Available online at www.sciencedirect.com
SCIENCE DIRECT®

LINEAR ALGEBRA AND ITS APPLICATIONS

ELSEVIER Linear Algebra and its Applications 403 (2005) 343–368
www.elsevier.com/locate/laa

Weak matrix majorization

Francisco D. Martínez Pería [a,*], Pedro G. Massey [a], Luis E. Silvestre [b]

[a] *Departamento de Matemática, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, CC 172, La Plata, Argentina*
[b] *Department of Mathematics, University of Texas at Austin, Austin, USA*

Received 19 May 2004; accepted 10 February 2005
Available online 31 March 2005
Submitted by R.A. Brualdi

Abstract

Given $X, Y \in \mathbb{R}^{n\times m}$ we introduce the following notion of matrix majorization, called *weak matrix majorization*,

$$X \succ_w Y \quad \text{if there exists a row-stochastic matrix } A \in \mathbb{R}^{n\times n} \text{ such that } AX = Y,$$

and consider the relations between this concept, strong majorization ($\succ_s$) and directional majorization ($\succ$). It is verified that $\succ_s \Rightarrow \succ \Rightarrow \succ_w$, but none of the reciprocal implications is true. Nevertheless, we study the implications $\succ_w \Rightarrow \succ_s$ and $\succ \Rightarrow \succ_s$ under additional hypotheses. We give characterizations of strong, directional and weak matrix majorization in terms of convexity.

We also introduce definitions for majorization between Abelian families of selfadjoint matrices, called *joint majorizations*. They are induced by the previously mentioned matrix majorizations. We obtain descriptions of these relations using convexity arguments.

AMS classification: Primary 15A51; 15A60; 15A45

Keywords: Multivariate and directional matrix majorizations; Row stochastic matrices; Mutually commuting selfadjoint matrices; Convex sets and functions

* Corresponding author.
E-mail addresses: francisco@mate.unlp.edu.ar (F.D. Martínez Pería), massey@mate.unlp.edu.ar (P.G. Massey), lsilvest@mail.ma.utexas.edu (L.E. Silvestre).

0024-3795/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2005.02.003

1. Introduction

Vector majorization in $\mathbb{R}^n$ has been widely applied both in different branches of mathematics (matrix analysis, statistics) and in other sciences like physics and economics. Also, different notions of matrix majorization between real $n \times m$ matrices have been considered in e.g., [11,12], Marshall and Olkin's classical book on majorization [10] and the recent papers [4,5,7,9]. Among them, we are interested in strong ($\succ_s$) and directional ($\succ$) majorization (see Remark 3.1 for some comments on the terminology). Given $X, Y \in M_{n,m}$ (the vector space of $n \times m$ real matrices) $X \succ_s Y$ if there exists a doubly-stochastic matrix $D \in \mathbb{R}^{n\times n}$ such that $DX = Y$; and $X \succ Y$ if the vector Xv majorizes $Yv \in \mathbb{R}^n$ for every $v \in \mathbb{R}^m$. In [5], Dahl gave a different concept of matrix majorization. For two matrices X and Y having m rows, X majorizes Y (in Dahl's sense) if there is a row-stochastic matrix A such that $XA = Y$. In Section 3 we introduce another related concept, *weak matrix majorization*: given $X, Y \in M_{n,m}$

$$X \succ_w Y \quad \text{if there exists a row-stochastic matrix } A \in \mathbb{R}^{n\times n} \text{ such that } AX = Y.$$

Although our definition of weak matrix majorization resembles to Dahl's majorization, they are quite different concepts. The main purpose of this work is to investigate the following items:

1.1. Describe weak matrix majorization and relate it with directional and strong matrix majorization

It turns out that weak matrix majorization has a simple geometrical interpretation. Indeed, this allows us to get an effective procedure to test the property and this is one of its advantages. It is well known that strong matrix majorization implies directional majorization; we prove that directional matrix majorization implies weak matrix majorization and give examples showing that, in general, the reciprocal implications are not true. Nevertheless, we study conditions under which these implications can be reversed; this problem has interest on its own, and has been considered in several articles, for example [7,11,12]. These issues are considered along Sections 3.1 and 3.3.

1.2. Find new characterizations for directional and strong matrix majorization

In Section 3.2 we use elementary facts of convexity theory in order to obtain new characterizations of matrix majorizations. In particular, we get a simple and effective criterium to determine whether $X \succ Y$. Another description of the different matrix majorizations, involving the comparison of traces of different families of matrices, is given at the end of this section. In Section 3.4 we consider the equivalence relations associated to them and we find the minimal matrices with respect to the different matrix majorizations.

1.3. Study different possible extensions of majorization between selfadjoint matrices to families of commuting selfadjoint matrices

Let $M_n(\mathbb{C})$ be the algebra of $n \times n$ matrices with complex entries. An *Abelian family* is an ordered family of mutually commuting selfadjoint matrices in $M_n(\mathbb{C})$. In Section 4 we introduce three different majorizations between Abelian families which we call *joint majorizations*. Many of the results previously obtained in Section 3 are restated in this context and some characterizations of these relations are given in terms of convexity.

2. Preliminaries

2.1. Notations

We denote by $M_{n,m} = M_{n,m}(\mathbb{R})$ (resp. $M_n = M_n(\mathbb{R})$) the real vector space of $n \times m$ (resp. $n \times n$) matrices with real entries and $M_{n,m}(\mathbb{C})$ (resp. $M_n(\mathbb{C})$) the complex vector space of $n \times m$ (resp. $n \times n$) matrices with complex entries. $GL(n)$ denotes the group of invertible $n \times n$ matrices (with real entries) and the group of permutations of order n is denoted by $\mathbb{S}_n$.

The vectors in $\mathbb{R}^n$ (or $\mathbb{C}^n$) are considered as column vectors. Nevertheless, we sometimes describe a vector as $v = (v_1, \ldots, v_n) \in \mathbb{C}^n$. The elements of the canonical basis are denoted $e_1, \ldots, e_n \in \mathbb{R}^n$. Given $x \in \mathbb{R}^n$, $\mathbb{R}x$ denotes the real subspace spanned by x and $\mathbb{C}x$ is the complex subspace spanned by x.

For $X \in M_{n,m}$, $R_i(X)$ (or shortly, X_i) denotes the ith row of X and $C_i(X)$ denotes the ith column of X. Also we will consider the sets of rows and columns of X

$$R(X) = \{R_i(X) : i = 1, \ldots, n\} \quad \text{and} \quad C(X) = \{C_i(X) : i = 1, \ldots, m\}.$$

Given $X \in M_{n,m}(\mathbb{C})$, $X^{\mathrm{t}} \in M_{m,n}(\mathbb{C})$ denotes its transpose, $X^* \in M_{m,n}(\mathbb{C})$ denotes its adjoint and $X^\dagger \in M_{m,n}(\mathbb{C})$ is the *Moore–Penrose* pseudoinverse of X. The dimension of the range of X is noted rank (X).

Given $S \subseteq \mathbb{R}^n$ we denote by $\mathrm{co}(S)$ the convex hull of S, i.e. the set of convex combinations of elements of S. We shall use the following terminology: the convex hull of a finite number of points in $\mathbb{R}^n$ is called a *polytope*. A polytope generated by affinely independent points is called a *simplex*.

If $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbb{C}^n$ then $\langle x, y\rangle$ denotes their inner product i.e, $\langle x, y\rangle = \sum_{i=1}^n x_i \overline{y_i}$. Given $A \in M_n(\mathbb{C})$, we say A is positive semidefinite if $\langle Ax, x\rangle \geqslant 0$ for every $x \in \mathbb{C}^n$. The canonical trace in $M_n(\mathbb{C})$ is denoted by tr.

If $v_1, \ldots, v_k \in \mathbb{C}^n$ then we denote by $v_1 \wedge \ldots \wedge v_k \in \bigwedge^k \mathbb{C}^n$ their antisymmetric product. Given $A \in M_n(\mathbb{C})$, denote $\bigwedge^k A$ the kth antisymmetric power of A. It is well known that $(\bigwedge^k A)(\bigwedge^k B) = \bigwedge^k (AB)$ and $(\bigwedge^k A)^* = \bigwedge^k (A^*)$ for $A, B \in M_n(\mathbb{C})$ (see for example [3]).

2.2. Nonnegative matrices

Let $A = (a_{ij}) \in M_{n,m}$. We say that A is nonnegative (resp. positive) if every $a_{ij} \geqslant 0$ (resp. $a_{ij} > 0$) and denote it $A \geqslant 0$ (resp. $A > 0$). Notice that the condition "A is nonnegative" is quite different to "A is positive semidefinite".

A nonnegative matrix $A \in M_n$ with the property that all its row sums are 1 is said to be *row-stochastic*. If we denote by $e \in \mathbb{R}^n$ the vector with all components 1, the set of row-stochastic matrices in M_n is the polytope characterized by

$$RS(n) = \{A \in M_n : A \geqslant 0, Ae = e\}.$$

A row-stochastic matrix $A \in M_n$ with the property that A^t is also row-stochastic is said to be *doubly-stochastic*. The set of doubly-stochastic matrices is also a polytope in M_n and is characterized by

$$DS(n) = \{D \in M_n : D \geqslant 0, De = e, D^t e = e\}.$$

The group of permutation matrices in M_n is contained in $DS(n)$. Birkhoff's theorem shows that these are the extremal points of the set of doubly-stochastic matrices.

