] and briefly for the special case of $p = 2$, we have been interested in characterising the set of $n \times n$ PSD doubly stochastic matrices with nonnegative pth roots, denoted by $K_n^{1/p}$. First, in Theorem 3.2.3, we obtained new sufficient condition for this problem that improves the existing one for the case $p = 2$. We have not been able to give a full characterisation of $K_n^{1/p}$, but we prove some geometrical properties of the set. More precisely, we show in Theorem 3.3.6, that $K_n^{1/p}$ is not convex for $n \geq 3$ but star convex with respect to J_n. In addition, we show that a positive semi-definite doubly stochastic matrix that is the inverse of an M-matrix belongs to $K_n^{1/p}$ for any integer p. As a result, a construction of some elements of $K_n^{1/p}$ was identified in Theorem 3.4.3. One of the problems that is related to the doubly stochastic roots problem is the symmetric inverse eigenvalue problem for doubly stochastic matrices. This was discussed in Section 3.5. Two types of symmetric doubly stochastic matrices can be defined. PSD doubly stochastic matrices which can be obtained using the concept of Soules matrices[122] and which are elements of $K_n^{1/p}$, for any integer p, as Theorem 3.5.3 shows. And the other type, symmetric doubly stochastic matrices obtained from orthogonal matrices which are not any Soules matrix. It seems that there is not yet a systematic way to examine the belonging of these matrices to $K_n^{1/p}$. However, in Theorem 3.6.4, a general characterisation of $K_n^{1/p}$ is given for the particular case of $n = 3$ and $p = 2$. As a generalisation, Theorem 3.6.7, explores a family of matrices that belongs to $K_n^{1/2}$ for any n.

The connection between generating elements of $K_n^{1/p}$ and the inverse eigenvalue problem led us naturally to study SDIEP in this thesis. Many sufficient conditions for this problem have been given over the years, beginning with the Perfect and Mirsky work [97], then with Soules sufficient condition [122]. However, the problem of characterising the spectra of all symmetric doubly stochastic matrices seems unlikely to be solved in the near future. We have introduced new recursive sufficient conditions for constructing symmetrically realisable lists (see Theorem 4.3.5 and Theorem 4.3.9). These sufficient conditions are shown to be not a refinement of Soules but new independent conditions. In light of these result, it was natural to seek symmetrically normalised Suleimanova realisable lists. Indeed, Chapter 5 introduced three sequences of sufficient conditions (see Theorem 5.3.3, Theorem 5.3.4, and Theorem

5.3.5) for normalised Suleimanova spectrum to be the spectrum of doubly stochastic matrices. We could mention the sequence S_p, that is proved to refine Soules condition for the case of normalised Suleimanova spectra , as we show in Theorem 5.4.4.

6.2 Futur work

In a similar way to the work presented in Chapter 2, one can further investigate the positive semi-definite property of numerical similarity matrices. As most practitioners will care about having a vector-based measure, we strongly believe that by a good choice of the measure μ, we can construct several general PSD matrices that could help us to deduce the positive semi-definiteness of many numerical similarity measures. Moreover, it seems that the main method in proving the strictly positive definiteness of Tversky similarity family of matrices could as well be employed for proving that the complete $(2^n-1)\times(2^n-1)$ similarity matrix F whose elements are given by

$$F(A,B)=\frac{|(A\triangle B)^c|}{a|A\triangle B|+b|(A\triangle B)^c|},$$

for any non empty subsets A and B of E, is also positive definite for any positive integer $n>1$, and for all $a\geq b>0$.

As the reader can see, the results in the thesis are concerned with the theory of matrices as linear algebra has proved useful for machine learning, AI and deep learning problems, as we have briefly discussed. The next step may be to consider some numerical applications that show the utility of the positive semi-definite property and doubly stochasticity. Using new distance functions derived from PSD similarity matrices can produce noticeably superior results then algorithms using the simple Euclidean distance. We can try the use of K-means with an Euclidean distance derived from a PSD similarity matrix via the transformation $d=\sqrt{1-s}$ or the use of a generalised Euclidean distance d_M where the weighted matrix M is one of the PSD similarity matrices considered in Chapter 2. Such distance d_M can be used as well to improve clustering by learning a doubly stochastic data similarity matrix [130]. Moreover, exploring the structure of the spectrum of symmetric doubly stochastic matrices is of a great importance in defining clusters using the consensus clustering [75].

Additional to these research subjects that rise directly from the study, one can investigate more on the symmetric doubly stochastic inverse eigenvalue problem, that is still open. Our next step would be to solve Conjecture 4.3.13 that will lead to many new sufficient conditions for SDIEP. As a result, we could characterise new elements of $K_n^{1/p}$ for any integers n and p.

Moreover, it would be interesting to see how to use a similar recursive method for generating sufficient conditions for RDIEP and DIEP.

Appendix A

We will give the proof of Observation 4.3.11.

Proof. Let n be even. We distinguish between the cases $n = 8p$, $n = 8p+2$, $n = 8p+4$ and $n = 8p+6$.

Case 1: For $n = 8p$, if σ satisfies Soules then,

$$
\begin{aligned}
0 &\leq \frac{1}{n} + \frac{1}{n}\lambda_2 + \sum_{i=1}^{4p-1} \frac{\lambda_{n-2i+2}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \sum_{i=p}^{4p-1} \frac{\lambda_{n-2i+2}}{(i+1)i} + \sum_{i=1}^{p-1} \frac{\lambda_{n-2i+2}}{(i+1)i} \\
&\leq \frac{1}{n} + \frac{1}{n}\lambda_2 + \sum_{i=p}^{4p-1} \frac{\lambda_{n-2i+2}}{(i+1)i} + \sum_{i=1}^{p-1} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{1}{4p(4p-1)}\lambda_4 + \frac{1}{(4p-2)(4p-1)}\lambda_6 + \sum_{i=p}^{4p-3} \frac{\lambda_{n-2i+2}}{(i+1)i} + \sum_{i=1}^{p-1} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&\leq \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{1}{4p(4p-1)}\lambda_4 + \frac{1}{(4p-2)(4p-1)}\lambda_4 + \sum_{i=p}^{4p-3} \frac{\lambda_8}{(i+1)i} + \sum_{i=1}^{p-1} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{1}{4p(4p-1)}\lambda_4 + \frac{1}{(4p-2)(4p-1)}\lambda_4 + \frac{3p-2}{p(4p-2)}\lambda_8 + \sum_{i=1}^{p-1} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{1}{4p(4p-1)}\lambda_4 + \frac{1}{(4p-2)(4p-1)}\lambda_4 + \left(\frac{4p-3}{2p(4p-2)} + \frac{1}{4p}\right)\lambda_8 + \sum_{i=1}^{p-1} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&\leq \frac{1}{n} + \frac{1}{n}\lambda_2 + \left(\frac{1}{4p(4p-1)} + \frac{1}{(4p-2)(4p-1)} + \frac{4p-3}{2p(4p-2)}\right)\lambda_4 + \frac{1}{4p}\lambda_8 + \sum_{i=1}^{p-1} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{2}{n}\lambda_4 + \frac{\frac{n}{4} - \left[\frac{n+4}{8}\right]}{\left(\frac{n}{4}\right)\left[\frac{n+4}{8}\right]}\lambda_8 + \sum_{i=1}^{\left[\frac{n+4}{8}\right]-1} \frac{\lambda_{n-8i+8}}{(i+1)i}.
\end{aligned}
$$

Therefore, σ verifies New condition 2.

Case 2: For $n = 8p+2$, if σ satisfies Soules then,

$$
\begin{aligned}
0 &\le \frac{1}{n} + \frac{1}{n}\lambda_2 + \sum_{i=1}^{4p} \frac{\lambda_{n-2i+2}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \sum_{i=p+1}^{4p} \frac{\lambda_{n-2i+2}}{(i+1)i} + \sum_{i=1}^{p} \frac{\lambda_{n-2i+2}}{(i+1)i} \\
&\le \frac{1}{n} + \frac{1}{n}\lambda_2 + \sum_{i=p+1}^{4p} \frac{\lambda_{n-2i+2}}{(i+1)i} + \sum_{i=1}^{p} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{1}{4p(4p+1)}\lambda_4 + \frac{1}{4p(4p-1)}\lambda_6 + \sum_{i=p+1}^{4p-2} \frac{\lambda_{n-2i+2}}{(i+1)i} + \sum_{i=1}^{p} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&\le \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{1}{4p(4p+1)}\lambda_4 + \frac{1}{4p(4p-1)}\lambda_6 + \sum_{i=p+1}^{4p-2} \frac{\lambda_{8}}{(i+1)i} + \sum_{i=1}^{p} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{1}{4(4p+1)}\lambda_4 + \frac{1}{4p(4p-1)}\lambda_6 + \frac{3p-2}{(p+1)(4p-1)}\lambda_8 + \sum_{i=1}^{p} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{1}{4p(4p+1)}\lambda_4 + \frac{1}{4p(4p-1)}\lambda_6 + \left(\frac{2p-2}{(4p-1)(2p+1)} + \frac{p}{(p+1)(2p+1)}\right)\lambda_8 \\
&\quad + \sum_{i=1}^{p} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&\le \frac{1}{n} + \frac{1}{n}\lambda_2 + \left(\frac{1}{4p(4p+1)} + \frac{1}{4p(4p-1)} + \frac{2p-2}{(4p-1)(2p+1)}\right)\lambda_4 + \frac{p}{(p+1)(2p+1)}\lambda_8 \\
&\quad + \sum_{i=1}^{p} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{2(n-2)}{n(n+2)}\lambda_4 + \frac{\frac{n+2}{4} - \left[\frac{n+6}{8}\right]}{\left(\frac{n+2}{4}\right)\left[\frac{n+6}{8}\right]}\lambda_8 + \sum_{i=1}^{\left[\frac{n+6}{8}\right]-1} \frac{\lambda_{n-8i+8}}{(i+1)i}.
\end{aligned}
$$

Therefore σ satisfies New condition 2.

Case 3: For $n = 8p+4$, if σ satisfies Soules then,

$$
\begin{aligned}
0 &\leq \frac{1}{n} + \frac{1}{n}\lambda_2 + \sum_{i=1}^{4p+1} \frac{\lambda_{n-2i+2}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \sum_{i=p+1}^{4p+1} \frac{\lambda_{n-2i+2}}{(i+1)i} + \sum_{i=1}^{p} \frac{\lambda_{n-2i+2}}{(i+1)i} \\
&\leq \frac{1}{n} + \frac{1}{n}\lambda_2 + \sum_{i=p+1}^{4p+1} \frac{\lambda_{n-2i+2}}{(i+1)i} + \sum_{i=1}^{p} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{1}{(4p+1)(4p+2)}\lambda_4 + \frac{1}{4p(4p+1)}\lambda_6 + \sum_{i=p}^{4p-1} \frac{\lambda_{n-2i+2}}{(i+1)i} + \sum_{i=1}^{p} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&\leq \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{1}{(4p+1)(4p+2)}\lambda_4 + \frac{1}{4p(4p+1)}\lambda_6 + \sum_{i=p}^{4p-1} \frac{1}{(i+1)i}\lambda_8 + \sum_{i=1}^{p} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{1}{(4p+1)(4p+2)}\lambda_4 + \frac{1}{4p(4p+1)}\lambda_6 + \frac{3p-1}{4p(p+1)}\lambda_8 + \sum_{i=1}^{p} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{1}{(4p+1)(4p+2)}\lambda_4 + \frac{1}{4p(4p+1)}\lambda_6 + \left(\frac{2p-1}{4p(2p+1)} + \frac{p}{(p+1)(2p+1)}\right)\lambda_8 \\
&\quad + \sum_{i=1}^{p} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&\leq \frac{1}{n} + \frac{1}{n}\lambda_2 + \left(\frac{1}{(4p+1)(4p+2)} + \frac{1}{4p(4p+1)} + \frac{2p-1}{4p(2p+1)}\right)\lambda_4 + \frac{p}{(p+1)(2p+1)}\lambda_8 \\
&\quad + \sum_{i=1}^{p} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{2}{n}\lambda_4 + \frac{\frac{n}{4} - \left[\frac{n+4}{8}\right]}{\left(\frac{n}{4}\right)\left[\frac{n+4}{8}\right]}\lambda_8 + \sum_{i=1}^{\left[\frac{n+4}{8}\right]-1} \frac{\lambda_{n-8i+8}}{(i+1)i}.
\end{aligned}
$$

Therefore σ satisfies New condition 2.

