Stochastic Inequalities
IMS Lecture Notes – Monograph Series
Volume 22 (1993)

MULTIVARIATE MAJORIZATION BY POSITIVE COMBINATIONS[1]

By HARRY JOE and JOSEPH VERDUCCI

University of British Columbia and Ohio State University

Multivariate majorization orderings are used to compare matrices according to their dispersiveness. When applied to matrices whose rows represent distributions of different resources, the ordering that appears to be most useful is called majorization by positive linear combinations (PC-majorization). Two matrices are PC-majorized if all positive linear combinations of the rows are ordered by ordinary vector majorization. Properties of PC-majorization are derived; an algorithm is given to determine whether or not one matrix is PC-majorized by another; and elementary operations that reduce a matrix in the PC-ordering are explained.

1. Introduction

The key idea of majorization is to pre-order vectors according to a universal standard of dispersiveness. That is, any reasonable measure of dispersiveness of the components of a vector should imply an ordering that is consistent with the pre-ordering of majorization. The universality of the majorization ordering is well illustrated by the hundreds of applications mentioned in Marshall and Olkin (1979), and many other sources.

Several attempts have been made to extend majorization to a pre-ordering of matrices. However, there appears to be no 'universal' extension, but rather several different extensions that are useful for different purposes. For example, Joe (1985) uses a 'vectorized' generalization to describe association in contingency tables, and Tong (1989) uses uniform majorization (described below) to obtain probability inequalities for rectangles. Several other multivariate majorization orderings may be found in the books by Marshall and Olkin (1979) and Arnold (1987).

In this paper, we study multivariate majorization orderings that can be interpreted as orderings of distribution of wealth of several resources, with lower in the ordering meaning closer to equal division of the resources. Our

[1]Research supported by an NSERC Canada grant and a U. S. National Science Foundation grant.

AMS 1991 *subject classifications.* Primary: 15A45; Secondary: 15A39, 15A48.

Key words and phrases. Convexity, dispersion, linear programming, matrix inequalities.

orderings are on $m \times n$ matrices (m resources and n individuals) of real numbers. One potential area of application is to the monitoring and management of economic and ecological systems, where interest focuses on interventions that lead to more equitable consumption of resources, especially when some resources may become scarce. In such a situation the total value of the resources consumed (called *wealth*) by an individual or species may fluctuate, depending on current availability of each resource. The essential criterion for judging if a redistribution of resources leads to something 'universally' more equitable is that the vector describing the distribution of wealth changes to something smaller in the majorization pre-ordering, *no matter what values are assigned to each resource.* This idea is described in more detail in Arnold (1987, pp. 60-61).

The formal definition and basic properties of our proposed PC-majorization ordering are given in Section 2. Section 3 contains an algorithm to determine whether two given $m \times n$ matrices are ordered or not. The theory behind elementary methods for reducing matrices in this ordering is developed in Section 4. Section 5 contains a preliminary study of the set of matrices that are PC-majorized by a given matrix, and suggests how this leads to a more general method of reduction.

2. Definitions and Basic Properties

We first define vector majorization and give some of its equivalent forms (see Marshall and Olkin (1979) and Arnold (1987) for details) that we will use. Then we define multivariate majorization by positive comparisons (PC-majorization), and relate it to other forms of multivariate majorization. Through examples and results, we motivate PC-majorization as providing the most useful interpretion for distributions of several resources. Our notation follows Arnold (1987).

DEFINITION 2.1 Let $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ be n–dimensional row vectors. Let the ordered x_i and y_i be denoted by $x_{[1]} \geq \cdots \geq x_{[n]}$ and $y_{[1]} \geq \cdots \geq y_{[n]}$. Then $\mathbf{x}$ is majorized by $\mathbf{y}$ (written $\mathbf{x} \prec \mathbf{y}$) if

$$\sum_{i=1}^{k} x_{[i]} \leq \sum_{i=1}^{k} y_{[i]}, \quad k = 1, \ldots, n-1$$

and $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$.

Equivalent definitions of $\mathbf{x} \prec \mathbf{y}$ are: (a) $\mathbf{x} = \mathbf{y}D$, where D is a doubly stochastic $n \times n$ matrix (each row and column having nonnegative entries that sum to 1), and (b) $\mathbf{x}$ is in the convex hull of the vectors which are permutations of $\mathbf{y}$.