Theorem (Birkhoff). *$D \in M_n$ is a doubly-stochastic matrix if and only if, for some $k \in \mathbb{N}$, there are permutation matrices $P_1, \ldots, P_k \in M_n$ and nonnegative scalars $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ such that $\alpha_1 + \cdots + \alpha_k = 1$ and*

$$D = \sum_{j=1}^{k} \alpha_j P_j.$$

2.3. Vector majorization

If $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, denote by $x^\uparrow$ and $x^\downarrow$ the vectors obtained by rearranging the entries of x in increasing and decreasing order, respectively. Given two vectors $x, y \in \mathbb{R}^n$, we say that x *mayorizes* y, and denote it $x \succ y$, if

$$\sum_{i=1}^{k} x_i^\downarrow \geqslant \sum_{i=1}^{k} y_i^\downarrow \quad k = 1, \ldots, n-1 \quad \text{and} \quad \sum_{k=1}^{n} x_i = \sum_{k=1}^{n} y_i. \tag{2.1}$$

The next theorem shows some known characterizations of vector majorization (see, for example, Bhatia's book [3]). Recall that a function $f : \mathbb{R}^n \to \mathbb{R}$ is *symmetric* (*or permutation invariant*) if $f(x) = f(Px)$ for every $x \in \mathbb{R}^n$ and every $n \times n$ permutation matrix P.

Theorem (P1). *Let $x, y \in \mathbb{R}^n$. The following are equivalent*:

1. $x \succ y$;
2. *For every convex symmetric function $f : \mathbb{R}^n \to \mathbb{R}$ we have $f(x) \geqslant f(y)$*;

3. *y belongs to the convex hull of the vectors obtained by permuting the entries of x;*
4. *There exists a doubly-stochastic $n \times n$ matrix D such that $y = Dx$.*

3. Matrix majorizations

Given two matrices $X, Y \in M_{n,m}$ we consider the following definitions of matrix majorization:

- Y is *strongly majorized* by X, denoted $X \succ_s Y$, if there exists $D \in DS(n)$ such that $DX = Y$.
- Y is *directionally majorized* by X, denoted $X \succ Y$, if for all $v \in \mathbb{R}^m$, $Xv \succ Yv$.

Remark 3.1. In [10] Marshall and Olkin define, for matrices $X, Y \in M_{n,m}$, Y to be majorized by X if there is $D \in DS(m)$ such that $XD = Y$. This notion was latter referred to in [2,5] as *multivariate majorization*. Thus, the notion of strong majorization given above corresponds to multivariate majorization of the *transposed* matrices. In [7] although strong majorization is considered, they still call it multivariate majorization. On the other hand, directional majorization has been considered in [7,9,12], for example.

When $X, Y \in M_{n,1}$, i.e. X and Y are vectors in $\mathbb{R}^n$, strong and directional matrix majorizations coincide with vector majorization. In this case, the Schur–Horn theorem (see [6]) states that strong matrix majorization (and then also directional majorization) is equivalent to the existence of a unitary matrix $U \in M_n(\mathbb{C})$ such that $(U \circ \overline{U})^{\mathrm{t}} X = Y$, where "$\circ$" denotes the Schur matrix product. But in general, given $X, Y \in M_{n,m}$, it is well known that the existence of a doubly-stochastic matrix $D \in DS(n)$ such that $DX = Y$ does not imply the existence of a unitary matrix $U \in M_n(\mathbb{C})$ such that $(U \circ \overline{U})^{\mathrm{t}} X = Y$ (see [10, p. 431]).

3.1. Weak matrix majorization

We introduce the following notion of matrix majorization.

Definition. Given two matrices $X, Y \in M_{n,m}$ we say that Y is *weakly majorized* by X, and write $X \succ_w Y$, if there exists $A \in RS(n)$ such that $AX = Y$.

We have considered square row-stochastic matrices only, but there are non-square row-stochastic matrices too. Say $A \in M_{n,m}$ is row-stochastic if A is nonnegative and all its row sums equal 1. Although we will not consider it in the rest of the paper, the definition of weak majorization can be extended to pairs of matrices with the same number of columns but different number of rows as follows: let $X \in M_{n,m}$ and $Y \in M_{p,m}$ then $X \succ_w Y$ if there exists a row-stochastic matrix $A \in M_{p,n}$ such that $AX = Y$.

Remark 3.2. Given $X, Y \in M_{n,m}$ consider the two m-tuples of vectors $(x_i)_{i=1\ldots m}$ and $(y_i)_{i=1\ldots m}$ in $\mathbb{R}^n$ defined by

$$x_i = C_i(X), \quad y_i = C_i(Y), \qquad i = 1, \ldots, m.$$

Then, it is easy to prove the following equivalences:

1. $X \succ_w Y$ if and only if there exists $A \in RS(n)$ such that $Ax_i = y_i$ for every $i = 1, \ldots, m$.
2. $X \succ Y$ if and only if $\sum_{j=1}^m \alpha_j x_j \succ \sum_{j=1}^m \alpha_j y_j$ for any m-tuple of scalars $(\alpha_1, \ldots, \alpha_m) \in \mathbb{R}^m$.
3. $X \succ_s Y$ if and only if there exists $D \in DS(n)$ such that $Dx_i = y_i$ for every $i = 1, \ldots, m$.

Therefore, each matrix majorization can be considered as a relation between the (ordered) m-tuples of column vectors $(x_i)_{i=1\ldots m}$ and $(y_i)_{i=1\ldots m}$.

It is clear that strong majorization implies directional majorization. Next we give a characterization of weak majorization and use it to prove that directional majorization implies weak majorization.

Proposition 3.3. *Let $X, Y \in M_{n,m}$. Then,*

(i) *$X \succ_w Y$ if and only if $R(Y) \subseteq \mathrm{co}(R(X))$;*
(ii) *If $X \succ Y$ then $X \succ_w Y$.*

Proof. (i) Let $X, Y \in M_{n,m}$ and $A \in M_n$. Then $AX = Y$ if and only if

$$R_i(Y) = \sum_{k=1}^n a_{ik} R_k(X), \quad i = 1, \ldots, n.$$

Therefore, if there exists $A \in RS(n)$ such that $AX = Y$ then $R(Y) \subseteq \mathrm{co}(R(X))$. On the other hand, if $R(Y) \subseteq \mathrm{co}(R(X))$ then, by the equation above, we can construct the rows of a matrix $A \in RS(n)$ such that $AX = Y$.

(ii) Let $X, Y \in M_{n,m}$ such that $X \succ Y$ and suppose that exists $1 \leqslant i \leqslant n$ such that $R_i(Y) \notin \mathrm{co}(R(X))$. Then, there exists an hyperplane which separates $R_i(Y)$ from $\mathrm{co}(R(X))$ i.e., there exist $v \in \mathbb{R}^m$ and $t > 0$ such that

$$\langle R_i(Y), v\rangle \geqslant t \quad \text{and} \quad \langle R_j(X), v\rangle < t \qquad \text{for all } j = 1, \ldots, n.$$

But this contradicts the vector majorization $Xv \succ Yv$ because

$$((Yv)^{\downarrow})_1 \geqslant (Yv)_i = \langle R_i(Y), v\rangle \geqslant t > ((Xv)^{\downarrow})_1.$$

Therefore, $X \succ_w Y$. $\square$

Remark 3.4. As a consequence of Proposition 3.3 we get an efficient method to check whether $X \succ_w Y$ holds. Indeed, by item (i), we only have to check if each row of Y can be written as a convex combination of the rows of X. For this one can solve a linear programming problem with variables being the convex weights to be found. Nevertheless, for small matrices, this can also be done using a graphic approach (see Remark 3.14).

Although the weak matrix majorization $X \succ_w Y$, for $X, Y \in M_{n,m}$, can be considered as an algebraic relation between the *columns* of X and Y (see Remark 3.2), in Proposition 3.3 we obtain a geometrical characterization of this relation in terms of the *rows* of X and Y.

The following examples show that, in general, the different matrix majorizations are not equivalent.

Example 1. $X \succ_w Y$ does not imply $X \succ Y$.

Let

$$X = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad Y = \begin{pmatrix} 1 & 0 \\ 1/2 & 1/2 \end{pmatrix}.$$

Then, if we take $A = Y \in RS(n)$, it is clear that $AX = Y$. Therefore $X \succ_w Y$. On the other hand, if we consider $v = (2, 1)$ then $Xv \not\succ Yv$. So that, $X \not\succ Y$.

Example 2. $X \succ Y$ does not imply $X \succ_s Y$.

It is a known fact. Indeed, there is an example in [11] due to Horn. Our example uses smaller matrices. Actually, we shall see in Corollary 3.22 that this is the minimum number of rows and columns required to lack the implication. Let

$$X = \begin{pmatrix} 0 & 3 & -3 & 0 \\ 0 & -2 & -2 & 4 \end{pmatrix}^t \quad \text{and} \quad Y = \begin{pmatrix} 2 & -2 & 0 & 0 \\ 0 & 0 & -2 & 2 \end{pmatrix}^t.$$

Then $X \succ Y$ but $X \not\succ_s Y$. The proof of this fact will be given in Remark 3.14.

In the next proposition we state several elementary properties of weak matrix majorization. The proof is omitted, it only requires elementary arguments.

Proposition 3.5. *Let $X, Y, Z \in M_{n,m}$. Then,*

1. $X \succ_w X$.
2. *If $X \succ_w Y$ and $Y \succ_w Z$ then $X \succ_w Z$.*
3. *If $X \succ_w Y$ then $X[I] \succ_w Y[I]$ for each $I \subset \{1, \dots, m\}$, where $X[I]$ is the submatrix of X whose columns are the columns of X indexed by the elements in I.*

4. *If* $X \succ_w Y$ *and* $R \in M_{m,p}$ *then* $XR \succ_w YR$.
5. *If* $X \succ_w Y$ *and* $P, Q \in M_n$ *are permutation matrices, then* $PX \succ_w QY$.
6. *If* $X \succ_w Y$ *then* $\operatorname{rank}(X) \geqslant \operatorname{rank}(Y)$.

Proposition 3.6. *Let* $X, Y \in M_{n,m}$ *and suppose that* $\operatorname{rank}(X) = n$. *The following are equivalent*:

(i) $X \succ_w Y$.
(ii) $YX^\dagger \in RS(n)$ *and* $\ker(X) \subseteq \ker(Y)$.