Case 4: For $n = 8p+6$, if σ satisfies Soules then,

$$
\begin{aligned}
0 &\leq \frac{1}{n} + \frac{1}{n}\lambda_2 + \sum_{i=1}^{4p+2} \frac{\lambda_{n-2i+2}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \sum_{i=p+1}^{4p+2} \frac{\lambda_{n-2i+2}}{(i+1)i} + \sum_{i=1}^{p} \frac{\lambda_{n-2i+2}}{(i+1)i} \\
&\leq \frac{1}{n} + \frac{1}{n}\lambda_2 + \sum_{i=p+1}^{4p+2} \frac{\lambda_{n-2i+2}}{(i+1)i} + \sum_{i=1}^{p} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{1}{(4p+2)(4p+3)}\lambda_4 + \frac{1}{(4p+2)(4p+1)}\lambda_6 + \sum_{i=p+1}^{4p} \frac{\lambda_{n-2i+2}}{(i+1)i} + \sum_{i=1}^{p} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&\leq \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{1}{(4p+2)(4p+3)}\lambda_4 + \frac{1}{(4p+2)(4p+1)}\lambda_6 + \sum_{i=p+1}^{4p} \frac{\lambda_8}{(i+1)i} + \sum_{i=1}^{p} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{1}{(4p+2)(4p+3)}\lambda_4 + \frac{1}{(4p+2)(4p+1)}\lambda_6 + \frac{3p}{(p+1)(4p+1)}\lambda_8 + \sum_{i=1}^{p} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{1}{(4p+2)(4p+3)}\lambda_4 + \frac{1}{(4p+2)(4p+1)}\lambda_6 + \left(\frac{2p-1}{(4p+1)(2p+2)} + \frac{1}{2p+2}\right)\lambda_8 \\
&\quad + \sum_{i=1}^{p} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&\leq \frac{1}{n} + \frac{1}{n}\lambda_2 + \left(\frac{1}{(4p+2)(4p+3)} + \frac{1}{(4p+2)(4p+1)} + \frac{2p-1}{(4p+1)(2p+2)}\right)\lambda_4 + \frac{1}{2p+2}\lambda_8 \\
&\quad + \sum_{i=1}^{p} \frac{\lambda_{n-8i+8}}{(i+1)i} \\
&= \frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{2(n-2)}{n(n+2)}\lambda_4 + \frac{\frac{n+2}{4} - \left[\frac{n+6}{8}\right]}{\left(\frac{n+2}{4}\right)\left[\frac{n+6}{8}\right]}\lambda_8 + \sum_{i=1}^{\left[\frac{n+6}{8}\right]-1} \frac{\lambda_{n-8i+8}}{(i+1)i}.
\end{aligned}
$$

Therefore σ satisfies New condition 2. □

Appendix B: Synthèse en français

La théorie des matrices s'est développée rapidement au cours des dernières décennies en raison de son large éventail d'applications et de ses nombreux liens avec différents domaines des mathématiques, de l'économie, de l'apprentissage automatique et du traitement du signal. L'une des principales caractéristiques qu'une matrice puisse posséder est la propriété semi-définie positive. La classe de telles matrices est fondamentale dans l'analyse matricielle et apparaît naturellement dans de nombreuses applications telles que l'analyse de données, l'analyse complexe, la mécanique et l'ingénierie. Les matrices semi-définies positives sont utilisées comme matrices de covariance en statistique, en tant que noyaux en apprentissage automatique et en tant que tenseurs en imagerie médicale. En outre, la propriété semi-definie positive est désormais importante lorsqu'on considère les matrices de similarité [35, 36, 131], matrices utilisées pour quantifier la ressemblance des éléments d'un espace de données. La propriété mentionnée nous permet de construire des distances métriques entre les éléments d'un ensemble de données.

D'autre part, les matrices non-négatives sont une partie intégrante de la théorie des matrices, initiées par Oskar Perron [98] et par George Frobenius [33] au début du XXe siècle. Parmi les matrices non-négatives, la classe des matrices doublement stochastiques est d'importance particulière. En effet, elle est particulièrement riche en applications dans d'autres domaines des mathématiques, tels que la théorie des graphes, la combinatoire, l'analyse numérique, ainsi que dans d'autres domaines tels que l'économie, la statistique, et la théorie des communications [14, 75, 124, 130, 134]. Un problème classique de l'analyse matricielle implique l'étude des racines des matrices. Plus particulièrement, trouver des racines qui préservent la non-négativité des matrices non-négatives [73, 74, 89, 125] et récemment des matrices stochastiques [37, 38] joue un rôle important dans plusieurs applications. En combinant la propriété doublement stochastique et la propriété semi-définie positive dans une matrice A, il est naturel de rechercher les conditions pour lesquelles la p-ième racine de A, pour tout entier naturel $p \geq 2$, est aussi une matrice doublement stochastique semi-définie positive. Un autre objet d'étude intéressant dans la théorie des matrices doublement stochastiques est celui du problème inverse des valeurs propres qui, étant un ensemble de nombres

complexes (ou réels) σ, cherche des conditions nécessaires et suffisantes pour qu'une matrice doublement stochastique ait pour spectre σ.

Dans cette thèse, nous apporterons des réponses aux trois problèmes cités ci-dessus concernant des matrices semi-définies positives et des matrices doublement stochastiques.

Définitions et présentation des domaines de recherche

Nous allons d'abord introduire quelques définitions et terminologies avant de mettre en évidence de manière plus détaillée les trois axes de recherche que nous avons abordées dans cette thèse.

Une matrice symétrique A dont les éléments sont des nombres réels, est dite semie-définie positive (PSD) si pour tout vecteur $x \in \mathbb{R}^n$, on a $x^T Ax \geq 0$. En particuler, A est dite définie positive (PD), si $x^T Ax > 0$, pout tout vecteur non nul $x \in \mathbb{R}^n$. Nous verrons plus loin d'autres cractéristiques des matrices semi-définies positives (voir [43]) qui s'avèrent utiles pour notre recherche.

Une matrice carrée réelle est dite doublement quasi-stochastique si la somme de chacune de ses lignes et colonnes est égale à 1. Une matrice doublement quasi-stochastique non-négative est appelée doublement stochastique (ou bi-stochastique). Clairement, une matrice $n \times n$ non-négative A est doublement stochastique si et seulement si

$$Ae_n = e_n$$

et

$$e_n A^T = e_n$$

ou de manière équivalente

$$AJ_n = J_n A = J_n,$$

ou J_n est la matrice $n \times n$ dont les éléments sont égaux à $\frac{1}{n}$ et e_n est le vecteur de taille n avec chaque élément est égale à $\frac{1}{\sqrt{n}}$, c'est-à-dire $e_n = \frac{1}{\sqrt{n}}(1, ..., 1)^T$.

La notion de similarité est un concept fondamental dans de nombreux domaines de recherche et applications tel que l'apprentissage automatique, la recherche d'information et l'analyse de données. En pratique, les similarités sont évaluées par une mesure qui quantifie la ressemblance entre dans des objets. Nous pouvons distinguer différentes mesures en fonction du type de données auquel elles s'appliquent, c'est-à-dire des données binaires, numériques ou structurées. Formallement, si $\mathcal{X}$ désigne l'espace des données, une mesure de similarité S est une fonction de $\mathcal{X} \times \mathcal{X}$ dans $\mathbb{R}$ qui satisfait les propriétés suivantes :

- Positivité : $S(x,y) \geq 0$ pour tout $x,y \in \mathcal{X}$
- Symétrie: $S(x,y) = S(y,x)$ pour tout $x,y \in \mathcal{X}$,
- Maximalité : $S(x,y) \leq S(x,x)$ pour tout $x,y \in \mathcal{X}$.

Une mesure de similarité donne lieu à une matrice de similarité lorsque l'on considère des paires de variables ou des unités d'échantillons.

Après avoir defini les notions de bases, le corps de notre travail est logiquement divisé en trois parties qui abordent les trois problèmes suivants, connectés mais qui s'avèrent d'une importance indépendante.

Motivé par ses applications , la première partie de cette thèse traite la propriété semi-définie positive des matrices de similarité. En fait, les matrices de similarité qui sont semi-définies positives revêtent une importance particulière en raison de leur capacité à définir des distances métriques. Cependant, peu de matrices de similarité se sont avérées semi-définies positives [35, 36]. Pour cela nous analysons les propriétés des matrices de similarité et plus particulièrement la propriété semi-définie positive. L'idée étant d'abord de définir les différents types de mesures de similarité, puis collecter les plus fréquentes trouvées dans la littérature. En utilisant une approche différente pour définir les matrices de similarité en tant qu'opérations théoriques entre des sous-ensembles d'un ensemble fini, de nombreuses matrices de similarité sont désormais démontrées semi-définies positives. D'autres propriétes sont aussi étudiées telles que les propriétés (strictement) définie et semi-définie 3-positive. En outre, une discussion détaillée des nombreuses applications de toutes ces propriétés dans divers domaines est établie.

Un type particulier de matrice de similarité peut être une matrice doublement stochastique. Outre leurs applications en combinatoire, les matrices doublement stochastiques, qui représentent un sous-ensemble particulier des matrices stochastiques, constituent un outil très utile en probabilités et en statistiques. En fait, une matrice stochastique $B = (b_{ij})$ peut être considérée comme une matrice de transition d'une chaîne de Markov où b_{ij} est la probabilité d'aller de l'état i à j. Un des nombreux problèmes étudiés dans la théorie des matrices stochastiques consiste à trouver les racines p-ièmes d'une matrice stochastique [37, 38] étant donné que le calcul de telles racines a conduit à de nombreuses applications dans le domaine des mathématiques financières et des maladies chroniques. Nous étendons dans la deuxieme partie de la thèse, l'analyse au cas des matrices PSD doublement stochastiques. Notre principal objectif est de caractériser des ensembles de matrices semi-définies positives doublement stochastique dont la racine p-ième est aussi doublement stochastique. Une des méthodes utilisée pour caractériser de telles matrices est l'utilisation du problème inverse des valeurs propres des matrices symétriques doublement stochastiques (ou SDIEP) qui

cherche à déterminer l'ensemble des listes de nombres réels qui peuvent être les spectres de matrices doublement stochastiques. Vu la relation étroite existante entre le problème de la recherche de racines et le problème inverse des valeurs propre, la troisième partie de la thèse est consacrée à l'étude de SDIEP. Dans la suite, nous présentons les principaux enjeux et résultats apportés dans cette thèse.

La propriété semi-définie positive des matrices de similarité

La première partie de la thèse est consascrée à l'étude de la propriété semi-definie positive des matrices de similarités car elle permet de définir de nouvelles distances. En fait, un coefficient de similarité est utilisé pour décrire quantitativement la ressemblance entre les éléments d'un ensemble de données dans de nombreux domaines d'applications, tels que l'intégration des informations [5], la théorie des hypergraphes [9, 10] et l'analyse d'images [15]. Cependant, pour certaines applications tel que le clustering, il est plus utile d'avoir des distances métriques [23]. Dans la suite, nous allons cité quelques applications et motivations décrites dans la thèse et qui mettent en valeur l'importance de ces distances définies à partir des mesures de similarité dans plusieurs domaines de l'apprentissage automatique et la théorie de l'évidence. De nombreuses mesures de similarité ont été proposées dans divers domaines, en particulier les coefficients binaires [131]. Cependant, peu de matrices de similarité se sont avérées semi-définies positives [35, 36]. L'idée est d'utiliser une approche différente en considérant les matrices de similarité en tant qu'opérations théoriques entre des sous-ensembles d'un ensemble fini, afin de pouvoir montrer qu'elles sont semi-définies positives.

Une grande variété de données peut être représentée par des variables binaires, qui expriment deux états possibles de l'échantillon (à savoir, présence/absence, vrai/ faux, homme /femme). Généralement parlant, une variable correspond à un objet ou à un individu et les scores binaires reflètent la présence ou l'absence de certaines caractéristiques de l'objet. Considérons donc deux vecteurs binaires $x = (x_1, x_2, ..., x_n)$ et $y = (y_1, y_2, ..., y_n)$ et prenons $X = \{1 \leq i \leq n / x_i = 1\}$, c'est-à-dire l'ensemble des caractéristiques présentes dans x et $Y = \{1 \leq i \leq n / y_i = 1\}$ les caractéristiques présentes dans y . Dans ce cadre, les mesures de similarité binaires peuvent être definies en utilisant les cardinaux des quatre quantités suivantes:

- $a = |X \cap Y|$
- $b = |X \cap \bar{Y}|$
- $c = |\bar{X} \cap Y|$

- $d = |\bar{X} \cap \bar{Y}|$.