We now go on to matrices. An $m \times n$ matrix of reals is interpreted as a distribution of m resources among n people; the i^{th} row is the distribution of the i^{th} resource among n individuals and the j^{th} column is the vector of amounts of the m resources for the j^{th} individual. Negative values correspond to debts. Arnold (1987) gives an interpretation of in terms of n individuals with money in m different currencies.

Throughout this paper, X and Y are real $m \times n$ matrices, and $\mathbb{R}_+^m$ is the set of m–dimensional row vectors with nonnegative elements. We emphasize the following pre-ordering of these matrices.

DEFINITION 2.2 $X \prec^{\text{PC}} Y$ or X is PC-majorized by Y if $\mathbf{a}X \prec \mathbf{a}Y$ for all $\mathbf{a} \in \mathbb{R}_+^m$.

We shall compare PC-majorization with three other kinds of multivariate majorization, defined as follows.

DEFINITIONS 2.3 $X \prec^{\text{UM}} Y$ or X is *uniformly majorized* by Y if $X = YD$ for a $n \times n$ doubly stochastic matrix D.

$X \prec^{\text{LC}} Y$ or X is majorized by Y through *linear combinations* if $\mathbf{a}X \prec \mathbf{a}Y$ for all $\mathbf{a} \in \mathbb{R}^m$.

$X \prec^{\text{MM}} Y$ or X is *marginally majorized* by Y if $\mathbf{x}_i \prec \mathbf{y}_i$, $i = 1, \ldots, m$, where $\mathbf{x}_i, \mathbf{y}_i$ are the i^{th} rows of X and Y respectively.

Arnold refers to our Definition 2.2 as $\prec^{\text{MO}}$ for Marshall–Olkin, but Marshall and Olkin (1979) have $\mathbf{a} \in \mathbb{R}^m$, which is a stronger condition. To avoid confusion, we refer to the Marshall–Olkin ordering by $\prec^{\text{LC}}$.

The following simple example illustrates PC-majorization and indicates why it is more appropriate for ordering distributions of resources than are the other three forms of multivariate majorization.

EXAMPLE 2.4 Let $X = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}$ and let $Y = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$.

First we verify that $X \prec^{\text{PC}} Y$. To do this we must show that for any $\mathbf{a} = (a_1, a_2) \in \mathbb{R}_+^m$,

$$(a_1 + 4a_2, 3a_1 + 2a_2) \prec (a_1 + 2a_2, 3a_1 + 4a_2);$$

but this follows immediately, since $3a_1 + 4a_2 \geq \max\{a_1 + 4a_2, 3a_1 + 2a_2\}$.

On the other hand, X is not uniformly majorized by Y, because the only doubly stochastic matrix that would leave the first row of Y unaltered is the identity matrix. Neither is X LC-majorized by Y, as can be seen by taking $\mathbf{a} = (1, -1)$. Most people would consider X to represent a more equitable distribution of resources than Y, since the second individual in Y is clearly the richer. Uniform and LC-majorization do not make such desirable distinctions.

It should also be clear that X and Y are equivalent with respect to marginal majorization; that is, $X \prec^{\mathrm{MM}} Y$ and $Y \prec^{\mathrm{MM}} X$. The fact that Y can be considered as 'smaller' than X, in the sense of marginal majorization, makes marginal majorization unsuitable as a pre-ordering for joint distributions of resources. □

The point of this example is that uniform and LC-majorization are too restrictive in determining which matrices should be ordered, and that marginal majorization is not restrictive enough. This is made clear in the next theorem, which formally relates these four types of multivariate majorization.

THEOREM 2.5 *PC-majorization possesses the following properties:*

Invariance by permutations. If $X \prec^{\mathrm{PC}} Y$, *then* $PXQ_1 \prec^{\mathrm{PC}} PYQ_2$, *for any choice of permutation matrices* P, Q_1, *and* Q_2.

Invariance by addition. If $X \prec^{\mathrm{PC}} Y$, *then* $X + (c_1, \ldots, c_m)^T \mathbf{e} \prec^{\mathrm{PC}} Y + (c_1, \ldots, c_m)^T \mathbf{e}$ *for real constants* c_i, *where* T *stands for transpose and* $\mathbf{e}$ *is a* $1 \times n$ *vector of ones.*

Marginal Majorization. If $X \prec^{\mathrm{PC}} Y$, *then* $X \prec^{\mathrm{MM}} Y$.