Proof. Suppose that $X \succ_w Y$, i.e. there exists $A \in RS(n)$ such that $AX = Y$. Since $\operatorname{rank}(X) = n$, $XX^\dagger = I_n$ and therefore $A = AXX^\dagger = YX^\dagger$. The equation $Y = AX$ clearly implies that $\ker(X) \subseteq \ker(Y)$.

Conversely, if $YX^\dagger \in RS(n)$ and $\ker(X) \subseteq \ker(Y)$, then $X^\dagger X$ is the orthogonal projection onto $\ker X^\perp \supseteq \ker Y^\perp$ and $YX^\dagger X = Y$. Hence $X \succ_w Y$. $\square$

Next, we consider weak matrix majorization when $X, Y \in M_n$, particularly when $X \in GL(n)$. The following corollary is a consequence of Proposition 3.6.

Corollary 3.7. *Suppose that* $X, Y \in M_n$ *and* $X \in GL(n)$. *Then,* $X \succ_w Y$ *if and only if* $YX^{-1} \in RS(n)$.

Proposition 3.8. *Let* $X, Y \in M_n$. *If* $X \succ_w Y$ *then* $|\det(X)| \geqslant |\det(Y)|$. *Moreover, if* $X \succ_w Y$ *and* $|\det(X)| = |\det(Y)| \neq 0$ *then there exists a permutation matrix* $P \in M_n$ *such that* $Y = PX$.

Proof. Let $S_X = \operatorname{co}(R(X) \cup \{0\})$ (resp. $S_Y = \operatorname{co}(R(Y) \cup \{0\})$) be the polytope generated by $R(X) \cup \{0\}$ (resp. $R(Y) \cup \{0\}$). Then, $\alpha_n|\det(X)|$ and $\alpha_n|\det(Y)|$, $\alpha_n \in \mathbb{R}^+$, are the volumes of S_X and S_Y, respectively. If $X \succ_w Y$ we have, by Proposition 3.3, that $S_Y \subseteq S_X$. Therefore $|\det(X)| \geqslant |\det(Y)|$.

Assume further that $X \succ_w Y$ and $|\det(X)| = |\det(Y)| \neq 0$. Then, using the terminology indicated in the Preliminaries, S_X and S_Y are simplexes with vertices $R(X) \cup \{0\}$ and $R(Y) \cup \{0\}$ respectively. Since $S_Y \subseteq S_X$ and $|\det(X)| = |\det(Y)| \neq 0$, they must coincide. In particular, they have the same vertices, meaning that X and Y have the same rows. $\square$

3.2. *Convexity and matrix majorizations*

We begin this section recalling a well known characterization of strong majorization in terms of convex functions. A proof of this result can be found in [5].

Theorem 3.9. *Let* $X, Y \in M_{n,m}$. *Then* $X \succ_s Y$ *if and only if, for every convex function* $f : V \to \mathbb{R}$ *we have*

$$\sum_{j=1}^{n} f(X_j) \geqslant \sum_{j=1}^{n} f(Y_j),$$

where $V \subseteq \mathbb{R}^m$ is a convex set such that $R(X) \cup R(Y) \subseteq V$.

Remark 3.10. We shall use the following elementary results about convex set and functions:

(i) given $z, w_i \in \mathbb{R}^m$ with $i = 1, \dots, n$,

$$z \in \mathrm{co}(\{w_i : i = 1, \dots, n\})$$
$$\text{if and only if } \max_{1 \leqslant i \leqslant n} \langle w_i, v\rangle \geqslant \langle z, v\rangle \quad \text{for all } v \in \mathbb{R}^m;$$

(ii) given two convex sets V_1 and V_2, $V_1 \subset V_2$ if and only if

$$\max_{x \in V_1} f(x) \leqslant \max_{x \in V_2} f(x)$$

for every convex function f defined over $V_1 \cup V_2$.

As a consequence of Proposition 3.3 and Remark 3.10 we obtain the following:

Corollary 3.11. *Let $X, Y \in M_{n,m}$. $X \succ_w Y$ if and only if*

$$\max_{1 \leqslant i \leqslant n} f(X_i) \geqslant \max_{1 \leqslant i \leqslant n} f(Y_i)$$

for every convex function $f : V \to \mathbb{R}$ where $V \subseteq \mathbb{R}^m$ is a convex set containing $R(X) \cup R(Y)$. Moreover, if we consider the linear functions $\phi_z : \mathbb{R}^m \to \mathbb{R}$, $z \in \mathbb{R}^m$ defined by $\phi_z(x) = \langle x, z\rangle$, $X \succ_w Y$ if and only if

$$\max_{1 \leqslant i \leqslant n} \phi_z(X_i) \geqslant \max_{1 \leqslant i \leqslant n} \phi_z(Y_i)$$

for every $z \in \mathbb{R}^m$.

The following theorem characterizes directional majorization between matrices in $M_{n,m}$, in terms of $\left[\frac{n}{2}\right] + 1$ polytopes, where $[r]$ is the greatest integer less than $r \in \mathbb{R}$.

Theorem 3.12. *Let $X, Y \in M_{n,m}$.$X \succ Y$ if and only if, for $k = 1, \dots, \left[\frac{n}{2}\right]$ and $k = n$, the set of averages of k different rows of Y is included in the convex hull of the set of averages of k different rows of X.*

Proof. Let $X, Y \in M_{n,m}$, and suppose that the set of averages of k different rows of Y is included in the convex hull of the set of averages of k different rows of X. Let $v \in \mathbb{R}^m$ and $1 \leqslant k \leqslant \left[\frac{n}{2}\right]$. Then, there exists a permutation $\sigma \in \mathbb{S}_n$ such that $(Yv)_{\sigma(1)} \geqslant \ldots \geqslant (Yv)_{\sigma(n)}$, where $(Yv)_i$ is the ith coordinate of $Yv \in \mathbb{R}^n$. By hypothesis, there exists a family $(c_\mu)_{\mu \in \mathbb{S}_n} \subseteq \mathbb{R}^+$ such that $\sum_{\mu \in \mathbb{S}_n} c_\mu = 1$ and

$$\frac{1}{k}\sum_{j=1}^{k} Y_{\sigma(j)} = \sum_{\mu \in \mathbb{S}_n} c_\mu \left(\frac{1}{k}\sum_{j=1}^{k} X_{\mu(j)}\right).$$

Therefore we have

$$\begin{aligned}\sum_{j=1}^{k}(Yv)_j^{\downarrow} &= k\left\langle \frac{1}{k}\sum_{j=1}^{k} Y_{\sigma(j)}, v\right\rangle = k\left\langle \sum_{\mu \in \mathbb{S}_n} c_\mu \left(\frac{1}{k}\sum_{j=1}^{k} X_{\mu(j)}\right), v\right\rangle \\ &= k\sum_{\mu \in \mathbb{S}_n} c_\mu \left\langle \frac{1}{k}\sum_{j=1}^{k} X_{\mu(j)}, v\right\rangle \leqslant k \max_{\mu \in \mathbb{S}_n}\left\langle \frac{1}{k}\sum_{j=1}^{k} X_{\mu(j)}, v\right\rangle \\ &= \sum_{j=1}^{k}(Xv)_j^{\downarrow}.\end{aligned}$$

Note that the hypothesis for $k = n$ implies

$$\sum_{j=1}^{n} Y_j = \sum_{j=1}^{n} X_j.$$

Let $\left[\frac{n}{2}\right] < k < n$, and let $\tau \in \mathbb{S}_n$ be a permutation such that $(Yv)_{\tau(1)} \leqslant \cdots \leqslant (Yv)_{\tau(n)}$. Again, by hypothesis, there exists $(d_\mu)_{\mu \in \mathbb{S}_n} \subseteq \mathbb{R}^+$, $\sum_{\mu \in \mathbb{S}_n} d_\mu = 1$, such that

$$\frac{1}{n-k}\sum_{j=1}^{n-k} Y_{\tau(j)} = \sum_{\mu \in \mathbb{S}_n} d_\mu \left(\frac{1}{n-k}\sum_{j=1}^{n-k} X_{\mu(j)}\right),$$

since $1 \leqslant n-k \leqslant \left[\frac{n}{2}\right]$. Then, we have

$$\begin{aligned}\sum_{j=1}^{k}(Yv)_j^{\downarrow} &= \left\langle \sum_{j=1}^{n} Y_j, v\right\rangle - \sum_{j=1}^{n-k}\langle Y_{\tau(j)}, v\rangle \\ &= \left\langle \sum_{j=1}^{n} X_j, v\right\rangle - (n-k)\left\langle \sum_{\mu \in \mathbb{S}_n} d_\mu \left(\frac{1}{n-k}\sum_{j=1}^{n-k} X_{\mu(j)}\right), v\right\rangle \\ &\leqslant \left\langle \sum_{j=1}^{n} X_j, v\right\rangle - (n-k)\min_{\mu \in \mathbb{S}_n}\left\langle \frac{1}{n-k}\sum_{j=1}^{n-k} X_{\mu(j)}, v\right\rangle \\ &= \sum_{j=1}^{k}(Xv)_j^{\downarrow},\end{aligned}$$

Therefore, $Xv \succ Yv$. Since $v \in \mathbb{R}^m$ was arbitrary then $X \succ Y$. On the other hand, let us suppose that $X \succ Y$. Given $\mu \in \mathbb{S}_n$,

$$\max_{\sigma \in \mathbb{S}_n} \left\langle \sum_{i=1}^{k} X_{\sigma(i)}, v \right\rangle = \sum_{i=1}^{k} (Xv)_i^{\downarrow} \geqslant \sum_{i=1}^{k} (Yv)_i^{\downarrow}$$
$$\geqslant \left\langle \sum_{i=1}^{k} Y_{\mu(i)}, v \right\rangle \quad \text{for all } v \in \mathbb{R}^n.$$

By Remark 3.10, we have that $\frac{1}{k}\sum_{i=1}^{k} Y_{\mu(i)}$ belongs to the convex hull of $\left\{\frac{1}{k}\sum_{i=1}^{k} X_{\sigma(i)} : \sigma \in \mathbb{S}_n\right\}$. □

Corollary 3.11 and Theorem 3.12 imply the following description of directional majorization in terms of weak majorization.