Il s'ensuit que toutes les mesures de similarité sont maintenant définies en fonction de la cardinalité des opérations théoriques entre sous-ensembles d'un ensemble fini. Par conséquent, sans perte de généralité, on peut supposer que $\mathcal{X} = P(E)$, l'ensemble de tous les sous-ensembles d'un ensemble fini E de cardinal n et une mesure de similarité S est définie comme une fonction de $P(E) \times P(E)$ dans $\mathbb{R}$. Soient $(A_i)_{1\leq i\leq m}$ m sous-ensembles non vides de E, alors la mesure de similarité S donne lieu à une matrice de similarité $M_S = \big(S(A_i, A_j)\big)_{1\leq i,j\leq m}$.

L'étape suivante consiste principalement à prouver que certaines classes de matrices sont PSD. Ca sera à son tour la base pour prouver qu'un certain nombre de matrices de similarité sont semi-définies positives. Pour arriver à notre but, nous avons construit des matrices de Gram en utilisant l'espace de Hilbert $\mathcal{L}^2(E, \mu, \mathbb{R})$ avec son produit scalaire $<,>$ donné par:

$$< f, g >= \int_E f.g \, d\mu,$$

pour tout $f, g \in \mathcal{L}^2(E, \mu, \mathbb{R})$ ou μ est une mesure positive sur E.

Par un choix adéquat des fonctions f et g et en prenant μ comme étant la mesure de dénombrement, nous avons construit des classes de matrices PSD. Nos principaux résultats sont données dans les théorèmes suivants (voir Theorem 2.4.5 et Theorem 2.4.7).

Théorème 1. Soit $E = \{e_1, \ldots, e_n\}$ un ensemble non vide, et $(A_i)_{1\leq i\leq m}$ m sous-ensembles de E. Alors ce qui suit est valable.

1) La matrice $(|A_i \cap A_j|)_{1\leq i,j\leq m}$ est PSD, et pour tout nombres reels a et b, la matrice

$$\big(a^2.|A_i \cap A_j| + ab.|A_i \triangle A_j| + b^2.|{A_i}^c \cap {A_j}^c|\big)_{1\leq i\leq m}$$

est également PSD.

2) Si $A_i \neq \emptyset$ pour tout i et $a \geq b > 0$, alors les 4 matrices suivantes sont PSD:

i) $\left(\frac{1}{a.|A_i \triangle A_j| + b.|A_i \cap A_j|}\right)_{1\leq i,j\leq m}$

ii) $\left(\frac{|A_i \cap A_j|}{a.|A_i \triangle A_j| + b.|A_i \cap A_j|}\right)_{1\leq i,j\leq m}$

iii) $\left(\frac{1}{a.|A_i \triangle A_j| + b.|(A_i \triangle A_j)^c|}\right)_{1\leq i,j\leq m}$

iv) $\left(\frac{|(A_i \triangle A_j)^c|}{a.|A_i \triangle A_j| + b.|(A_i \triangle A_j)^c|}\right)_{1\leq i,j\leq m}$.

Théorème 2. Soit $E = \{e_1, \ldots, e_n\}$ un ensemble non vide, et $(A_i)_{1 \leq i \leq m}$ m sous-ensembles de E. Alors la matrice

$$\left(|A_i \cap A_j||A_i^c \cap A_j^c| - |A_i \cap A_j^c||A_i^c \cap A_j|\right)_{1 \leq i,j \leq m}$$

est PSD.

Par consequent, les deux familles de matrices de similarités

i) $\left(T_\theta(A_i, A_j)\right)_{1 \leq i,j \leq m} \left(\frac{|A_i \cap A_j|}{|A_i \cap A_j| + \theta |A_i \triangle A_j|}\right)_{1 \leq i,j \leq m}$ et

ii) $\left(S_\theta(A_i, A_j)\right)_{1 \leq i,j \leq m} = \left(\frac{|(A_i \triangle A_j)^c|}{|(A_i \triangle A_j)^c| + \theta |A_i \triangle A_j|}\right)_{1 \leq i,j \leq m}$,

introduits par Gower et Legendre [36] sont désormais PSD pour $\theta \geq 1$.

De plus, 21 matrices de similarités trouvées dans la littérature (voir [23] et [131]), se sont aussi avérées PSD. Ci-dessous quelques exemples de mesures de similarité dont la matrice est PSD.

Mesure de similarité	Notation	Définition
Rogers and Tanimoto	S_{RT}	$\frac{a+d}{a+2(b+c)+d}$
Sokal and Michener	S_{SM}	$\frac{a+d}{a+b+c+d}$
Sokal and Sneath 3	S_{SS3}	$\frac{ad}{\sqrt{(a+b)(a+c)(b+d)(c+d)}}$
Rogot and Goldberg	S_{RG}	$\frac{a}{a+b+a+c} + \frac{d}{b+d+c+d}$
Hawkins and Dotson	S_{HD}	$\frac{1}{2}(\frac{a}{a+b+c} + \frac{d}{b+c+d})$
Jaccard	S_{Jac}	$\frac{a}{a+b+c}$
Gleason	S_{Gleas}	$\frac{2a}{2a+b+c}$
Ochiai	S_{Och}	$\frac{a}{\sqrt{a+b}\sqrt{a+c}}$
Braunn-Blanquet	S_{BB}	$\frac{a}{\max(a+b,a+c)}$

Il est important de noter que le chapitre 2 traite aussi la propriété semi-definie positive de certains mesures de similarité prenant des valeurs négatives. En fait, pour certaines des mesures de similarité définies dans la littérature, la contrainte de positivité est rejetée. À cette fin, une définition plus générale d'une mesure de similarité qui ne vérifie pas la condition de positivité est adoptée, dans ce qui suit.

Une fonction R de $\mathcal{X} \times \mathcal{X}$ à $\mathbb{R}$ est dite une similarité si R satisfait les conditions suivantes:

1. $R(x,y) = R(y,x)$ pour tout $x,y \in \mathcal{X}$,

2. $R(x,x) \geq 0$ pour tout $x \in \mathcal{X}$,

3. $|R(x,y)| \leq R(x,x)$ pour tout $x,y \in \mathcal{X}$.

On parle de mesures de corrélation et d'association dont les suivants sont semi-definies positives.

Mesure	**Notation**	**Definition**
Phi	R_{Phi}	$\frac{ad-bc}{\sqrt{(a+b)(a+c)(b+d)(c+d)}}$
Maxwell and Pilliner	R_{MP}	$\frac{2(ad-bc)}{(a+b)(b+d)+(a+c)(c+d)}$
Hamann [39]	R_{Ham}	$\frac{a-b-c+d}{a+b+c+d}$

Afin de distinguer entre les deux types de similarité, on a noté par S les similarités qui verifient la condition de positivité at par R celles qui peuvent prendre des valeurs négatives.

Parmi toutes les matrices de similarité vérifiées comme étant semi-définies positives, seule la matrice de similarité de Jaccard d'ordre $2^n - 1$ (i.e dont les éléments sont les valeurs de la similarité des paires de tous les sous-ensembles de E à l'exclusion de l'ensemble vide), est prouvée d'être (strictement)définie positive. Puisque chaque sous-matrice principale d'une matrice définie positive est définie positive, la matrice de Jaccard entre m sous-ensembles non vides et quelconques de E est également définie positive. Notre objective était d'étendre ce résultat à la famille Tversky des matrices de similarité, ce qui est mis en place dans le théorème suivant (voir Theorem 2.6.5).

Théorème 3. Soit E un ensemble fini non vide de dimension n. Soit $\mathbf{T}$ la matrice $(2^n - 1) \times (2^n - 1)$ de Tversky dont les éléments sont les coefficients de Tversky pour toute paire de sous-ensembles A et B de E (à l'exclusion de l'ensemble vide), définis par

$$T(A,B) = \frac{|A \cap B|}{a|A \triangle B| + b|A \cap B|}.$$

Alors $\mathbf{T}$ est définie positive pour tout entier positif $n > 1$, et pour tout $a \geq b > 0$.

Concernant les autres matrices de similarité, une condition suffisante sur les sous-ensembles de E pour lesquels toutes les matrices de similarité considérées dans notre thèse soient définies positives est établie (Theorem 2.6.2). En effet, soit S une matrice de similarité à valeurs positives (dont la mesure correspondante appartient aux exemples cités ci-dessus)

entre m sous-ensembles quelconques $(A_i)_i$ de E et tel que, pour tout $i \in \{1,2,...,m\}$, il existe x_i dans E avec $x_i \in A_i \setminus \cup_{j \neq i} A_j$. Alors S est une matrice définie positive.

Les matrices de similarité ne sont pas toutes PSD. Et dans certains cas, il est peut être difficile de prouver que des matrices de similarité particulières sont PSD. Pour cela, on peut donc vérifier si de telles matrices satisferont la propriété semi-définie 3-positive (3-PSD). On dit qu'une matrice A est semi-définie 3-positive si chaque sous-matrice d'ordre 3 est semi-définie positive. L'une des avantages de considérer un coefficient de similarité s qui est 3-PSD, est que cette condition permet d'obtenir une fonction définie à partir de s et remplissant l'inégalité triangulaire, c'est-à-dire une pseudométrique (à savoir $d_1 = \arccos(s)$ et $d_2 = \sqrt{1-s}$ [127]). Dans ce contexte, nous avons pu démontré que si $s : P(E) \times P(E) \to \mathbb{R}^+$ est un coefficient de similarité tel que $s(X,X) = k$, $\forall X \in P(E)$, pour $k > 0$ et $s(X,Y) + s(Y,Z) \leq s(X,Z) + k$, $\forall X,Y,Z \in P(E)$, alors s est semi-définie 3-positive (voir Theorem 2.7.4).

Bien entendu, il est nécessaire de mettre en évidence la motivation pour ce travail, ce qui est élaborée dans la dernière section du Chapitre 2. Nous allons mettre le point sur quelques applications de la propriété semi-définie positive, dans ce qui suit.

1. Le moyen le plus simple de déduire un coefficient de dissimilarité d à partir d'une mesure de similarité s comprise entre $[0,1]$ est d'utiliser $d = 1 - s$ ou une fonction décroissante de s. Cependant, une telle transformation n'est pas généralement une distance (métrique). De nombreux coefficients de similarité deviendront une distance, voire même une distance euclidienne , s'ils sont transformés conformément à $d = \sqrt{1-s}$. En fait, Gower et Legendre [36] ont prouvé que si S est une matrice de similarité semi-définie positive, la matrice de dissimilarité donnée par $D = \sqrt{1-S}$ est euclidienne. De telles distances euclidiennes peuvent être utilisées dans le clustering K-means [49, 61]. Les algorithmes K-means peuvent cesser de converger avec l'utilisation d'autres fonctions de distance .

2. La plupart des tâches d'apprentissage automatique reposent sur l'utilisation d'une distance métrique. Par exemple, le classifieur KNN [21] doit être fourni par une distance appropriée, à travers laquelle les points de données voisins peuvent être identifiés. La distance la plus utilisée pour comparer des données numériques est la distance euclidienne, qui suppose que chaque caractéristique des points de données est également importante et indépendante des autres. Cette hypothèse peut ne pas être toujours satisfaite dans les applications réelles. Une distance de qualité doit identifier les caractéristiques pertinentes parmi celles non pertinentes. Cependant, on peut envisager une transformation linéaire plus générale des données, définie par $d_A(x,y) = \sqrt{(x-y)^T A (x-y)}$ où $x,y \in \mathbb{R}^m$ et $A \in \mathbb{R}^{m \times m}$ est une matrice semi-

définie positive. Par exemple, en statistiques, une méthode efficace pour déterminer la similarité entre des ensembles de données est la distance de Mahalanobis. Cette distance ($d_{\sigma^{-1}}$), grâce à l'inverse de la matrice de covariance σ, prend en compte la corrélation entre les variables. Elle peut être généralisée en remplaçant la matrice de convariance par une autre, dont le rôle est d'intégrer les informations sur les ensembles de l'étude [136]. Pour définir ce type de métriques (respectivement pseudométriques), nous avons besoin de matrices PD (respectivement de matrices PSD). Les matrices de similarité sont de très bons candidats pour cela.