Uniform Reduction. If $X \prec^{\mathrm{UM}} Y$, *then* $X \prec^{\mathrm{LC}} Y$, *which in turn implies* $X \prec^{\mathrm{PC}} Y$.

PROOF Invariance with respect to P follows from the invariance of the domain $\mathbb{R}_+^m$ of the weighting vector $\mathbf{a}$. Invariance with respect to Q_1 and Q_2 follows from the permutation invariance of ordinary vector majorization. Also, vector majorization is invariant under addition of the same constant to all components, which makes PC-majorization invariant under the addition of constant rows.

Suppose $X \prec^{\mathrm{PC}} Y$. Let $\mathbf{a} \in \mathbb{R}_+^m$ be a vector with a 1 in the i^{th} position and 0 elsewhere. Then $\mathbf{a}X \prec \mathbf{a}Y$ is equivalent to $\mathbf{x}_i \prec \mathbf{y}_i$. By letting i go from 1 to m, $X \prec^{\mathrm{MM}} Y$.

Finally, suppose $X \prec^{\mathrm{UM}} Y$. Then there exists a doubly stochastic matrix D such that $X = YD$, and $\mathbf{a}X = (\mathbf{a}Y)D$ or $\mathbf{a}X \prec \mathbf{a}Y$ for all $\mathbf{a} \in \mathbb{R}^m$. Hence $X \prec^{\mathrm{LC}} Y$. Restricting $\mathbf{a}$ to $\mathbb{R}_+^m$ gives $X \prec^{\mathrm{PC}} Y$. □

We note also that the $\prec^{\mathrm{MM}}$ and $\prec^{\mathrm{UM}}$ pre-orderings possess the same invariance properties as $\prec^{\mathrm{PC}}$. Further relationships among these these multivariate majorization pre-orderings are developed in Section 4, in the context of finding elementary operations that reduce a matrix Y to something smaller in the PC-majorization ordering. In particular, we will give conditions under which the orderings $\prec^{\mathrm{MM}}$ and $\prec^{\mathrm{PC}}$ become equivalent, in which case a reduction can be obtained by reducing one row of Y in the usual vector majorization ordering. First, however, it is useful to know how each of the three multivariate majorization relations is verified.

3. Determining Whether Two given Matrices are Ordered

In this section we show for given matrices X, Y how to check whether one or more of the multivariate majorization orderings in Section 2 holds. Marginal majorization is easily checked since it is rowwise majorization. If $X \prec^{\mathrm{MM}} Y$, it is logical next to check if $X \prec^{\mathrm{PC}} Y$. PC-majorization would seem to require checking vector majorization for an infinite number of $\mathbf{a}$ but we show that a finite number of $\mathbf{a}$'s will do. Finally, if $X \prec^{\mathrm{PC}} Y$, the check for uniform majorization also requires some work because there could be zero, one, or more doubly stochastic matrices D such that $X = YD$.

The main difficulty in checking for PC-majorization is that the ordering of the components of $\mathbf{a}X$ and $\mathbf{a}Y$ change as $\mathbf{a}$ varies over $\mathbb{R}^m_+$. The following example illustrates the difficulty, offers some geometric intuition, and suggests the general solution.

EXAMPLE 3.1 Suppose that

$$X = \begin{bmatrix} 4 & 3 & 3 \\ 3 & 3 & 4 \end{bmatrix}, \quad Y = \begin{bmatrix} 5 & 4 & 1 \\ 2 & 3 & 5 \end{bmatrix}.$$

For any $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^m_+$, we must check if

$$(4a_1 + 3a_2, 3a_1 + 3a_2, 3a_1 + 4a_2) \prec (5a_1 + 2a_2, 4a_1 + 3a_2, a_1 + 5a_2).$$

Let $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ denote the columns of Y, so that $\mathbf{A} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$, and $\mathbf{C} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$ correspond to individuals A, B, and C. As the first resource, whose worth is measured by a_1, decreases in value relative to the second resource, the ranking of individuals in decreasing order of total wealth goes from (ABC) to (BAC) to (BCA) to (CBA). Assuming that we have already checked for marginal majorization, we claim that we need to check $\mathbf{a}X \prec \mathbf{a}Y$ only for 3 values of $\mathbf{a}$ corresponding to the three transitions in the rankings of the relative wealth of individuals A, B, and C.