Corollary 3.13. *Let $X, Y \in M_{n,m}$. $X \succ Y$ if and only if $\overline{X}(k) \succ_w \overline{Y}(k)$ for $k = 1, \ldots, \left[\frac{n}{2}\right]$ and $k = n$, where $\overline{X}(k)$ (respectively $\overline{Y}(k)$) is the matrix of $\frac{n!}{k!(n-k)!}$ rows, which are all possible averages of k different rows of X (respectively of Y).*

As a consequence of Corollary 3.13 and Remark 3.4 we get an efficient way to check whether $X \succ Y$ holds. Indeed, with the notation above, we only have to check if $\overline{X}(k) \succ_w \overline{Y}(k)$ for $k = 1, \ldots, [\frac{n}{2}]$ and $k = n$ (i.e., $\left[\frac{n}{2}\right] + 1$ instances of weak matrix majorization). But given such a k, we can use linear programming (as explained in Remark 3.4) to check whether $\overline{X}(k) \succ_w \overline{Y}(k)$ holds.

Remark 3.14. Let X, Y denote the matrices in Example 2. In order to verify that $X \succ Y$, by Corollary 3.13, we only have to verify that $\overline{X}(k) \succ_w \overline{Y}(k)$ for $k = 1, 2, 4$.

In first place, $\overline{X}(4) = (0, 0) = \overline{Y}(4) \in M_{1,2}$, so that, $\overline{X}(4) \succ_w \overline{Y}(4)$. Moreover,

$$\overline{X}(2) = \begin{pmatrix} 3/2 & -3/2 & 0 & 0 & 3/2 & -3/2 \\ -1 & -1 & 2 & -2 & 1 & 1 \end{pmatrix}^{t}$$

and

$$\overline{Y}(2) = \begin{pmatrix} 0 & 1 & 1 & -1 & -1 & 0 \\ 0 & -1 & 1 & -1 & 1 & 0 \end{pmatrix}^{t}.$$

Then, the graphics in Fig. 1 show the inclusion of the polygons that prove $\overline{X}(k) \succ_w \overline{Y}(k)$ for $k = 1, 2$. Therefore $X \succ Y$.

On the other hand, the convex function $f(x, y) = \max\left\{-y, \frac{y}{2} + x, \frac{y}{2} - x\right\}$ and Theorem 3.9 show that $X \not\succ_s Y$ in Example 2.

The next theorem gives characterizations of strong, directional and weak matrix majorization comparing the traces of certain matrices.

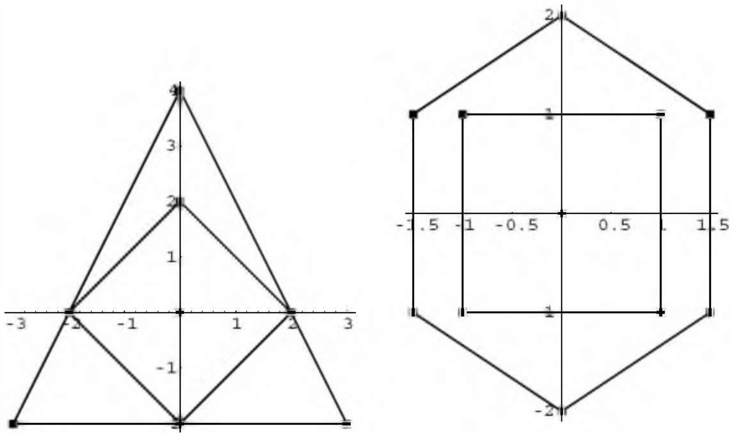

Fig. 1. Polygons corresponding to $k = 1$ and $k = 2$.

Theorem 3.15. *Let* $X, Y \in M_{n,m}$. *Then,*

1. $X \succ_s Y$ *if and only if for every* $Z \in M_{m,n}$ *there exists a permutation matrix* $P \in M_n$ *such that*
$$\operatorname{tr}(ZPX) \geqslant \operatorname{tr}(ZY).$$
2. $X \succ Y$ *if and only if, for every* $Z \in M_{m,n}$ *with* $\operatorname{rank}(Z) = 1$, *there exists a permutation matrix* $P \in M_n$ *such that*
$$\operatorname{tr}(ZPX) \geqslant \operatorname{tr}(ZY).$$
3. $X \succ_w Y$ *if and only if for every* $w \in \mathbb{R}^m$ *and every* $1 \leqslant i \leqslant n$, *there exists a permutation matrix* $P \in M_n$ *such that*
$$\operatorname{tr}(we_i^{\mathrm{t}}PX) \geqslant \operatorname{tr}(we_i^{\mathrm{t}}Y).$$

Proof. To prove 1. recall first that $M_{n,m}$ with the inner product given by $\langle X, Y\rangle = \operatorname{tr}(Y^{\mathrm{t}}X)$ can be identified with $\mathbb{R}^{n.m}$ endowed with the usual inner product.

By Birkhoff's theorem $X \succ_s Y$ is equivalent to the fact that Y belongs to the convex hull of the set $\{PX : P$ is a permutation matrix in $M_n\}$. By Remark 3.10 this is equivalent to the following: for every $Z \in M_{m,n}$ there exists a permutation matrix $P \in M_n$ such that
$$\operatorname{tr}(ZY) = \langle Y, Z^{\mathrm{t}}\rangle \leqslant \langle PX, Z^{\mathrm{t}}\rangle = \operatorname{tr}(ZPX).$$

To prove 2. note that, given $v \in \mathbb{R}^m$ then, $Xv \succ Yv$ is equivalent to the fact that Yv belongs to the convex hull of the set $\{PXv : P$ is a permutation matrix in $M_n\}$. By Remark 3.10 this is equivalent to the following: for every $w \in \mathbb{R}^n$ there exists a permutation matrix $P \in M_n$ such that $\langle PXv, w\rangle \geqslant \langle Yv, w\rangle$. Then we have
$$\operatorname{tr}(vw^{\mathrm{t}}PX) = \operatorname{tr}(w^{\mathrm{t}}PXv) = \langle PXv, w\rangle \geqslant \langle Yv, w\rangle = \operatorname{tr}(w^{\mathrm{t}}Yv) = \operatorname{tr}(vw^{\mathrm{t}}Y).$$

Since every rank one matrix $Z \in M_{m,n}$ can be expressed as $Z = vw^{\mathrm{t}}$ for $v \in \mathbb{R}^m$, $w \in \mathbb{R}^n$ we are done.

Item 3 follows in the same way. Recall that $X \succ_w Y$ is equivalent to $Y_j \in \mathrm{co}(R(X))$ for every $1 \leqslant j \leqslant n$ and note that $Y_j = Y^{\mathrm{t}} e_j$. Then, by Remark 3.10, this is equivalent to the following: for every $w \in \mathbb{R}^n$ and every $1 \leqslant j \leqslant n$ there exists a permutation matrix $Q \in M_n$ such that $\langle w, X^{\mathrm{t}} Q e_j \rangle \geqslant \langle w, Y^{\mathrm{t}} e_j \rangle$. So we have

$$\mathrm{tr}(w e_j^{\mathrm{t}} Q^{\mathrm{t}} X) = \langle w, X^{\mathrm{t}} Q e_j \rangle \geqslant \langle w, Y^{\mathrm{t}} e_j \rangle = \mathrm{tr}(w e_j^{\mathrm{t}} Y),$$

for every $w \in \mathbb{R}^n$. Taking $P = Q^{\mathrm{t}}$ we have the desired result. □

3.3. When weak majorization implies strong majorization

In this section we study conditions under which weak or directional matrix majorization implies strong matrix majorization. This problem has interest on its own, and has been considered in several articles, for example [7,11,12].

Proposition 3.16. *Let $X, Y \in M_{n,m}$ such that $X \succ Y$. Suppose that* $\mathrm{co}(R(X))$ *has only two extremal points. Then $X \succ_s Y$.*

Proof. Note that, as $\mathrm{co}(R(X))$ has only two extremal points, the points in $R(X)$ are contained in a line of $\mathbb{R}^m$. Then, the points of $R(Y)$ also belong to this line. Let $Z \in M_{n,m}$ be the matrix whose rows are all equal to $R_1(X)$. It is easy to see that $X \succ Y$ (resp. $X \succ_s Y$) if and only if $X - Z \succ Y - Z$ (resp. $X - Z \succ_s Y - Z$).

Therefore, we can suppose that $\mathrm{rank} X \leqslant 1$ and $\mathrm{rank} Y \leqslant 1$. If $X = 0$ the result is immediate. If $Y = 0$ and $\mathrm{rank} X = 1$ suppose that $Xe_1 \neq 0$ and consider the matrix $D \in DS(n)$ such that $D(Xe_1) = Ye_1 = 0$, then we have that $DX = 0 = Y$ since every column of X is a real multiple of $C_1(X) = Xe_1$. If $\mathrm{rank} Y = \mathrm{rank} X = 1$, let $x_1, y_1 \in \mathbb{R}^n$ and $x_2, y_2 \in \mathbb{R}^m$ such that $X = x_1 x_2^{\mathrm{t}}$ and $Y = y_1 y_2^{\mathrm{t}}$. Moreover, since $\mathbb{R} y_2 = \mathrm{ran}(Y^{\mathrm{t}}) = \mathrm{ran}(X^{\mathrm{t}}) = \mathbb{R} x_2$, we may assume that $y_2 = x_2$. Note that $X x_2 = \langle x_2, x_2 \rangle x_1$ and $Y x_2 = \langle x_2, x_2 \rangle y_1$.

Since $X \succ Y$, then $x_1 \succ y_1$ and there exists $D \in DS(n)$ such that $Dx_1 = y_1$. Hence

$$DX = Dx_1 x_2^{\mathrm{t}} = y_1 x_2^{\mathrm{t}} = Y$$

and $X \succ_s Y$. □

Given $X \in M_{n,m}$ we will denote $[X, e] \in M_{n,(m+1)}$ the matrix whose first (ordered) m columns are equal to those of X and its last column is the vector e. In [7], Hwang and Pyo proved the following theorem.