3. Outre les applications évidentes des matrices de similarité PSD sous la forme de matrices de covariance en apprentissage automatique, elles peuvent également être utilisées comme noyaux, dans les méthodes Kernel (dont le modèle le plus connu est Support vector machines, ou SVM) [20, 110]) largement utilisés pour les données non structurées ainsi que pour les données structurées telles que les graphes. La semi-définition positive des noyaux est essentielle pour garantir la convergence des algorithmes, tels que les SVM.

Les résultats de cette première partie du manuscrit ont fait l'objet d'une première publication dans [86].

Les racines p-ième des matrices semi-définies positives doublement stochastiques

Dans la deuxième partie de la thèse, nous avons combiné la propriété semi-définie positive et la propriété doublement stochastique dans une matrice A afin d'examiner la solution de l'équation $X^p = A$, où p est un entier positif.

Le calcul des fonctions matricielles peut être considéré comme l'un des problèmes les plus intéressants de l'analyse matricielle. En particulier, trouver les racines de certaines classes de matrices, notemment des matrices non-négatives est largement étudié (voir par exemple, [73, 74, 89, 125]). Récemment, la recherche des racines p-ième d'une matrice stochastique a été discutée dans [37, 38] étant donné que le calcul de telles racines a conduit à de nombreuses applications dans le domaine des mathématiques financières [46, 132]. En fait, les matrices stochastiques apparaissent dans les modèles de chaîne de Markov. Généralement, une matrice de transition (stochastique) sur un intervalle de temps est nécessaire, mais seule une matrice stochastique sur un intervalle de temps plus long est disponible. Une telle matrice de transition peut être obtenue en calculant la p-ième racineb de la matrice originale [17]. Nous avons étendu cette analyse pour les matrices PSD doublement stochastiques.

En effet, pour chaque $A \succeq 0$ et pour un entier positif p, il existe une unique matrice $X \succeq 0$ tel que $X^p = A$ [43]. La matrice X est appelée alors la p-ième racine semi-définie positive de A et est notée $A^{1/p}$. Alors que le calcul des racines p-ième d'une matrice stochastique implique beaucoup d'interrogations [41], comme dans quelles conditions une matrice stochastique donnée a-t-elle une p-ième racine? combien y a-t-il de racines? et comment elles peuvent être calculées, notre problème d'intérêt ici est plus restrictif, puisque nous nous intéressons à la racine unique semi-définie positive doublement stochastique d'une matrice semi-définie positive doublement stochastique. De plus, une telle matrice est toujours doublement quasi-stochastique [77], le problème est donc réduit à étudier la non-négativité de la matrice.

Dans un premier temps, deux théorèmes (voir Theorem 3.2.3 et Theorem 3.2.4) caractérisant de nouvelles conditions suffisantes pour que la racine carrée d'une matrice $n \times n$ semi-définie positive bistochastique soit non-négative, sont mises en place:

Théorème 4. Soit $A = (a_{ij})_{1 \leq i,j \leq n}$ une matrice semi-définie positive doublement stochastique avec au moins $(n-1)$ éléments diagonaux $\leqslant \frac{1}{n-1}$. Alors, il existe une matrice doublement stochastique $B = (b_{ij})$ telle que $B^2 = A$.

Ce théorème est en fait un raffinement du théorème par Marcus et Minc [67]. Un autre théorème donne une condition suffisante sur la trace de la matrice.

Théorème 5. Soit $A = (a_{ij})$ une matrice semi-définie positive doublement stochastique de taille n telle que $Tr(A) \leqslant \frac{n^2-n+2}{n^2-n}$. Alors, il existe une matrice doublement stochastique B telle que $B^2 = A$

Ensuite, si nous désignons par K_n, l'ensemble convexe de toutes les matrices de taille n semi-définies positives doublement stochastiques et pour tout entier positif p, nous définissons

$$K_n^{1/p} := \{A \in K_n : A^{1/p} \text{est doublement stochastique}\},$$

nous pouvons commencé l'étude géomètrique de l'ensemble mis en question, qui est $K_n^{1/p}$.

Dans ce contexte, on a démontré (voir Theorem 3.3.6) que pour tous entiers positifs p et n, on a :

1. K_n est convexe pour tout $n \geq 1$.
2. $K_2^{1/p}$ est convexe.
3. Pour $n \geq 3$, $K_n^{1/p}$ n'est pas convexe.

4. $K_n^{1/p}$ est étoilée par rapport à J_n, ou J_n est la matrice de taille n dont tous les éléments sont égaux à $\frac{1}{n}$.

En plus, si p,n et m sont des entiers positifs tels que $m < n$. Alors l'ensemble défini par

$$K_m^{1/p} \oplus K_{n-m}^{1/p} := \{X \oplus Y \text{ où} X \in K_m^{1/p} \text{ et} Y \in K_{n-m}^{1/p}\}$$

est un sous-ensemble de $K_n^{1/p}$, et il est étoilé par rapport à $J_m \oplus J_{n-m}$.

Le problème de la caractérisation de racines bistochastiques des matrices appartenant à K_n est toujours ouvert. Nous n'avons pas encore donné une caractérisation complète de toutes les matrices semi-définies positives doublement stochastiques et qui, étant donné p, ont des racines d'ordre p bistochastiques. Cependant, nous avons identifié des classes de matrices semi-définies positives bistochastiques appartenant à $K_n^{1/p}$. Le théorème suivant (Theorem 3.3.8) a été élaboré pour définir un ensemble convexe qui appartient à $K_n^{1/p}$.

Théorème 6. Soient $X_1 := J_n$, $X_2 := 1 \oplus J_{n-1}$,..., $X_{n-1} := I_{n-2} \oplus J_2$, $X_n := I_n$, et soit T l'enveloppe convexe des matrices

$$\{X_1, X_2, ..., X_n\}.$$

Alors T appartient à $K_n^{1/p}$ pour tout entier positif p

Une autre famille d'éléments appartenant à $K_n^{1/p}$ est définie en utilisant la théorie des M-matrices. Une matrice $A \in \mathbb{R}^{n\times n}$ est dite une M-matrice [30] si $A = sI - B$, où B est une matrice non-négative et $s \geq \rho(B)$, le rayon spectral de B. De plus, lorsque $s > \rho(B)$, alors A est nécessairement inversible et son inverse est une matrice non-négative [7]. De plus, $A^{1/p}$ est également une M-matrice pour tous les entiers positifs p par un résultat de Fiedler et Schneider [30]. Un théorème caractérisant une condition suffisante pour avoir une p-ième racine bistochastique est donné .

Théorème 7. Si la matrice doublement stochastique semi-définie positive A est l'inverse d'une M-matrice , alors l'unique p-ième racine de A, $A^{1/p}$, est doublement stochastique pour tout entier positif p.

En utilsant ce théorème, on a défini une classe de matrices appartenant à $K_n^{1/p}$. En fait, en considérant un entier $c > 1$, et une matrice doublement stochastique B de taille n, nous avons montré que la famille de matrices définie par

$$M_c = \frac{1}{c}\left(I_n - \frac{c-1}{c}B\right)^{-1} = \frac{1}{c}\sum_{k=0}^{\infty}\left(\frac{c-1}{c}B\right)^k$$

est semi-définie positive doublement stochastique dont la p-ième racine est doublement stochastique pour tout entier positif p (voir Theorem 3.4.3).

Un outil puissant pour générer des éléments dans $K_n^{1/p}$ est l'utilisation des valeurs propres des matrices bistochastiques. En effet, si A est une matrice symétrique doublement stochastique, alors par le théorème spectrale des matrices symétriques [43], il existe une matrice orthogonale V dont la première colonne est $e_n = \frac{1}{\sqrt{n}}(1,...,1)^T$ et telle que $A = V\text{diag}(1,\lambda_2,...,\lambda_n)V^T$. Par conséquent, si A est semi-définie positive, sa p-ième racine semi-définie positive est donnée par

$$A^{1/p} = V\text{diag}(1,\lambda_2^{1/p},...,\lambda_n^{1/p})V^T.$$

Il en résulte qu'on pourrait examiner les conditions dont les valeurs propres $\{1,\lambda_2,...,\lambda_n\}$ devraient satisfaire pour que A et $A^{1/p}$ soient doublement stochastiques.

En relation avec ce dernier constat, rappelons que la méthode principale pour résoudre le problème inverse des valeurs propres des matrices symétriques doublement stochastiques (SDIEP) consiste à prendre une matrice diagonale réelle $\Lambda = \text{diag}(1,\lambda_2,...,\lambda_n)$, avec $-1 \leq \lambda_i \leq 1$ pour $i = 2,...,n$, et une matrice orthogonale V dont la première colonne est e_n, puis explorer les conditions pour lesquelles $A = V\Lambda V^T$ est doublement stochastique. En conclusion, un lien évident entre SDIEP et la construction des éléments dans $K_n^{1/p}$ a été établi.

Dans cette direction, on a considéré deux types de matrices symétriques doublement stochastiques: des matrices symétriques doublement stochastiques qui peuvent être obtenues en utilisant le concept des matrices de Soules [122] et des matrices symétriques doublement stochastiques obtenues à partir des autres matrices orthogonales qui ne sont pas des matrices de Soules. En conséquence, nous avons montré que si S est une matrice $n \times n$ de Soules dont la première colonne est e_n et si $1 \geqslant \lambda_2 \geqslant ... \geqslant \lambda_n \geqslant 0$, alors la matrice définie par $A = S\text{diag}(1,\lambda_2,...,\lambda_n)S^T$ est dans $K_n^{1/p}$ pour tout entier positif p. Cependant, pour la seconde catégorie de matrices obtenues par une matrice orthogonale qui n'est pas une matrices de Soules, il semble qu'il n'existe pas encore une méthode systématique pour examiner l'appartenance de ces matrices à $K_n^{1/p}$. Un cas particulier est établi dans la dernière section du chapitre 3 où des propositions sont mises en place afin de donner une caractérisation de $K_n^{1/p}$ pour le cas particulier de $n = 3$ et $p = 2$ (voir Theorem 3.6.4).

Soient

$$a_t = \left(-\frac{1}{2}t^2 - \frac{1}{2}t\sqrt{-3t^2+2}\right),$$

$$b_t = -a_t - \frac{1}{3},$$

$$c_t = \left(-\frac{1}{2}t^2 + \frac{1}{2}t\sqrt{-3t^2+2} \right),$$

et

$$d_t = -c_t - \frac{1}{3},$$

pour tout $t \in \left[-\frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{3} \right]$. Etant donné $1 \geqslant \lambda_2 \geqslant \lambda_3 \geqslant 0$, on a définit E_{1t}, E_{2t}, E_{3t} et E_{4t} comme suit.