Theorem. *Let $X, Y \in M_{n,m}$ be such that $[Y, e][X, e]^\dagger$ has nonnegative entries. Then $X \succ Y$ if and only if $X \succ_s Y$.*

We extend this result by replacing $X \succ Y$ by $X \succ_w Y$ plus $e^t X = e^t Y$. Note that if $X \succ Y$ then $X \succ_w Y$ and $e^t X = e^t Y$ (see Corollary 3.13), but the other implication is not true (see Remark 3.19). Moreover, using the notion of weak matrix majorization we give a simpler proof.

Theorem 3.17. *Let $X, Y \in M_{n,m}$ and suppose that $[Y, e][X, e]^\dagger$ has nonnegative entries. If $X \succ_w Y$ and $e^t X = e^t Y$ then $X \succ_s Y$.*

In order to prove this theorem we are going to use the following lemma whose proof is straightforward from the definitions.

Lemma 3.18. *Let $X, Y \in M_{n,m}$ then*

$$\begin{aligned} X \succ_w Y \quad &\text{if and only if} \quad [X, e] \succ_w [Y, e], \\ X \succ Y \quad &\text{if and only if} \quad [X, e] \succ [Y, e], \\ X \succ_s Y \quad &\text{if and only if} \quad [X, e] \succ_s [Y, e], \\ e^t X = e^t Y \quad &\text{if and only if} \quad e^t[X, e] = e^t[Y, e]. \end{aligned}$$

Proof of Theorem 3.17. Let $Z = [X, e]$ and $W = [Y, e]$. Applying Lemma 3.18 we only have to prove that, if $WZ^\dagger$ has nonnegative entries, then $Z \succ_w W$ and $e^t Z = e^t W$ implies $Z \succ_s W$.

Suppose $Z \succ_w W$, then there exists a row-stochastic matrix A such that $W = AZ$. Multiplying both sides of the equation by $Z^\dagger$ we obtain:

$$WZ^\dagger = AZZ^\dagger = AP,$$

where P is the orthogonal projection onto the range of Z. Since $APZ = AZ = W$, we will conclude that $Z \succ_s W$ as soon as we prove that AP is doubly-stochastic. We know by hypothesis that $AP = WZ^\dagger$ has nonnegative entries. We are left to show that $APe = e$ and $e^t AP = e^t$. Since we chose $Z = [X, e]$, then e is in the image of Z and $Pe = e$. Therefore

$$APe = Ae = e$$

because A is row-stochastic. By hypothesis, $e^t Z = e^t W$, so

$$e^t AP = e^t WZ^\dagger = e^t ZZ^\dagger = e^t P = e^t.$$

Then AP is doubly-stochastic, $(AP)Z = W$, and by Lemma 3.18 also $(AP)X = Y$. $\square$

Remark 3.19. The condition $X \succ_w Y$ and $e^t X = e^t Y$ of Theorem 3.17 is weaker than the hypothesis $X \succ Y$ of Hwang–Pyo's theorem. In fact, let $X, Y \in M_{6,2}$ be given by

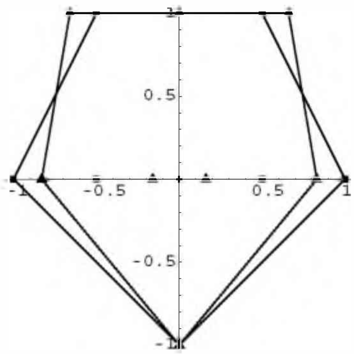

Fig. 2. $\mathrm{co}(\overline{Y}(2)) \not\subseteq \mathrm{co}(\overline{X}(2))$.

$$X = \begin{pmatrix} 0 & 1 & 1 & -1 & -1 & 0 \\ 1 & 1 & -1 & -1 & 1 & 1 \end{pmatrix}^{\mathrm{t}}$$

and

$$Y = \begin{pmatrix} 2/3 & 2/3 & 1 & -1 & -2/3 & -2/3 \\ 1 & 1 & -1 & -1 & 1 & 1 \end{pmatrix}^{\mathrm{t}}.$$

It is easy to show that $X \succ_w Y$ and $e^{\mathrm{t}}X = (0, 2) = e^{\mathrm{t}}Y$. However, Fig. 2 shows that $\overline{X}(2) \not\succ_w \overline{Y}(2)$ (where ■ represents the rows of $\overline{X}(2)$ and ▲ represents the rows of $\overline{Y}(2)$). Thus, by Corollary 3.13, $X \not\succ Y$.

The following results are consequences of Theorem 3.17.

Corollary 3.20. *Let $X, Y \in M_{n,m}$ and suppose that* $\mathrm{ran}([X, e]) = \mathbb{R}^n$. *If $X \succ_w Y$ and $e^{\mathrm{t}}X = e^{\mathrm{t}}Y$ then $X \succ_s Y$.*

Proof. It follows from Proposition 3.6 and Theorem 3.17. □

Corollary 3.21. *Let $X, Y \in M_{n,m}$ such that the rows of X are the vertices of a simplex. If $X \succ_w Y$ and $e^{\mathrm{t}}X = e^{\mathrm{t}}Y$ then $X \succ_s Y$.*

Proof. The fact that the rows of X generate a *simplex* is equivalent to the fact that the set $\{R_2(X) - R_1(X), \ldots, R_n(X) - R_1(X)\}$ is linearly independent. Then, the rank of the matrix

$$Z = \begin{pmatrix} 0 \\ R_2(X) - R_1(X) \\ \vdots \\ R_n(X) - R_1(X) \end{pmatrix}$$

is $n-1$. Therefore the subspace $\mathscr{S}$ spanned by the columns of Z has also dimension $n-1$ and $e \notin \mathscr{S}$. Using that $C_i(Z) = C_i(X) - x_{1i}e$, $1 \leqslant i \leqslant m$, we conclude that the set $\{C_1(X), \ldots, C_m(X), e\}$ span $\mathbb{R}^n$. Using Corollary 3.20, we get $X \succ_s Y$. $\square$

Corollary 3.22. *Let $X, Y \in M_{n,m}$ with $1 \leqslant n \leqslant 3$. Then, $X \succ Y$ implies $X \succ_s Y$.*

Proof. Let $X, Y \in M_{n,m}$, with $1 \leqslant n \leqslant 3$, such that $X \succ Y$. If $\mathrm{co}(R(X))$ is a segment, it follows from Proposition 3.16. Otherwise $n = 3$ and we have that $\mathrm{co}(R(X))$ is the triangle contained in $\mathbb{R}^m$ with vertices $X_i = R_i(X)$, $i = 1, 2, 3$, so we can apply Corollary 3.21. $\square$

Remark 3.23. Let $X, Y \in M_{n,m}$. Note that $X \succ Y$ is equivalent to $f(Xv) \geqslant f(Yv)$ for every $v \in \mathbb{R}^m$ and every convex symmetric function $f : \mathbb{R}^n \to \mathbb{R}$ (see Theorem (P1)). On the other hand, if we consider the convex symmetric functions $f_\infty(z_1, \ldots, z_n) = \max(z_1, \ldots, z_n)$ and $f_1(z_1, \ldots, z_n) = z_1 + \ldots + z_n$, then $X \succ_w Y$ is equivalent to $f_\infty(Xv) \geqslant f_\infty(Yv)$ for every $v \in \mathbb{R}^m$ (see item 3 of Theorem 3.15), while $e^{\mathrm{t}}X = e^{\mathrm{t}}Y$ is equivalent to $f_1(Xv) \geqslant f_1(Yv)$ for every $v \in \mathbb{R}^m$.

Assume now that $\mathrm{ran}([X, e]) = \mathbb{R}^n$. Corollary 3.20 says that if $X \succ_w Y$ and $e^{\mathrm{t}}X = e^{\mathrm{t}}Y$ then $X \succ Y$. We may re-write this result as follows: if $f_\infty(Xv) \geqslant f_\infty(Yv)$ and $f_1(Xv) \geqslant f_1(Yv)$ for every $v \in \mathbb{R}^m$ then, $f(Xv) \geqslant f(Yv)$ for every $v \in \mathbb{R}^m$ and every convex symmetric function $f : \mathbb{R}^n \to \mathbb{R}$. This reformulation of our result reminds the following theorem of interpolation theory: if $A \in M_{n,m}$ is such that $\|Av\|_\infty \leqslant \|v\|_\infty$ and $\|Av\|_1 \leqslant \|v\|_1$ for every $v \in \mathbb{R}^m$ then $\|Av\| \leqslant \|v\|$ for every $v \in \mathbb{R}^m$ and every gauge symmetric norm.

3.4. Equivalence relations associated to matrix majorizations

As we have already mentioned, matrix majorizations considered so far are preorder relations. Since $X \succ_w Y$ if and only if $R(Y) \subseteq \mathrm{co}(R(X))$, it is clear that the relation $X \succ_w Y$ and $Y \succ_w X$ is equivalent to $\mathrm{co}(R(X)) = \mathrm{co}(R(Y))$. The next theorem describes the equivalence relation associated to directional and strong matrix majorization.

Theorem 3.24. *Let $X, Y \in M_{n,m}$. Then the following are equivalent*:

(i) *There exists a permutation matrix $Q \in M_n$ such that $QX = Y$.*
(ii) *$X \succ_s Y$ and $Y \succ_s X$.*
(iii) *$X \succ Y$ and $Y \succ X$.*

Before proving this, we consider the following property of directional matrix majorization.

Lemma 3.25. *Let $X, Y \in M_{n,m}$ be such that $X \succ Y$ and $R_{i_0}(X) = R_{j_0}(Y)$. Let $\widetilde{X} \in M_{(n-1),m}$ (respectively $\widetilde{Y} \in M_{(n-1),m}$) denote the matrix obtained by deleting the i_0th row from X (respectively the j_0th row from Y). Then $\widetilde{X} \succ \widetilde{Y}$.*

Proof. It follows from the following fact (see Theorem (P1) in the preliminaries): if $x, y \in \mathbb{R}^r$ then, for every $\lambda \in \mathbb{R}$,

$$x \succ y \iff (x_1, \ldots, x_r, \lambda) \succ (y_1, \ldots, y_r, \lambda). \qquad \square$$

Proof of Theorem 3.24. The implications (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) are clear. So we only have to prove the implication (iii) $\Rightarrow$ (i). We use induction on the number of rows of X and Y. If $n = 1$ it is immediate: note that if $X, Y \in M_{1,m}$ then $X \succ Y$ implies $X = Y$.

In case that $n > 1$, note that if $X \succ Y$ and $Y \succ X$ then, $X \succ_w Y$ and $Y \succ_w X$. Therefore the convex hull of $R(X)$ coincides with that of $R(Y)$ and in particular they have the same extremal points. If z is an extremal point of $\mathrm{co}(R(X)) = \mathrm{co}(R(Y))$ then, $z = R_{i_0}(X) = R_{j_0}(Y)$ with $1 \leqslant i_0, j_0 \leqslant n$.

If $\widetilde{X}, \widetilde{Y} \in M_{(n-1),m}$ are as in the Lemma, then it holds that $\widetilde{X} \succ \widetilde{Y}$ and $\widetilde{Y} \succ \widetilde{X}$. By the inductive hypothesis, the rows of $\widetilde{X}$ are a reordering of the rows of $\widetilde{Y}$. Therefore the rows of X are a reordering of the rows of Y, which implies i). $\square$

The following Corollary is an analogue of Theorem 3.24 for weak matrix majorization, in the particular case that $X, Y \in GL(n)$. It is a consequence of Proposition 3.8.

Corollary 3.26. *Let $X, Y \in M_n$ with $Y \in GL(n)$. Then the following are equivalent*:

(i) *There exists a permutation matrix $Q \in M_n$ such that $QX = Y$.*
(ii) *$X \succ_w Y$ and $Y \succ_w X$.*

Next, we determine the minimal matrices with respect to the preorders that we have considered so far. In this context, a *minimal* element with respect to a preorder relation $\ll$ in a set P is an element $m \in P$ such that, given $n \in P$, if $n \ll m$ then $m \ll n$.

Proposition 3.27. *$X \in M_{n,m}$ is minimal with respect to any of the preorder $\succ_w$, $\succ$ or $\succ_s$ if and only if $X_1 = \cdots = X_n$, that is, all the rows of X coincide.*

Proof. If $R(X) = \{v\}$, for $v \in \mathbb{R}^m$, then $\mathrm{co}(R(X)) = \{v\}$. Then, if $X \succ_w Y$ it is clear that $X = Y$. On the other hand, let $X \in M_{n,m}$ be a matrix with at least two different rows. Then $R(X)$ contains two different points (in $\mathbb{R}^m$). If $D \in DS(n)$ is the matrix with all entries equal to $1/n$ we have that $Y = DX \prec_s X$. Moreover, since $R_1(Y) = R_2(Y) = \cdots = R_n(Y)$, then $\mathrm{co}(R(Y))$ contains only one point, so

$R(X) \not\subset \operatorname{co}(R(Y))$. Therefore $Y \not\succ_w X$ and X is not minimal with respect to any of the matrix majorizations. $\square$

4. Joint majorizations

In [1] Ando considers the majorization relation between selfadjoint matrices. Indeed, if $a, b \in H(n)$, the set of selfadjoint matrices of $M_n(\mathbb{C})$, let $\lambda(a), \lambda(b) \in \mathbb{R}^n$ denote the vectors of eigenvalues of a and b respectively, counted with multiplicity. Then *a majorizes b* (in Ando's sense) if $\lambda(a) \succ \lambda(b)$; in this case we write $a \succsim b$. Among many others, we can cite the following characterizations of majorization between selfadjoint matrices.

Theorem 4.1. *Let $a, b \in H(n)$. Then the following are equivalent*:

1. $a \succsim b$.
2. *For every convex function $f : (\alpha, \beta) \to \mathbb{R}$, such that $\sigma(a) \cup \sigma(b) \subseteq (\alpha, \beta)$, we have that* $\operatorname{tr} f(a) \geqslant \operatorname{tr} f(b)$.
3. *b belongs to the convex hull of the unitary orbit of a.*

The goal of this section is to define and characterize some possible extensions of majorization in $H(n)$, which we call *joint majorizations*. Many results in this section are based on previously obtained results about matrix majorizations.

4.1. Joint majorization between Abelian families in $H(n)$

By an *Abelian family* we mean an ordered family $(a_i)_{i=1,\ldots,m}$ of selfadjoint matrices in $M_n(\mathbb{C})$ such that

$$a_i a_j = a_j a_i, \quad i, j = 1, \ldots, m.$$

In order to introduce the joint majorizations we consider the following well known facts: if $(a_i)_{i=1,\ldots,m}$ and $(b_i)_{i=1,\ldots,m}$ are two Abelian families in $M_n(\mathbb{C})$ then, there exist unitary matrices $U, V \in M_n(\mathbb{C})$ such that

$$U^* a_i U = D_{\lambda(a_i)}, \quad V^* b_i V = D_{\lambda(b_i)}, \qquad i = 1, \ldots, m.$$

where D_x denotes the diagonal matrix with main diagonal $x \in \mathbb{R}^n$. In this case $\lambda(a_i)$ is the vector of eigenvalues corresponding to a_i, counted with multiplicity, in some order depending on U. Consider the matrices $A, B \in M_{n,m}$ given by

$$C_i(A) = \lambda(a_i), \quad C_i(B) = \lambda(b_i), \qquad i = 1, \ldots, m.$$

Definition. Let $(a_i)_{i=1,\ldots,m}, (b_i)_{i=1,\ldots,m} \subseteq M_n(\mathbb{C})$ be two Abelian families and let $A, B \in M_{n,m}$ be defined as above. We say that the family $(a_i)_{i=1,\ldots,m}$ *jointly* *weakly majorizes* (respectively *jointly strongly majorizes*, *jointly majorizes*) the family $(b_i)_{i=1,\ldots,m}$ and write

$$(a_i)_{i=1,\ldots,m} \succ_w (b_i)_{i=1,\ldots,m}$$

(respectively $(a_i)_{i=1,\ldots,m} \succ_s (b_i)_{i=1,\ldots,m}$, $(a_i)_{i=1,\ldots,m} \succ (b_i)_{i=1,\ldots,m}$) if $A \succ_w B$ (respectively $A \succ_s B$, $A \succ B$).

A few words concerning the definition are in order. First, note that if U, W are two unitary matrices that diagonalize the family $(a_i)_{i=1,\ldots,m}$ simultaneously then there exists a permutation matrix Q such that

$$U^*a_iU = Q^*W^*a_iWQ, \quad i = 1, \ldots, m.$$

Thus, if $A' \in M_{n,m}$ denotes the matrix whose columns $C_i(A')$ are the main diagonals of the matrices W^*a_iW, then $A = QA'$. That is, the definition above does not depend on the unitary U and the notions are well defined. This also shows that the set of rows $R(A)$ does not depend on the unitary U. This set is called the *joint spectrum* of the family and denoted by $\sigma(a_1, \ldots, a_m)$. Moreover, if $f : V \to \mathbb{C}$ is such that $\sigma(a_1, \ldots, a_m) \subseteq V$ then we consider

$$f(a_1, \ldots, a_m) = UD_xU^*$$

where D_x is the diagonal matrix with main diagonal $x = (f(R_1(A)), \ldots, f(R_n(A))) \in \mathbb{C}^n$. Note that $f(a_1, \ldots, a_m) \in M_n(\mathbb{C})$ does not depend on U. We say that the matrix $f(a_1, \ldots, a_m)$ is obtained from the family $(a_i)_{i=1,\ldots,m}$ by functional calculus.

From now on, whenever the context makes it clear, we shall not write the subindex corresponding to the family of matrices and simply write $(a_i) \succ_w (b_i)$ (resp. $(a_i) \succ_s (b_i)$, $(a_i) \succ (b_i)$).

Proposition 4.2. *Let $(a_i)_{i=1,\ldots,m}$ and $(b_i)_{i=1,\ldots,m}$ be two Abelian families in $M_n(\mathbb{C})$. Then*

1. $(a_i) \succ_w (b_i)$ *if and only if* $\mathrm{co}(\sigma(b_1, \ldots, b_m)) \subseteq \mathrm{co}(\sigma(a_1, \ldots, a_m))$, *where* $\mathrm{co}(S)$ *denotes the convex hull of the set* $S \subseteq \mathbb{R}^m$.
2. $(a_i) \succ (b_i)$ *if and only if, for every* $\gamma_1, \ldots, \gamma_m \in \mathbb{R}$ *it holds*

   $$\gamma_1a_1 + \cdots + \gamma_ma_m \succsim \gamma_1b_1 + \cdots + \gamma_mb_m \text{ (in Ando's sense).}$$

3. $(a_i) \succ_s (b_i)$ *if and only if there exist* $k \in \mathbb{N}$, *unitary matrices* $W_1, \ldots, W_k \in M_n(\mathbb{C})$ *and nonnegative numbers* $\mu_1, \ldots, \mu_k$, $\sum_{j=1}^k \mu_j = 1$, *such that*

   $$b_i = \sum_{j=1}^k \mu_jW_j^*a_iW_j, \quad \textit{for } 1 \leqslant i \leqslant m. \tag{4.1}$$

Proof. Items 1. and 2. are mostly consequences of the definitions, so the proof is omitted. The proof of item 3. is postponed until the proof of Theorem 4.5. □

4.2. *Characterizations of joint majorizations*

In this subsection we consider some characterizations of the different notions of joint majorization introduced so far. We begin with the following elementary lemma. Recall that a linear map $T : S \to M_n(\mathbb{C})$ from a linear subspace $S \subseteq M_n(\mathbb{C})$ is called *unital* if $T(I) = I$, where I denotes the identity matrix; *positive* if $T(a)$ is positive semidefinite whenever a is positive semidefinite, and *trace preserving* if $\operatorname{tr}(T(a)) = \operatorname{tr}(a)$ for every $a \in S$.