1. $E_{1t} = \left\{ O_{1t}\text{diag}(1,\lambda_2,\lambda_3)O_{1t}^T,\ t \in \left]-\frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6}\right[\Big/ \frac{1}{3} + c_t\sqrt{\lambda_2} + d_t\sqrt{\lambda_3} \geq 0 \right\},$

2. $E_{2t} = \left\{ O_{1t}\text{diag}(1,\lambda_2,\lambda_3)O_{1t}^T,\ t \in \left[-\frac{\sqrt{6}}{6}, 0\right] \Big/ \frac{1}{3} + (t^2 - \frac{1}{2})\sqrt{\lambda_2} + (-t^2 + \frac{1}{6})\sqrt{\lambda_3} \geq 0 \right\},$

3. $E_{3t} = \left\{ O_{2t}\text{diag}(1,\lambda_2,\lambda_3)O_{2t}^T,\ t \in \left]\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}\right] \Big/ \frac{1}{3} + a_t\sqrt{\lambda_2} + b_t\sqrt{\lambda_3} \geq 0 \right\},$

   et,

4. $E_{4t} = \left\{ O_{2t}\text{diag}(1,\lambda_2,\lambda_3)O_{2t}^T,\ t \in \left]0, \frac{\sqrt{6}}{6}\right] \Big/ \frac{1}{3} + (t^2 - \frac{1}{2})\sqrt{\lambda_2} + (-t^2 + \frac{1}{6})\sqrt{\lambda_3} \geq 0 \right\},$

Alors on a

$$K_3^{1/2} = E_{1t} \cup E_{2t} \cup E_{3t} \cup E_{4t}.$$

Une version généralisée fonctionnant pour tout entier naturel n permet de définir des éléments dans $K_n^{1/2}$. En fait, il est impossible pour le moment de pouvoir caractériser l'ensemble des matrices orthogonales qui diagonalisent l'ensemble des matrices symétriques doublement stochastiques. Pourtant, en considérant la matrice orthogonale suivante:

$$V_n = \begin{pmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{(n-1)(n-2)}} & \cdots & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{(n-1)(n-2)}} & \cdots & \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{(n-1)(n-2)}} & \cdots & -\frac{2}{\sqrt{6}} & 0 \\ \vdots & \vdots & \vdots & & \vdots & \\ \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{(n-1)(n-2)}} & \cdots & 0 & 0 \\ \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n(n-1)}} & \frac{-(n-2)}{\sqrt{(n-1)(n-2)}} & \cdots & 0 & 0 \\ \frac{1}{\sqrt{n}} & \frac{-(n-1)}{\sqrt{n(n-1)}} & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

qui est une matrice de Soules [122], on a pu construire deux familles de matrices orthogonales V_{1t} and V_{2t} (voir Section 3.6). Par conséquent, on a caractérisé 4 sous-ensembles de $K_n^{1/2}$ pour tout entier positif n (voir Theorem 3.6.7). Etant donné $1 = \lambda_1 \geqslant \lambda_2 \geqslant ... \geqslant \lambda_n \geqslant 0$, on a

défini les ensembles F_{1t}, F_{2t}, F_{3t} et F_{4t} suivants:

$$-F_{1t} = \Bigg\{ V_{1t}\text{diag}(1,\lambda_2,...,\lambda_n)V_{1t}^T, t \in \left]-\frac{\sqrt{6}}{3},-\frac{\sqrt{6}}{6}\right[\text{ tel que } \\ \frac{1}{n}+\frac{1}{n(n-1)}\sqrt{\lambda_2}+...+\frac{1}{(4)(3)}\sqrt{\lambda_{n-2}}+c_t\sqrt{\lambda_{n-1}}+d_t\sqrt{\lambda_n}\geqslant 0\Bigg\},$$

$$-F_{2t} = \Bigg\{ V_{1t}\text{diag}(1,\lambda_2,...,\lambda_n)V_{1t}^T, t \in \left[-\frac{\sqrt{6}}{6},0\right] \text{ tel que } \\ \frac{1}{n}+\frac{1}{n(n-1)}\sqrt{\lambda_2}+...+\frac{1}{(4)(3)}\sqrt{\lambda_{n-2}}+(t^2-\frac{1}{2})\sqrt{\lambda_{n-1}}+(-t^2+\frac{1}{6})\sqrt{\lambda_n}\geq 0\Bigg\},$$

$$-F_{3t} = \Bigg\{ V_{2t}\text{diag}(1,\lambda_2,...,\lambda_n)V_{2t}^T, t \in \left]\frac{\sqrt{6}}{6},\frac{\sqrt{6}}{3}\right] \text{ tel que } \\ \frac{1}{n}+\frac{1}{n(n-1)}\sqrt{\lambda_2}+...+\frac{1}{(4)(3)}\sqrt{\lambda_{n-2}}+a_t\sqrt{\lambda_{n-1}}+b_t\sqrt{\lambda_n}\geqslant 0\Bigg\},$$

et,

$$-F_{4t} = \Bigg\{ V_{2t}\text{diag}(1,\lambda_2,..,\lambda_n)V_{2t}^T, t \in \left]0,\frac{\sqrt{6}}{6}\right] \text{ tel que } \\ \frac{1}{n}+\frac{1}{n(n-1)}\sqrt{\lambda_2}+...+\frac{1}{(4)(3)}\sqrt{\lambda_{n-2}}+(t^2-\frac{1}{2})\sqrt{\lambda_{n-1}}+(-t^2+\frac{1}{6})\sqrt{\lambda_n}\geq 0\Bigg\}.$$

Alors F_{1t}, F_{2t}, F_{3t} et F_{4t} appartiennent $K_n^{1/2}$.

Les résultats de cette deuxième partie du manuscrit ont fait l'objet d'une publication dans [87].

Le problème inverse des valeurs propres des matrices symétriques doublement stochastiques

Le lien établi entre le problème visant à trouver des racines non-négatives des matrices PSD doublement stochastiques et le problème inverse des valeurs propres, nous a naturellement

amenés à étudier le problème inverse des valeurs propres des matrices symétriques doublement stochastiques, ou en anglais, SDIEP (symmetric doubly stochastic inverse eigenvalue problem) dans la troisième partie de la thèse.

Pour l'instant, le problème SDIEP n'a été résolu que pour le cas des matrices d'ordre $n = 3$ par Perfect et Mirsky [97] et reste ouvert pour les cas $n \geq 4$. Cependant, on peut trouver dans la littérature de nombreux ouvrages sur le SDIEP, donnant plusieurs conditions suffisantes (voir [58]) . Jusqu'à présent, il existe deux méthodes principales pour résoudre le SDIEP. La première consiste à prendre une matrice diagonale $\Lambda = \text{diag}(1, \lambda_2, ..., \lambda_n)$ et une matrice non singulière P, puis à explorer les conditions pour lesquelles $A = P^{-1}\Lambda P$ est doublement stochastique. La deuxième méthode entre dans la catégorie de la construction de nouvelles matrices doublement stochastiques à partir de matrices de tailles plus petites avec des spectres connus. Notre travail dans la thèse concerne la deuxième méthode on a utilisé des constructions d'une matrice doublement stochastique à partir de deux blocs de matrices doublement stochastiques (voir Theorem 4.2.2, Lemma 4.2.3 et Lemma 4.2.4). On se limite ici à citer la construction suivante à titre d'exemple:

Soit A une matrice $n \times n$ doublement stochastique dont les valeurs propres sont données par $1, \lambda_2, ..., \lambda_n$ et soit B une matrice $m \times m$ doublement stochastique de valeurs propres $1, \mu_2, ..., \mu_m$ tel que $m \geq n$. Alors pour tout $\rho \geq 0$ et pout tout $\alpha \geq 0$ tels que ρ et α ne s'annulent pas simultanément, la matrice $(m+n) \times (m+n)$,C definie par

i) Pour $m \geq n$,

$$C = \frac{1}{\alpha + \frac{\rho m}{\sqrt{mn}}} \begin{pmatrix} \alpha A & \rho e_n e_m^T \\ \rho e_m e_n^T & (\alpha + \rho \frac{m-n}{\sqrt{mn}})B \end{pmatrix}$$

est doublement stochastique dont les valeurs propres sont

$$1, \frac{\alpha\sqrt{mn} - \rho n}{\alpha\sqrt{mn} + \rho m}, \frac{\alpha}{\alpha + \frac{\rho m}{\sqrt{mn}}}\lambda_2, ..., \frac{\alpha}{\alpha + \frac{\rho m}{\sqrt{mn}}}\lambda_n, \frac{\alpha\sqrt{mn} + \rho(m-n)}{\alpha\sqrt{mn} + \rho m}\mu_2, ..., \frac{\alpha\sqrt{mn} + \rho(m-n)}{\alpha\sqrt{mn} + \rho m}\mu_m.$$

Dans un premier temps, nous avons considéré une condition suffisante donnée dans [1] où les auteurs ont utilisé une méthode récursive de construction de matrices doublement stochastiques (similaire à la construction citée ci-dessus) et ont affirmé que leur nouvelle condition suffisante établie permet d'améliorer la condition de Soules deja existante [122].

Théorème 8. [1] Soient $n \geq 5$ et $1 \geq \lambda_2 \geq .... \geq \lambda_n \geq -1$. Si

$$\frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{\frac{n}{2} - [\frac{n+2}{4}]}{(\frac{n}{2})[\frac{n+2}{4}]}\lambda_4 + \sum_{k=1}^{[\frac{n+2}{4}]-1} \frac{\lambda_{n-4k+4}}{(k+1)k} \geq 0, \tag{1}$$

pour n pair et

$$\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_3+\sum_{k=1}^{\left[\frac{n+3}{4}\right]-1}\frac{\lambda_{n-4k+4}}{(k+1)k}\geq 0, \tag{2}$$

pour n impair, alors il existe une matrice de taille n symétrique doublement stochastique dont les valeurs propres sont $1,\lambda_2,\lambda_3,...,\lambda_n$.

Nous avons donné un contre-exemple au théorème précédent dans le cas où n est impair. En fait, la liste suivante de nombres réels $\sigma=(1,1,1,-\frac{2}{3},-\frac{2}{3},-\frac{2}{3},-1)$ satisfait les conditions du théorème précédent . Donc, par [1], il existe une matrice A de taille 7 symétrique doublement stochastique dont le spectre est σ. D'autre part, σ ne peut pas être le spectre d'une matrice non-négative car sinon, selon le théorème de Perron-Frobenius (voir [77]), il existerait une matrice non-négative réductible de spectre σ. Cependant, il n'y a aucun moyen de diviser σ en 3 sous-ensembles, chaque sous-ensemble ayant une somme non-négative. Cela conduit à une contradiction.

En outre, nous avons présenté une condition suffisante alternative (notée New condition 1) pour le cas impair en utilisant la même méthode récursive pour construire des matrices doublement stochastiques. Cependant, cette condition ne sera pas une amélioration de la condition de Soules, mais plutôt une condition indépendante comme nous l'avons montré par plusieurs observations et exemples. La condition alternative est la suivante (voir Theorem 4.3.5):

Théorème 9. Soit $n\geq 5$ un entier impair et soit $1\geq\lambda_2\geq....\geq\lambda_n\geq -1$. Si

$$\frac{1}{n}+\frac{n-1}{n(n+1)}\lambda_2+\frac{\frac{n+1}{2}-\left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_4+\sum_{k=1}^{\left[\frac{n+3}{4}\right]-1}\frac{\lambda_{n-4k+4}}{(k+1)k}\geq 0,,$$

alors il existe matrice symétrique doublement stochastique dont les valeurs propres sont $1,\lambda_2,\lambda_3,...,\lambda_n$.

Par un procédé similaire, on a pu établir une nouvelle condition suffisante indépendante de Soules (voir Theorem 4.3.9) pour SDIEP.

En fait, on peut appliquer le même processus de façon récursive pour obtenir des conditions suffisantes plus générales pour le SDIEP basées sur la division euclidienne de n par 2^k avec k entier positif. Afin de mettre en évidence ce processus récursif, nous avons élaboré la conjecture suivante.

Conjecture 1. Soient k et n deux entiers positifs tel que $n \geq 2^k$, et soit $1 \geq \lambda_2 \geq ... \geq \lambda_n \geq -1$. Pour tout $i = 1, ..., k$, on note α_{i-1} le reste de la division euclidienne de n par 2^{i-1} et par

$$r_i^{(n)} = \begin{cases} 2^{i-1} - \alpha_{i-1} & \text{if } \alpha_{i-1} \neq 0 \\ 0 & \text{if } \alpha_{i-1} = 0. \end{cases}$$

Si

$$\frac{1}{n} + \frac{n - \left[\frac{n+1}{2}\right]}{n\left[\frac{n+1}{2}\right]}\lambda_2 + \frac{\frac{n+r_2^{(n)}}{2} - \left[\frac{n+2+r_2^{(n)}}{4}\right]}{\left(\frac{n+r_2^{(n)}}{2}\right)\left[\frac{n+2+r_2^{(n)}}{4}\right]}\lambda_4 + \frac{\frac{n+r_3^{(n)}}{4} - \left[\frac{n+4+r_3^{(n)}}{8}\right]}{\left(\frac{n+r_3^{(n)}}{4}\right)\left[\frac{n+4+r_3^{(n)}}{8}\right]}\lambda_8 + ...$$

$$... + \frac{\frac{n+r_k^{(n)}}{2^{k-1}} - \left[\frac{n+2^{k-1}+r_k^{(n)}}{2^k}\right]}{\left(\frac{n+r_k^{(n)}}{2^{k-1}}\right)\left[\frac{n+2^{k-1}+r_k^{(n)}}{2^k}\right]}\lambda_{2^k} + \sum_{i=1}^{\left[\frac{n+2^{k-1}+r_k^{(n)}}{2^k}\right]-1} \frac{\lambda_{n-2^k i+2^k}}{(i+1)i} \geq 0,$$

Alors il existe une matrice $n \times n$ symétrique doublement stochastique dont les valeurs propres sont $1, \lambda_2, ..., \lambda_n$.

En conclusion, pour $k = 1$, on obtient la condition de Soules. Les cas $k = 2$ et $k = 3$, correspondent aux deux conditions suffisantes qu'on a démontré dans notre chapitre. Quant aux cas où $k > 3$, il y a pas pour l'instant un moyen systématique de procéder de la sorte. Cependant, la conjecture est illustrée par des exemples.

Le dernier chapitre du manuscrit est consacré à l'étude de SDIEP pour un type particulier de spectres, le spectre normalisé de Suleimanova. On appelle $\sigma = \{1, \lambda_2, ..., \lambda_n\} \subset \mathbb{R}$ un spectre normalisé de Suleimanova si $1 + \lambda_2 + ... + \lambda_n \geq 0$ et $0 \geq \lambda_i \geq -1$ pour tout $i = 2, ..., n$. De cette définition, la question suivante découle:

"Si σ est un spectre normalisé de Suleimanova σ est-il toujours réalisable par une matrice doublement stochastique?"

Les auteurs de [52] ont prouvé que la réponse était oui pour tous les ordres de Hadamard (c'est à dire pour $N = 2^k$) et que, dans ce cas, la matrice réalisée était symétrique. Nous avons procédé dans cette étude pour apporter des réponses pour n quelconque. Cela revient à étudier ce qu'on a appelé NS-SDIEP (normalised Suleimanova symmetric doubly stochastic inverse eigenvalue problem) qui est le problème qui cherche les spectres normalisés de Suleimanova qui sont réalisables par une matrice symétrique doublement stochastique.

Notre premier résultat montre que pour n entier impair, la réponse est négative. En fait, si n est impair, alors $\lambda = (1, 0, ..., 0, -1)$ ne peut pas être le spectre d'une matrice

$n \times n$ doublement stochastique. En plus, pour $n = 3$, nous avons défini une infinité de spectres normalisés de Suleimanova non réalisables par des matrices symétriques doublement stochastiques.

La région de $\mathbb{R}^3$ qui contient tous les spectres de Suleimanova normalisés, ordonnés de façon décroissante, réalisables par des matrices doublement stochastiques, est

$$Convcoque[(1,-1/2,-1/2),(1,0,0),(1,0,-2/3)],$$

tandis que la région de tous les spectres de Suleimanova normalisés, ordonnés par ordre décroissant, qui ne sont pas réalisables par des matrices doublement stochastiques, est

$$Convcoque[(1,-1/2,-1/2),(1,0,-1),(1,0,-2/3)] \setminus [(1,-1/2,-1/2),(1,0,-2/3)],$$

où $[(1,-1/2,-1/2),(1,0,-2/3)]$ est le segment de droite joignant $(1,-1/2,-1/2)$ à $(1,0,-2/3)$ et *Convcoque* représente une coque convexe (voir Corollary 5.2.6).

D'une autre part, en utilisant aussi une méthode récursive, on a introduit trois séquences de conditions suffisantes (M_p, S_p et N_p pour $p \geq 2$) pour que le spectre normalisé de Suleimanova soit le spectre de matrices doublement stochastiques. M_p, S_p et N_p sont données respectivement dans les théorèmes suivants (voir Theorem 5.3.3, Theorem 5.3.4 Et Theorem 5.3.5).

Théorème 10. Soient p et n deux entiers positifs tel que $p \geq 2$ et $n \geq 2^p$, et soit $0 \geq \lambda_2 \geq ... \geq \lambda_n \geq -1$. Si

$$\frac{1}{n} + \frac{1}{n(n-1)}\lambda_2 + ... + \frac{1}{2^p(2^p+1)}\lambda_{n-2^p+1} + \frac{1}{2^p}(\lambda_{n-2^p+2} + ... + \lambda_{n-1} + \lambda_n) \geq 0,$$

alors il existe une matrice $n \times n$ symétrique doublement stochastique dont les valeurs propres sont $1, \lambda_2, ..., \lambda_n$.

Théorème 11. Soient p and n deux entiers positifs tels que $p \geq 2$ et $n \geq 2^{p+1}$, et soit $0 \geq \lambda_2 \geq ... \geq \lambda_n \geq -1$. Si

$$\frac{1}{n} + \frac{n-k-1}{n(k+1)}\lambda_2 + \sum_{i=2^p}^{k} \frac{\lambda_{n-2i+2}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1} \lambda_{n-2i+2} \geq 0$$

tient avec $n = 2k+2$ pour n pair (avec la convention que $\sum_{i=2^p}^{k} \frac{\lambda_{n-2i+2}}{(i+1)i} = 0$ pour $n = 2^{p+1}$), et $n = 2k+1$ pour n impair, alors il existe une matrice $n \times n$ symétrique doublement stochastique dont les valeurs propres sont $1, \lambda_2, ..., \lambda_n$.

Théorème 12. Soient p et n deux entiers positifs tels que $p \geq 2$ et $n \geq 2^{p+2}$, et soit $0 \geq \lambda_2 \geq ... \geq \lambda_n \geq -1$. Si

$$\frac{1}{n} + \frac{1}{n}\lambda_2 + \frac{\frac{n}{2} - \left[\frac{n+2}{4}\right]}{\left(\frac{n}{2}\right)\left[\frac{n+2}{4}\right]}\lambda_4 + \sum_{i=2^p}^{\left[\frac{n+2}{4}\right]-1} \frac{\lambda_{n-4i+4}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1} \lambda_{n-4i+4} \geq 0,$$

tient pour n pair (avec la convention que $\sum_{i=2^p}^{\left[\frac{n+2}{4}\right]-1} \frac{\lambda_{n-4i+4}}{(i+1)i} = 0$ pour $n = 2^{p+2}$) et

$$\frac{1}{n} + \frac{n-1}{n(n+1)}\lambda_2 + \frac{\frac{n+1}{2} - \left[\frac{n+3}{4}\right]}{\left(\frac{n+1}{2}\right)\left[\frac{n+3}{4}\right]}\lambda_4 + \sum_{i=2^p}^{\left[\frac{n+3}{4}\right]-1} \frac{\lambda_{n-4i+4}}{(i+1)i} + \frac{1}{2^p}\sum_{i=1}^{2^p-1} \lambda_{n-4i+4} \geq 0,$$

tient pour n impair, alors il existe une matrice $n \times n$ symétrique doublement stochastique dont les valeurs propres sont $1, \lambda_2, ..., \lambda_n$.

Bien entendu, pour mettre en évidence l'utililité de ces conditions, il faut étudier les relations d'inclusion entre les éléments de M_p, S_p et N_p et les comparer aux conditions préexistantes notemment la condition de Perfect et Mirsky, notée M, la condition de Soules, noté S et la condition New condition 1 que nous avons élaboré, noté N. Dans ce contexte, un théoréme (voir Theorem 5.4.4) a été établi dont la conclusion est la suivante:

Pour un spectre normalisé de Suleimanova, on a:

- $PM \subset M_2 \subset M_3 \subset ... \subset M_p \subset M_{p+1} \subset ....$
- $S \subset S_2 \subset S_3 \subset ... \subset S_p \subset S_{p+1} \subset ....$
- $N \subset N_2 \subset N_3 \subset ... \subset N_p \subset N_{p+1} \subset ....$

Des exemples qui illustrent que ces conditions peuvent revêtir une grande importance pour NS-SDIEP, ont été mis en place. Il est utile de mentionner que la famille S_p pour $p \geq 2$ est une amélioration de la condition de Soules S pour les spectres normalisés de Suleimanova.

Finalement, on a démontré que ces dernières ne sont pas nécessaires pour qu'un spectre normalisé soit réalisable par une matrice symétrique doublement stochastique. Le spectre $\sigma = \{1, 0, 0, -\frac{1}{6}, -\frac{1}{6}, -\frac{11}{18}\}$ ne vérifie pas les conditions M, S, N et $M2$ mais il est le spectre d'une matrice symétrique doublement stochastique car il vérifie le théorème suivant que nous avons établi.

Théorème 13. Pour deux puissances non-négatives de 2, n et m avec $m \geq n$, soit $\sigma = \{1, \lambda_2, ..., \lambda_{n+m}\} \subset \mathbb{R}$, tel que $0 \geq \lambda_2 \geq ... \geq \lambda_{n+m} \geq -1$. Si il existe une permutation π de

$\{\lambda_2, ..., \lambda_{n+m}\}$ tels que $\pi(\lambda_2, ..., \lambda_{n+m}) = (\mu_2, ..., \mu_{n+m})$ et

$$\mu_2 > -\frac{n}{m},$$

$$\mu_3 + ... + \mu_{n+1} \geq -\frac{m\mu_2 + n}{m+n},$$

et

$$\mu_{n+2} + ... + \mu_{n+m} \geq -\frac{n\mu_2 + m}{m+n},$$

alors σ est réalisable par une matrice $(n+m) \times (n+m)$ symétrique doublement stochastique

Les résultats de cette partie du manuscrit ont fait l'objet d'une publication dans [88].

References

[1] I. Adeli, M. Taheri, and M. M. Moghadam. A recursive method for constructing doubly stochastic matrices and inverse eigenvalue problem. *Linear Algebra and its Applications*, 537:318–331, 2018.

[2] M. R. Anderberg. *Cluster analysis for applications: probability and mathematical statistics: a series of monographs and textbooks*, volume 19. Academic Press, Inc., New York, NY., 2014.

[3] E. Andrade, C. Manzaneda, and M. Robbiano. Realizable lists on a class of nonnegative matrices. *Linear Algebra and its Applications*, 551:36–56, 2018.

[4] R. B. Bapat, R. B. Bapat, T. Raghavan, et al. *Nonnegative matrices and applications*, volume 64. Cambridge University Press, 1997.

[5] G. Benson, A. Levy, S. Maimoni, D. Noifeld, and B. R. Shalom. LCSk: A refined similarity measure. *Theoretical Computer Science*, 638:11–26, 2016.

[6] A. Berlinet and C. Thomas-Agnan. *Reproducing kernel Hilbert spaces in probability and statistics*. Springer Science & Business Media, 2011.

[7] A. Berman and R. J. Plemmons. *Nonnegative matrices in the mathematical sciences*, volume 9. Siam, 1994.

[8] R. Bhatia. *Positive definite matrices*, volume 16. Princeton University Press, Princeton, NJ, 2009.

[9] I. Bloch and A. Bretto. Mathematical morphology on hypergraphs, application to similarity and positive kernel. *Computer Vision and Image Understanding*, 117(4):342–354, 2013.

[10] I. Bloch, A. Bretto, and A. Leborgne. Robust similarity between hypergraphs based on valuations and mathematical morphology operators. *Discrete Applied Mathematics*, 183:2–19, 2015.

[11] M. Bouchard, A.-L. Jousselme, and P.-E. Doré. A proof for the positive definiteness of the jaccard index matrix. *International Journal of Approximate Reasoning*, 54(5):615–626, 2013.

[12] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[13] J. Braun-Blanquet et al. Plant sociology. the study of plant communities. *Authorized English translation of Pflanzensoziologie*, 1932.

[14] R. A. Brualdi. Some applications of doubly stochastic matrices. *Linear Algebra and its Applications*, 107:77–100, 1988.

[15] D. Brunet, J. Vass, E. R. Vrscay, and Z. Wang. Geodesics of the structural similarity index. *Applied Mathematics Letters*, 25(11):1921–1925, 2012.

[16] D. M. Cardoso, I. Gutman, E. A. Martins, and M. Robbiano. A generalization of Fiedler's lemma and some applications. *Linear and Multilinear Algebra*, 59(8):929–942, 2011.

[17] T. Charitos, P. R. de Waal, and L. C. van der Gaag. Computing short-interval transition matrices of a discrete-time markov chain from partially observed data. *Statistics in Medicine*, 27(6):905–921, 2008.

[18] S. Chen, B. Ma, and K. Zhang. On the similarity metric and the distance metric. *Theoretical Computer Science*, 410(24-25):2365–2376, 2009.