Lemma 4.3. *Let $\mathscr{D}$ be the diagonal algebra in $M_n(\mathbb{C})$ and let $T : \mathscr{D} \to \mathscr{D}$ be a positive unital map. Then there exists $E \in RS(n)$ such that*

$$T(D_x) = D_{Ex}. \tag{4.2}$$

If, in addition, T is trace preserving then $E \in DS(n)$. On the other hand, if T is as in (4.2) for some $E \in RS(n)$ (respectively $E \in DS(n)$), then T is a positive unital map (resp. trace preserving positive unital map).

Proof. We identify $\mathscr{D}$ with $\mathbb{C}^n$ as vector spaces by the map $D_x \mapsto x$, where D_x is the diagonal matrix with main diagonal $x \in \mathbb{C}^n$. Therefore, under this identification, T induces a linear transformation $\tilde{T}$ on $\mathbb{C}^n$ by $\tilde{T}x = \sum_{i=1}^n T(D_x)_{ii} e_i$, where $\{e_i\}_{i=1,\dots,n}$ is the canonical basis of $\mathbb{C}^n$. Let E be the matrix of $\tilde{T}$ with respect to the canonical basis. Then E satisfies $Ee = e$ and $Ex \geqslant 0$ whenever $x \geqslant 0$, where $y \geqslant 0$ means that all coordinates of $y \in \mathbb{R}^n$ are nonnegative. Therefore $E \in RS(n)$ and $T(D_x) = D_{Ex}$. Moreover, if T is trace preserving then $\operatorname{tr}(Ex) = \operatorname{tr}(x)$, where $\operatorname{tr}(y) = y_1 + \cdots + y_n$. This implies that $E \in DS(n)$. The converse is clear. □

We shall make use of the following well known result.

Lemma 4.4. *Let $\mathscr{A} \subseteq M_n(\mathbb{C})$ be a unital *-subalgebra of $M_n(\mathbb{C})$. Then there exists a trace preserving positive unital map $\Psi : M_n(\mathbb{C}) \to \mathscr{A}$ such that $\Psi(a) = a$ for all $a \in \mathscr{A}$.*

In what follows, given $(a_i)_{i=1,\dots,m} \subseteq M_n(\mathbb{C})$, $C^*(a_1, \dots, a_m)$ denotes the unital *-subalgebra generated by the a_i's, that is, the smallest unital *-subalgebra $\mathscr{A}$ of $M_n(\mathbb{C})$ such that $a_i \in \mathscr{A}$, $i = 1, \dots, m$. It is clear that if $(a_i)_{i=1,\dots,m}$ is an Abelian family in $M_n(\mathbb{C})$ and U is a unitary matrix that simultaneously diagonalizes this family then U diagonalizes any $a \in C^*(a_1, \dots, a_m)$ i.e, $U^*aU = D_x$ for some $x \in \mathbb{C}^n$. Therefore, $C^*(a_1, \dots, a_m) \subseteq U\mathscr{D}U^* = \{UD_xU^* : x \in \mathbb{C}^n\}$.

Theorem 4.5. *Let $(a_i)_{i=1,\dots,m}$ and $(b_i)_{i=1,\dots,m}$ be two Abelian families in $M_n(\mathbb{C})$. Then*

1. $(a_i) \succ_w (b_i)$ *if and only if there exists a positive unital map*
$$T : C^*(a_1, \dots, a_m) \to C^*(b_1, \dots, b_m)$$
such that $T(a_i) = b_i$ for every $i = 1, \dots, m$.
2. $(a_i) \succ (b_i)$ *if and only if, for every $k = 1, \dots, \left[\frac{n}{2}\right]$ and $k = n$ we have*
$$\left(\log\left[\bigwedge^k \exp(a_i)\right]\right)_{i=1,\dots,m} \succ_w \left(\log\left[\bigwedge^k \exp(b_i)\right]\right)_{i=1,\dots,m}.$$
3. $(a_i) \succ_s (b_i)$ *if and only if there exists a trace preserving positive unital map*
$$T : C^*(a_1, \dots, a_m) \to C^*(b_1, \dots, b_m)$$
such that $T(a_i) = b_i$ for every $i = 1, \dots, m$.

Proof. Let $U, V \in M_n(\mathbb{C})$ be unitary matrices such that
$$U^*a_iU = D_{\lambda(a_i)}, \quad V^*b_iV = D_{\lambda(b_i)}, \qquad i = 1, \dots, m,$$
where $\lambda(a_i), \lambda(b_i) \in \mathbb{R}^n$. As we have mentioned before, if $a \in C^*(a_1, \dots, a_m)$ then $U^*aU \in \mathscr{D}$.

1. Suppose there exists a positive unital map $T : C^*(a_1, \dots, a_m) \to C^*(b_1, \dots, b_m)$ such that $T(a_i) = b_i$ for every $i = 1, \dots, m$. Let $\tilde{T} : \mathscr{D} \to \mathscr{D}$ be defined by
$$\tilde{T}(D_x) = V^*T(\Psi(UD_xU^*))V,$$
where $\Psi : M_n(\mathbb{C}) \to C^*(a_1, \dots, a_m)$ is the map obtained in Lemma 4.4. Note that $\tilde{T}$ is a positive unital map such that $\tilde{T}(D_{\lambda(a_i)}) = D_{\lambda(b_i)}$, $i = 1, \dots, m$. By Lemma 4.3 there exists $E \in RS(n)$ such that $E\lambda(a_i) = \lambda(b_i)$, $i = 1, \dots, m$. Then, we conclude that $(a_i) \succ_w (b_i)$ (see Remark 3.2).

On the other hand, if $(a_i) \succ_w (b_i)$ let $E \in RS(n)$ be such that $E\lambda(a_i) = \lambda(b_i)$, $i = 1, \dots, m$. Let $T : U\mathscr{D}U^* \to C^*(b_1, \dots, b_m)$ be defined by
$$T(UD_xU^*) = \Phi(VD_{Ex}V^*),$$
where $\Phi : M_n(\mathbb{C}) \to C^*(b_1, \dots, b_m)$ is the map obtained as in Lemma 4.4. Then, the result follows considering the restriction of T to $C^*(a_1, \dots, a_m)$.

2. Note that $\bigwedge^k U$ is unitary and diagonalizes the family $(\bigwedge^k a_i)$. Let $A \in M_{n,m}$ be such that, for $1 \leqslant i \leqslant m$, $C_i(A) = \lambda(a_i)$. For $1 \leqslant k \leqslant n$, let $A_{(k)}$ be the $\frac{n!}{k!(n-k)!} \times m$ matrix such that $C_i(A_{(k)}) = \lambda(i, k)$, where
$$\bigwedge^k U^*\left(\log\left[\bigwedge^k \exp(a_i)\right]\right)\bigwedge^k U = D_{\lambda(i,k)}, \quad 1 \leqslant i \leqslant m.$$
We shall show that, for $1 \leqslant k \leqslant n$, $A_{(k)} = k \cdot \overline{A}(k)$ (up to a permutation of rows) where $\overline{A}(k)$ is as in Corollary 3.13. Let $1 \leqslant j \leqslant n$ and denote by $u_j = C_j(U)$ the

columns of U. Then, in order to compute the rows of $A_{(k)}$ we just have to note that, for every $1 \leqslant i \leqslant m$ and any choice of $1 \leqslant l_1 < \cdots < l_k \leqslant n$, $u_{l_1} \wedge \cdots \wedge u_{l_k}$ is an eigenvector of $\log \left[\bigwedge^k \exp(a_i)\right]$ corresponding to the eigenvalue

$$\log \left[\prod_{j=1}^{k} \exp(\lambda(a_i)_{l_j}) \right] = \lambda(a_i)_{l_1} + \cdots + \lambda(a_i)_{l_k},$$

where $\lambda(a_i)_{l_j}$ is the l_jth entry of the vector $\lambda(a_i)$. The equality $A_{(k)} = k \cdot \overline{A}(k)$ easily follows from this fact. Therefore, the result follows from the hypothesis ($A_{(k)} \succ_w B_{(k)}$ for $k = 1, \ldots, \left[\frac{n}{2}\right]$ and $k = n$) and Corollary 3.13.

3. The proof given for the first part of the theorem can easily be extended to prove this statement. $\square$

Example 3. Recall that a *system of projections* in $M_n(\mathbb{C})$ is a family $\{P_i\}_{i=1,\ldots,k}$ of orthogonal projections such that $\sum_{i=1}^{k} P_i = I$. Given such a family, the associated *pinching*, $\mathscr{C} : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ is given by

$$\mathscr{C}(A) = \sum_{i=1}^{k} P_i A P_i.$$

$\mathscr{C}$ is an example of a trace preserving positive unital map. In particular, if P_i is the orthogonal projection onto $\mathbb{C}e_i$, $i = 1, \ldots, n$, then the pinching associated to this system of projections is called the *diagonal pinching* and noted $\mathscr{C}_0$.

In [10, p. 331–332], Marshall and Olkin gave an example of multivariate majorization that we now rewrite in terms of strong joint majorization (in this context, it is a consequence of item 3. of Theorem 4.5):

Let $(a_i)_{i=1,\ldots,m}$ be an Abelian family in $M_n(\mathbb{C})$ and let $\mathscr{C}_0$ denote the diagonal pinching. Then, $(a_i) \succ_s (\mathscr{C}_0(a_i))$.

It is worth to notice that, given an Abelian family $(a_i)_{i=1,\ldots,m}$, the above result is not true for an arbitrary pinching $\mathscr{C}$ since the family $(\mathscr{C}(a_i))_{i=1,\ldots,m}$ may not be Abelian.

We are now going to complete the proof of Proposition 4.2.