[19] S.-S. Choi, S.-H. Cha, and C. C. Tappert. A survey of binary similarity and distance measures. *Journal of Systemics, Cybernetics and Informatics*, 8(1):43–48, 2010.

[20] M. B. Christopher. *Pattern recognition and machine learning.* Springer-Verlag New York, 2016.

[21] T. Cover and P. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1):21–27, 1967.

[22] A. B. Cruse. A note on symmetric doubly-stochastic matrices. *Discrete Mathematics*, 13(2):109–119, 1975.

[23] M. M. Deza and E. Deza. *Encyclopedia of distances*. Springer Science & Business Media, 2012.

[24] M. Doolittle. The verification of predictions. *Bulletin of the Philosophical Society of Washington*, 7:122–127, 1985.

[25] H. E. Driver and A. L. Kroeber. *Quantitative expression of cultural relationships*. University of California Press Berkeley, 1932.

[26] R. Ellard and H. Šmigoc. Connecting sufficient conditions for the symmetric nonnegative inverse eigenvalue problem. *Linear Algebra and its Applications*, 498:521–552, 2016.

[27] L. Elsner, R. Nabben, and M. Neumann. Orthogonal bases that lead to symmetric nonnegative matrices. *Linear Algebra and its Applications*, 271(1-3):323–343, 1998.

[28] D. P. Faith. Asymmetric binary similarity measures. *Oecologia*, 57(3):287–290, 1983.

[29] M. Fiedler. Eigenvalues of nonnegative symmetric matrices. *Linear Algebra and its Applications*, 9:119–142, 1974.

[30] M. Fiedler and H. Schneider. Analytic functions of m-matrices and generalizations. *Linear and Multilinear Algebra*, 13(3):185–201, 1983.

[31] J. L. Fleiss. Measuring agreement between two judges on the presence or absence of a trait. *Biometrics*, pages 651–659, 1975.

[32] S. A. Forbes. *On the local distribution of certain Illinois fishes: an essay in statistical ecology*, volume 7. Illinois State Laboratory of Natural History, 1907.

[33] G. Frobenius. *Über Matrizen aus nicht negativen Elementen*. Preussische Akademie der Wissenschaften Berlin: Sitzungsberichte der Preußischen Akademie der Wissenschaften zu Berlin. Reichsdr., 1912.

[34] H. A. Gleason. Some applications of the quadrat method. *Bulletin of the Torrey Botanical Club*, pages 21–33, 1920.

[35] J. C. Gower. A general coefficient of similarity and some of its properties. *Biometrics*, pages 857–871, 1971.

[36] J. C. Gower and P. Legendre. Metric and euclidean properties of dissimilarity coefficients. *Journal of Classification*, 3(1):5–48, 1986.

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

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

[39] U. Hamann. Merkmalsbestand und verwandtschaftsbeziehungen der farinosae: ein beitrag zum system der monokotyledonen. *Willdenowia*, pages 639–768, 1961.

[40] R. P. Hawkins and V. A. Dotson. Reliability scores that delude: An alice in wonderland trip through the misleading characteristics of interobserver agreement scores in interval recording. *E. Ramp, G. Semb (Eds.),Behavior Analysis: Areas of Research and Application, Englewood Cliffs, NJ: Prentice-Hall, Inc*, 197(5), 1973.

[41] N. J. Higham and L. Lin. On *p*th roots of stochastic matrices. *Linear Algebra and its Applications*, 435(3):448–463, 2011.

[42] M. E. Hohn. Binary coefficients: A theoretical and empirical study. *Journal of the International Association for Mathematical Geology*, 8(2):137–150, 1976.

[43] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, 1990.

[44] Z. Hubalek. Coefficients of association and similarity, based on binary (presence-absence) data: an evaluation. *Biological Reviews*, 57(4):669–689, 1982.

[45] S.-G. Hwang and S.-S. Pyo. The inverse eigenvalue problem for symmetric doubly stochastic matrices. *Linear Algebra and its Applications*, 379:77–83, 2004.

[46] R. B. Israel, J. S. Rosenthal, and J. Z. Wei. Finding generators for markov chains via empirical transition matrices, with applications to credit ratings. *Mathematical Finance*, 11(2):245–265, 2001.

[47] P. Jaccard. Étude comparative de la distribution florale dans une portion des alpes et des jura. *Bull Soc Vaudoise Sci Nat*, 37:547–579, 1901.

[48] D. A. Jackson, K. M. Somers, and H. H. Harvey. Similarity coefficients: measures of co-occurrence and association or simply measures of occurrence ? *The American Naturalist*, 133(3):436–453, 1989.

[49] A. K. Jain. Data clustering: 50 years beyond k-means. *Pattern Recognition Letters*, 31(8):651–666, 2010.

[50] C. R. Johnson. Row stochastic matrices similar to doubly stochastic matrices. *Linear and Multilinear Algebra*, 10(2):113–130, 1981.

[51] C. R. Johnson, C. Marijuán, P. Paparella, and M. Pisonero. The NIEP. In *Operator Theory, Operator Algebras, and Matrix Theory*, pages 199–220. Springer, 2018.

[52] C. R. Johnson and P. Paparella. Perron spectratopes and the real nonnegative inverse eigenvalue problem. *Linear Algebra and its Applications*, 493:281–300, 2016.

[53] I. Kaddoura and B. Mourad. On a conjecture concerning the inverse eigenvalue problem of 4×4 symmetric doubly stochastic matrices. In *International Mathematical Forum*, volume 3, pages 1513–1519. Citeseer, 2008.

[54] M. Katz. On the extreme points of a certain convex polytope. *Journal of Combinatorial Theory*, 8(4):417–423, 1970.

[55] K. Kim, N. Ormes, and F. Roush. The spectra of nonnegative integer matrices via formal power series. *Journal of the American Mathematical Society*, 13(4):773–806, 2000.

[56] T. Laffey and E. Meehan. A refinement of an inequality of Johnson, Loewy and London on nonnegative matrices and some applications. *ELA. The Electronic Journal of Linear Algebra [electronic only]*, 3, 01 1998.

[57] T. J. Laffey. Realizing matrices in the nonnegative inverse eigenvalue problem. *Matrices and Group Representations (Coimbra, 1998)*, 19:21–32, 1999.

[58] Y.-J. Lei, W.-R. Xu, Y. Lu, Y.-R. Niu, and X.-M. Gu. On the symmetric doubly stochastic inverse eigenvalue problem. *Linear Algebra and its Applications*, 445:181–205, 2014.

[59] M.-J. Lesot, M. Rifqi, and H. Benhadda. Similarity measures for binary and numerical data: a survey. *International Journal of Knowledge Engineering and Soft Data Paradigms*, 1(1):63–84, 2008.

[60] L. Lin. *Roots of stochastic matrices and fractional matrix powers*. PhD thesis, Manchester Institute for Mathematical Sciences, The University of Manchester, 2011.

[61] S. Lloyd. Least squares quantization in pcm. *IEEE Transactions on Information Theory*, 28(2):129–137, 1982.

[62] J. E. Loevinger. A systematic approach to the construction and evaluation of tests of ability. *Psychological monographs*, 61(4):i, 1947.

[63] R. Loewy. A note on the real nonnegative inverse eigenvalue problem. *Electronic Journal of Linear Algebra*, 31(1):765–773, 2016.

[64] R. Loewy and D. London. A note on an inverse problem for nonnegative matrices. *Linear and Multilinear Algebra*, 6(1):83–90, 1978.

[65] D. J. MacKay and D. J. Mac Kay. *Information theory, inference and learning algorithms*. Cambridge University press, 2003.

[66] C. Manzaneda, E. Andrade, and M. Robbiano. Realizable lists via the spectra of structured matrices. *Linear Algebra and its Applications*, 534:51–72, 2017.

[67] M. Marcus and H. Minc. Some results on doubly stochastic matrices. *Proceedings of the American Mathematical Society*, 13(4):571–579, 1962.

[68] C. Marijuán and M. Pisonero. On sufficient conditions for the RNIEP and their margins of realizability. *Electronic Notes in Discrete Mathematics*, 46:201–208, 2014.

[69] C. Marijuan, M. Pisonero, and R. Soto. A map of sufficient conditions for the real nonnegative inverse eigenvalue problem. *Linear Algebra and its Applications*, 426:690–705, 10 2007.

[70] C. Marijuán, M. Pisonero, and R. L. Soto. A map of sufficient conditions for the symmetric nonnegative inverse eigenvalue problem. *Linear Algebra and its Applications*, 530:344–365, 2017.

[71] A. Maxwell and A. Pilliner. Deriving coefficients of reliability and agreement for ratings. *British Journal of Mathematical and Statistical Psychology*, 21(1):105–116, 1968.

[72] B. H. McConnaughey. *The determination and analysis of plankton communities*. Lembaga Penelitian Laut, 1964.

[73] J. J. McDonald and P. Paparella. Matrix roots of imprimitive irreducible nonnegative matrices. *Linear Algebra and Its Applications*, 498:244–261, 2016.

[74] J. J. McDonald, P. Paparella, and M. J. Tsatsomeros. Matrix roots of eventually positive matrices. *Linear Algebra and Its Applications*, 456:122–137, 2014.

[75] C. D. Meyer and C. D. Wessell. Stochastic data clustering. *SIAM Journal on Matrix Analysis and Applications*, 33(4):1214–1236, 2012.

[76] H. Minc. Inverse elementary divisors problem for doubly stochastic matrices. *Linear and Multilinear Algebra*, 11(2):121–131, 1982.

[77] H. Minc. *Nonnegative matrices*. Wiley, 1988.

[78] B. Mourad. *Generalized Doubly-stochastic Matrices and Inverse Eigenvalue Problems*. University of New South Wales, 1998.

[79] B. Mourad. An inverse problem for symmetric doubly stochastic matrices. *Inverse Problems*, 19(4):821, 2003.

[80] B. Mourad. On a lie-theoretic approach to generalized doubly stochastic matrices and applications. *Linear and Multilinear Algebra*, 52(2):99–113, 2004.

[81] B. Mourad. A note on the boundary of the set where the decreasingly ordered spectra of symmetric doubly stochastic matrices lie. *Linear Algebra and its Applications*, 416(2-3):546–558, 2006.

[82] B. Mourad. On a spectral property of doubly stochastic matrices and its application to their inverse eigenvalue problem. *Linear Algebra and Its Applications*, 436(9):3400–3412, 2012.

[83] B. Mourad. Generalization of some results concerning eigenvalues of a certain class of matrices and some applications. *Linear and Multilinear Algebra*, 61(9):1234–1243, 2013.

[84] B. Mourad, H. Abbas, and M. S. Moslehian. A note on the inverse spectral problem for symmetric doubly stochastic matrices. *Linear and Multilinear Algebra*, 63(12):2537–2545, 2015.

[85] B. Mourad, H. Abbas, A. Mourad, A. Ghaddar, and I. Kaddoura. An algorithm for constructing doubly stochastic matrices for the inverse eigenvalue problem. *Linear Algebra and its Applications*, 439(5):1382–1400, 2013.

[86] R. Nader, A. Bretto, B. Mourad, and H. Abbas. On the positive semi-definite property of similarity matrices. *Theoretical Computer Science*, 755:13–28, 2019.

[87] R. Nader, B. Mourad, H. Abbas, and A. Bretto. On the positive semi-definite pth roots of positive semi-definite doubly stochastic matrices. *Linear and Multilinear Algebra (Submitted)*.

[88] R. Nader, B. Mourad, A. Bretto, and H. Abbas. A note on the real inverse spectral problem for doubly stochastic matrices. *Linear Algebra and its Applications*, 569:206–240, 2019.

[89] H. D. Noble, S. C. McClain, and R. A. Chipman. Mueller matrix roots depolarization parameters. *Applied Optics*, 51(6):735–744, 2012.

[90] A. Ochiai. Zoogeographic studies on the soleoid fishes found in japan and its neighbouring regions. *Bulletin of Japanese Society of Scientific Fisheries*, 22:526–530, 1957.

[91] C. S. Ong, X. Mary, S. Canu, and A. J. Smola. Learning with non-positive kernels. In *Proceedings of the twenty-first international conference on Machine learning*, page 81. ACM, 2004.

[92] P. Paparella. Realizing suleimanova-type spectra via permutative matrices. *Electronic Journal of Linear Algebra*, 31(1):306–312, 2016.