Proof of Proposition 4.2. Assume that there exist $k \in \mathbb{N}$, unitary matrices $W_1, \ldots, W_k \in M_n(\mathbb{C})$ and nonnegative numbers $\mu_1, \ldots, \mu_k$, $\sum_{j=1}^{k} \mu_j = 1$, such that Eq. (4.1) holds. Then, we define $T : C^*(a_1, \ldots, a_m) \to C^*(b_1, \ldots, b_m)$ by

$$T(a) = \Phi\left(\sum_{j=1}^{k} \mu_j W_j^* a W_j\right),$$

where $\Phi : M_n(\mathbb{C}) \to C^*(b_1, \ldots, b_m)$ is obtained as in Lemma 4.4. It is clear that T is a trace preserving positive unital map, so by Theorem 4.5 we get $(a_i) \succ_s (b_i)$.

On the other hand, if $(a_i) \succ_s (b_i)$ let U, V, $\lambda(a_i)$, $\lambda(b_i)(1 \leqslant i \leqslant m)$ as in the proof of Theorem 4.5. Then, there exists $E \in DS(n)$ such that $E\lambda(a_i) = \lambda(b_i)$ for $1 \leqslant i \leqslant m$ (see Remark 3.2). By Birkhoff's Theorem there exist $k \in \mathbb{N}$, permutation matrices $P_1, \ldots, P_k \in DS(n)$ and nonnegative numbers $\mu_1, \ldots, \mu_k \in \mathbb{R}$, $\sum_{j=1}^{k} \mu_j = 1$ such that $E = \sum_{j=1}^{k} \mu_j P_j$. Then, for $1 \leqslant i \leqslant m$ we have

$$b_i = V^* D_{\lambda(b_i)} V = V^* D_{E\lambda(a_i)} V = V^* \left(\sum_{j=1}^{k} \mu_j \, P_j^{\mathrm{t}} D_{\lambda(a_i)} P_j \right) V$$
$$= \sum_{j=1}^{k} \mu_j \, (U P_j V)^* \, a_i \, (U P_j V). \qquad \square$$

4.3. Joint majorizations and convex functions

In this section we consider characterizations of the joint majorizations in terms of the functional calculus described before Proposition 4.2. Given an arbitrary family of square matrices $(a_i)_{i=1,\ldots,m} \subseteq M_n(\mathbb{C})$, the *(first) joint numerical range* (see [8]) is defined by

$$W(a_1, \ldots, a_m) = \{(v^* a_1 v, \ldots, v^* a_m v) : v \in \mathbb{C}^n, v^* v = 1\}.$$

We shall relate the joint numerical range $W(a_1, \ldots, a_m)$ to the joint spectrum $\sigma(a_1, \ldots, a_m)$ of an Abelian family.

Lemma 4.6. *Let $(a_i)_{i=1,\ldots,m}$ be an Abelian family. Then,*

$$W(a_1, \ldots, a_m) = \mathrm{co}(\sigma(a_1, \ldots, a_m)).$$

Proof. Note that $W(a_1, \ldots, a_m)$ is invariant under unitary conjugation of the a_i's by a fixed unitary $U \in M_n$. So we can assume that $a_i = D_{\lambda(a_i)}$, $i = 1, \ldots, m$. If $v^* v = 1$ we have

$$(v^* a_1 v, \ldots, v^* a_m v) = \left(\sum_{j=1}^{n} |v_j|^2 \lambda_j(a_1), \ldots, \sum_{j=1}^{n} |v_j|^2 \lambda_j(a_m) \right)$$
$$= \sum_{j=1}^{n} |v_j|^2 \left(\lambda_j(a_1), \ldots, \lambda_j(a_m) \right),$$

where $\sum_{j=1}^{n} |v_j|^2 = 1$. The lemma follows from this fact. $\square$

Proposition 4.7. *Let $(a_i)_{i=1,\ldots,m}$ and $(b_i)_{i=1,\ldots,m}$ be two Abelian families. Then, the following are equivalent:*

1. $(a_i) \succ_w (b_i)$.
2. $W(b_1, \ldots, b_m) \subseteq W(a_1, \ldots, a_m)$.
3. *For every convex function* $f : V \to \mathbb{R}$ *it holds*

$$\|f(a_1, \ldots, a_m)\| \geqslant \|f(b_1, \ldots, b_m)\|,$$

where $V \subseteq \mathbb{R}^m$ *is a convex set containing* $\sigma(a_1, \ldots, a_m)$ *and* $\sigma(b_1, \ldots, b_m)$.

Proof. 1. $\Leftrightarrow$ 2. follows from Lemma 4.6 and item 1. of Proposition 4.2. On the other hand, 1. $\Leftrightarrow$ 3. follows from Corollary 3.11. $\square$

The following proposition is a consequence of item 2. in Theorem 4.5 and Proposition 4.7.

Proposition 4.8. *Let* $(a_i)_{i=1,\ldots,m}$ *and* $(b_i)_{i=1,\ldots,m}$ *be two Abelian families. Then,* $(a_i) \succ (b_i)$ *if and only if, for* $k = 1, \ldots, \left[\frac{n}{2}\right]$ *and* $k = n$ *we have*

$$\left\| f\left(\log\left[\bigwedge^k \exp(a_1)\right], \ldots, \log\left[\bigwedge^k \exp(a_m)\right]\right)\right\| \geqslant \left\| f\left(\log\left[\bigwedge^k \exp(b_1)\right], \ldots, \log\left[\bigwedge^k \exp(b_m)\right]\right)\right\|$$

for every convex function $f : V \to \mathbb{R}$, *where* $V \subseteq \mathbb{R}^m$ *is a convex set containing* $\sigma(a_1, \ldots, a_m)$ *and* $\sigma(b_1, \ldots, b_m)$.

The following proposition is a restatement of Theorem 3.9 in this context.

Proposition 4.9. *Let* $(a_i)_{i=1,\ldots,m}$ *and* $(b_i)_{i=1,\ldots,m}$ *be two Abelian families. Then,* $(a_i) \succ_s (b_i)$ *if and only if, for every convex function* $f : V \to \mathbb{R}$ *it holds that*

$$\operatorname{tr} f(a_1, \ldots, a_m) \geqslant \operatorname{tr} f(b_1, \ldots, b_m),$$

where $V \subseteq \mathbb{R}^m$ *is a convex set containing* $\sigma(a_1, \ldots, a_m)$ *and* $\sigma(b_1, \ldots, b_m)$.

4.4. Equivalence relations associated to joint majorizations

The joint majorizations considered so far are preorder relation among Abelian families in $M_n(\mathbb{C})$. The next theorem describes the equivalence relations associated to these preorders.

Theorem 4.10. *Let* $(a_i)_{i=1,\ldots,m}$ *and* $(b_i)_{i=1,\ldots,m}$ *be two Abelian families in* $M_n(\mathbb{C})$.

(a) *The following are equivalent*:
  (i) $(a_i) \succ_w (b_i)$ *and* $(b_i) \succ_w (a_i)$.
  (ii) $W(a_1, \ldots, a_m) = W(b_1, \ldots, b_m)$.

(b) *The following are equivalent*:

(i) *There exists a unitary matrix* $W \in M_n$ *such that* $W^* a_i W = b_i$ *for every* $i = 1, \ldots, m$.

(ii) $(a_i) \succ_s (b_i)$ *and* $(b_i) \succ_s (a_i)$.

(iii) $(a_i) \succ (b_i)$ *and* $(b_i) \succ (a_i)$.

Proof. (a) It is an immediate consequence of Proposition 4.7.

(b) Note that the inner automorphism $\alpha : C^*(a_1, \ldots, a_m) \to C^*(b_1, \ldots, b_m)$ induced by W is a trace preserving, positive unital map. Therefore (i) implies (ii). Clearly (ii) $\Rightarrow$ (iii). On the other hand, if $(a_i) \succ (b_i)$ and $(b_i) \succ (a_i)$, by Theorem 3.24 there exits a permutation matrix $Q \in M_n$ such that

$$V(Q^{\mathrm{t}}(U^* a_i U)Q)V^* = b_i, \quad i = 1, \ldots, m,$$

where $U, V \in M_n$ are as in the proof of Theorem 4.5. Therefore, by taking $W = UQV^*$ we have completed the proof. $\square$

Acknowledgments

We would like to thank Professor Demetrio Stojanoff for his advise and suggestions regarding the results herein. We would also like to thank the referee for many suggestions that helped us improve this paper.

References

[1] T. Ando, Majorization, doubly stochastic matrices and comparison of eigenvalues, Lecture Note, Hokkaido University, 1982.

[2] L. Beasley, S. Lee, Linear operators preserving multivariate majorization, Linear Algebra Appl. 304 (2000) 141–159.

[3] R. Bhatia, Matrix Analysis, Springer-Verlag, New York, 1997.

[4] G.-S. Cheon, Y.-H. Lee, The doubly stochastic matrices of a multivariate majorization, J. Korean Math. Soc. 32 (4) (1995) 857–867.

[5] G. Dahl, Matrix majorization, Linear Algebra Appl. 288 (1999) 53–73.

[6] R. Horn, C. Johnson, Matrix Analysis, Cambridge University Press, 1985.

[7] S.-G. Hwang, S.-S. Pyo, Matrix majorization via vector majorization, Linear Algebra Appl. 332–334 (2001) 15–21.

[8] C.-K. Li, Y.-T. Poon, Convexity of the joint numerical range, SIAM J. Matrix Anal. Appl. 21 (1999) 668–678.

[9] C.-K. Li, E. Poon, Linear operators preserving directional majorization, Linear Algebra Appl. 325 (1–3) (2001) 141–146.

[10] A.W. Marshall, I. Olkin, Inequalities: Theory of Majorization and its Applications, Academic Press, New York, 1979.

[11] S. Sherman, A correction to "On a conjecture concerning doubly stochastic matrices", Proc. Amer. Math. Soc. 5 (1954) 998–999.
[12] S. Shreiber, On a result of Sherman concerning doubly stochastic matrices, Proc. Amer. Math. Soc. 9 (1958) 350–353.