[93] K. Pearson. On the coefficient of racial likeness. *Biometrika*, 18(1/2):105–117, 1926.

[94] K. Pearson and D. Heron. On theories of association. *Biometrika*, 9(1/2):159–315, 1913.

[95] H. Perfect et al. Methods of constructing certain stochastic matrices. *Duke Mathematical Journal*, 20(3):395–404, 1953.

[96] H. Perfect et al. Methods of constructing certain stochastic matrices. ii. *Duke Mathematical Journal*, 22(2):305–311, 1955.

[97] H. Perfect and L. Mirsky. Spectral properties of doubly-stochastic matrices. *Monatshefte für Mathematik*, 69(1):35–57, 1965.

[98] O. Perron. Zur theorie der matrices. *Mathematische Annalen*, 64(2):248–263, 1907.

[99] R. Reams. Constructions of trace zero symmetric stochastic matrices for the inverse eigenvalue problem. *Electronic Journal of Linear Algebra*, 27(1):359, 2014.

[100] J. Recasens. On a conjecture concerning positive semi-definiteness. *Theoretical Computer Science*, 2019.

[101] D. J. Rogers and T. T. Tanimoto. A computer program for classifying plants. *Science*, 132(3434):1115–1118, 1960.

[102] E. Rogot and I. D. Goldberg. A proposed index for measuring agreement in test-retest studies. *Journal of Chronic Diseases*, 19(9):991–1006, 1966.

[103] O. Rojo and H. Rojo. Constructing symmetric nonnegative matrices via the fast fourier transform. *Computers & Mathematics with Applications*, 45(10-11):1655–1672, 2003.

[104] O. Rojo and H. Rojo. Some results on symmetric circulant matrices and on symmetric centrosymmetric matrices. *Linear Algebra and its Applications*, 392:211–233, 2004.

[105] P. F. RUSSELL, T. R. Rao, et al. On habitat and association of species of anopheline larvae in south-eastern madras. *Journal of the Malaria Institute of India*, 3(1), 1940.

[106] L. Sahu and B. R. Mohan. An improved k-means algorithm using modified cosine distance measure for document clustering using mahout with hadoop. In *Industrial and Information Systems (ICIIS), 2014 9th International Conference on*, pages 1–5. IEEE, 2014.

[107] B. Schölkopf and A. J. Smola. *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press, 2002.

[108] M. Schultz and T. Joachims. Learning a distance metric from relative comparisons. In *Advances in neural information processing systems*, pages 41–48, 2004.

[109] T. Seidl and H.-P. Kriegel. Efficient user-adaptable similarity search in large multimedia databases. In *VLDB*, volume 97, pages 506–515, 1997.

[110] J. Shawe-Taylor and N. Cristianini. *Kernel methods for pattern analysis*. Cambridge university press, 2004.

[111] G. G. Simpson. Mammals and the nature of continents. *American Journal of Science*, 241(1):1–31, 1943.

[112] A. Singh, A. Yadav, and A. Rana. K-means with three different distance metrics. *International Journal of Computer Applications*, 67(10), 2013.

[113] R. Sinkhorn. Concerning the magnitude of the entries in a doubly stochastic matrix. *Linear and Multilinear Algebra*, 10(2):107–112, 1981.

[114] R. Sinkhorn and P. Knopp. Concerning nonnegative matrices and doubly stochastic matrices. *Pacific Journal of Mathematics*, 21(2):343–348, 1967.

[115] J. R. Smith and S.-F. Chang. Automated binary texture feature sets for image retrieval. In *Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings., 1996 IEEE International Conference on*, volume 4, pages 2239–2242. IEEE, 1996.

[116] P. H. Sneath, R. R. Sokal, et al. *Numerical taxonomy. The principles and practice of numerical classification.* W.H. Freeman, 1973.

[117] R. Sokal and P. Sneath. Principles of numerical taxonomy. 1963.

[118] R. R. Sokal. A statistical method for evaluating systematic relationship. *University of Kansas Science Bulletin*, 38:1409–1438, 1958.

[119] T. Sørensen. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on danish commons. *Kongelige Danske Videnskabernes Selskab Biologiske Skrifter*, 5:1–34, 1948.

[120] T. Sorgenfrei. *Molluscan Assemblages from the Marine Middle Miocene of South Jutland and Their Environments... Dansk Sammandrag: Molluskselskaber Fra Det Marine Mellem Miocæn i Syd-Og Sønderjylland Og Deres Miljø*. 1958.

[121] R. Soto. The inverse spectrum problem for positive generalized stochastic matrices. *Computers & Mathematics with Applications*, 43(6-7):641–656, 2002.

[122] G. W. Soules. Constructing symmetric nonnegative matrices. *Linear and Multilinear Algebra*, 13(3):241–251, 1983.

[123] H. Suleımanova. Stochastic matrices with real characteristic numbers. In *Doklady Akad. Nauk SSSR (NS)*, volume 66, pages 343–345, 1949.

[124] V. Sunder and N. J. Wildberger. Actions of finite hypergroups and examples. In *Harmonic Analysis and Hypergroups*, pages 145–163. Springer, 1998.

[125] B.-S. Tam and P.-R. Huang. Nonnegative square roots of matrices. *Linear Algebra and its Applications*, 498:404–440, 2016.

[126] D. Tarlow, K. Swersky, L. Charlin, I. Sutskever, and R. Zemel. Stochastic k-neighborhood selection for supervised and unsupervised learning. In *International Conference on Machine Learning*, pages 199–207, 2013.

[127] M. S. Tomas, C. Alsina, and J. Rubio-Martinez. Pseudometrics from three-positive semidefinite similarities. *Fuzzy Sets and Systems*, 157(17):2347–2355, 2006.

[128] J. D. Tubbs. A note on binary template matching. *Pattern Recognition*, 22(4):359–365, 1989.

[129] A. Tversky. Features of similarity. *Psychological Review*, 84(4):327, 1977.

[130] F. Wang, P. Li, A. C. König, and M. Wan. Improving clustering by learning a bi-stochastic data similarity matrix. *Knowledge and Information Systems*, 32(2):351–382, 2012.

[131] M. J. Warrens et al. *Similarity coefficients for binary data: properties of coefficients, coefficient matrices, multi-way metrics and multivariate coefficients*. Psychometrics and Research Methodology Group, Leiden University Institute for Psychological Research, Faculty of Social Sciences, Leiden University, 2008.

[132] F. V. Waugh and M. E. Abel. On fractional powers of a matrix. *Journal of the American Statistical Association*, 62(319):1018–1021, 1967.

[133] K. Q. Weinberger and L. K. Saul. Distance metric learning for large margin nearest neighbor classification. *Journal of Machine Learning Research*, 10(Feb):207–244, 2009.

[134] F. Witte. Doubly stochastic matrices and sequential data association part i. In *Information Linkage Between Applied Mathematics and Industry*, pages 641–646. Elsevier, 1979.

[135] G. Wu, E. Y. Chang, and Z. Zhang. An analysis of transformation on non-positive semidefinite similarity matrix for kernel machines. In *Proceedings of the 22nd International Conference on Machine Learning*, volume 8, 2005.

[136] E. P. Xing, M. I. Jordan, S. J. Russell, and A. Y. Ng. Distance metric learning with application to clustering with side-information. In *Advances in Neural Information Processing Systems*, pages 521–528, 2003.

[137] G. U. Yule. On the association of attributes in statistics: with illustrations from the material of the childhood society, &c. *Phil. Trans. R. Soc. Lond. A*, 194(252-261):257–319, 1900.

[138] G. U. Yule. On the methods of measuring association between two attributes. *Journal of the Royal Statistical Society*, 75(6):579–652, 1912.

[139] F. Zhang. *Matrix theory: basic results and techniques*. Springer Science & Business Media, 2011.

[140] S.-x. Zhu, T.-x. Gu, and X.-p. Liu. Solving inverse eigenvalue problems via house-holder and rank-one matrices. *Linear Algebra and its Applications*, 430(1):318–334, 2009.

Title: A study concerning the positive semi-definite property for similarity matrices and for doubly stochastic matrices with some applications.

Abstract: Matrix theory has shown its importance by its wide range of applications in different fields such as statistics, machine learning, economics and signal processing. This thesis concerns three main axis related to two fundamental objects of study in matrix theory and that arise naturally in many applications, that are positive semi-definite matrices and doubly stochastic matrices. One concept which stems naturally from machine learning area and is related to the positive semi-definite property, is the one of similarity matrices. This thesis will explore the latter desirable structure for a list of similarity matrices found in the literature. Moreover, we present new results concerning the strictly positive definite and the three positive semi-definite properties of particular similarity matrices. A detailed discussion of the many applications of all these properties in various fields is also established. On the other hand, an interesting research field in matrix analysis involves the study of roots of stochastic matrices which is important in Markov chain models in finance and healthcare. We extend the analysis of this problem to positive semi-definite doubly stochastic matrices. Our contributions include some geometrical properties of the set of all positive semi-definite doubly stochastic matrices of order n with nonnegative pth roots for a given integer p, denoted by $K_n^{1/p}$. We also present methods for finding elements in $K_n^{1/p}$ by making use of the theory of M-Matrices and the symmetric doubly stochastic inverse eigenvalue problem (SDIEP), which is also of independent interest. In the context of the SDIEP, we present some new results along this line. In particular, we propose to use a recursive method on constructing doubly stochastic matrices from smaller size matrices with known spectra to obtain new independent sufficient conditions for SDIEP. Finally, we focus our attention on the realizability by a symmetric doubly stochastic matrix of normalised Suleimanova spectra which is a normalized variant of the spectra introduced by Suleimanova.

Keywords: Similarity matrices ; Positive semi-definite matrices ; Distance and dissimilarity ; Machine learning applications ; Doubly stochastic matrices ; Matrix roots ; Inverse eigenvalue problem ; Normalised Suleimanova spectra.

Titre: Une étude concernant la propriété semi-définie positive des matrices de similarité et des matrices doublement stochastiques avec certaines applications.

Résumé: La théorie des matrices s'est développée rapidement au cours des dernières décennies en raison de son large éventail d'applications et de ses nombreux liens avec différents domaines des mathématiques, de l'économie, de l'apprentissage automatique et du traitement du signal. Cette thèse concerne trois axes principaux liés à deux objets d'étude fondamentaux de la théorie des matrices et apparaissant naturellement dans de nombreuses applications, à savoir les matrices semi-définies positives et les matrices doublement stochastiques. Un concept qui découle naturellement du domaine de l'apprentissage automatique et qui est lié à la propriété semi-définie positive est celui des matrices de similarité. Cette thèse explorera la propriété semi-définie positive pour une liste de matrices de similarité trouvées dans la littérature. De plus, nous présentons de nouveaux résultats concernant les propriétés définie positive et semi-définie trois-positive de certains matrices de similarité. Une discussion détaillée des nombreuses applications de tous ces propriétés dans divers domaines est également établie. D'autre part, un problème récent de l'analyse matricielle implique l'étude des racines des matrices stochastiques, ce qui s'avère important dans les modèles de chaîne de Markov en finance. Nous étendons l'analyse de ce problème aux matrices doublement stochastiques semi-définies positives. Nous montrons d'abord certaines propriétés géométriques de l'ensemble de toutes les matrices semi-définies positives doublement stochastiques d'ordre n ayant la p-ième racine doublement stochastique pour un entier donné p, noté par $K_n^{1/p}$. En utilisant la théorie des M-matrices et le problème inverse des valeurs propres des matrices symétriques doublement stochastiques (SDIEP), nous présentons également quelques méthodes pour trouver des éléments appartenent à $K_n^{1/p}$. Dans le contexte du SDIEP, nous présentons quelques nouveaux résultats le long de cette ligne. En particulier, nous proposons d'utiliser une méthode récursive de construction de matrices doublement stochastiques afin d'obtenir de nouvelles conditions suffisantes indépendantes pour SDIEP. Enfin, nous concentrons notre attention sur les spectres normalisés de Suleimanova, qui constituent un cas particulier des spectres introduits par Suleimanova.

Mots clés: Matrices de similarité ; Matrices semi-définies positives ; Distance et dissimilarité ; Apprentissage automatique ; Matrices doublement stochastiques ; Racines d'une matrice ; Le problème inverse des valeurs propres ; Spectre de Suleimanova.