Some geometric intuition can be obtained by visualizing the column vectors $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ as points in the euclidean plane, and $\mathbf{a}$ as determining a ray from the origin through the point $\mathbf{a}$. For any given $\mathbf{a}$, the wealth of individuals A, B, and C is proportional to the orthogonal projection of the points $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ onto the ray determined by $\mathbf{a}$. A transition in rankings occurs when the ray is orthogonal to a line connecting some pair of points among $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$.

The transition from (ABC) to (BAC) occurs at $\mathbf{a} = (1,1)$; from (BAC) to (BCA) at $\mathbf{a} = (3,4)$; and from (BCA) to (CBA) at $\mathbf{a} = (2,3)$. In particular, for $\mathbf{a} = (3,4), \mathbf{a}X = (24,21,25)$ and $\mathbf{a}Y = (23,24,23)$. Therefore Y does not PC-majorize X.

The argument why attention can be restricted to the above three values of $\mathbf{a}$ is best given in the general setting that we now develop.

THEOREM 3.2 *Suppose that $X \prec^{\mathrm{MM}} Y$, where the columns of Y are denoted by $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$. For each (i,j) with $1 \leq i \leq j \leq n$, let*

$$T_{ij} = \{\mathbf{a} \in \mathbb{R}_+^m : \sum_k a_k = 1 \quad \text{and} \quad \mathbf{a}(\mathbf{Y}_i - \mathbf{Y}_j) = 0\}$$

and

$$T = \cup_{i<j} T_{ij}.$$

Then $X \prec^{\mathrm{PC}} Y$ if and only if $\mathbf{a}X \prec \mathbf{a}Y$ for every $\mathbf{a} \in T$.

PROOF First note that $\mathbf{a}X \prec \mathbf{a}Y$ if and only if $c\mathbf{a}X \prec c\mathbf{a}Y$ for every $c > 0$. Thus it is enough to check the definition of PC-majorization for all $\mathbf{a}$ such that $\sum_j a_j = 1$.

For each $k = 1, \ldots, m$, define

$$f_k(\mathbf{a}) = \sum_{i=1}^{k} (\mathbf{a}Y)_{[i]}$$

to be the sum of the k largest components of $\mathbf{a}Y$. We may also express $f_k(\mathbf{a})$ as

$$f_k(\mathbf{a}) = \max_Q \sum_{i=1}^{k} (\mathbf{a}YQ)_i$$

where the maximum is taken over all $n \times n$ permutation matrices. Expressed as a maximum of linear functions, it is clear that $f_k(\mathbf{a})$ is convex and piecewise linear. Transitions from one linear section to another can occur only when at least two of the components $\mathbf{a}\mathbf{Y}_i$ and $\mathbf{a}\mathbf{Y}_j$ are equal. A similar argument shows that

$$g_k(\mathbf{a}) = \sum_{i=1}^{k} (\mathbf{a}X)_{[i]},$$

which is the sum of the k largest components of $\mathbf{a}X$, is also convex.

To establish the PC-majorization $X \prec^{\mathrm{PC}} Y$, it is enough to show that $f_k(\mathbf{a}) \geq g_k(\mathbf{a})$ for each $\mathbf{a} \in \mathcal{H} \equiv \{\mathbf{a} \in \mathbb{R}_+^m : \sum_j a_j = 1\}$. The assumption that $X \prec^{\mathrm{MM}} Y$ implies that $f_k(\mathbf{a}) \geq g_k(\mathbf{a})$ at the extreme points $(1, \ldots, 0), \ldots, (0, \ldots, 1)$ of $\mathcal{H}$. Now suppose that $g_k(\mathbf{a}) > f_k(\mathbf{a})$ at some point $\mathbf{a} \in \mathcal{H}$. We may assume that the components of $\mathbf{a}Y$ are distinct, since otherwise $\mathbf{a} \in T$. Let $\mathcal{H}_{\mathbf{a}}$ denote the subset of $\mathcal{H}$ that contains $\mathbf{a}$ and over which $f_k(\cdot)$ is linear. This is a convex set, defined by

$$\mathcal{H}_{\mathbf{a}} = \{\mathbf{b} \in \mathcal{H} : \mathbf{b}_{\pi(i)} \geq \mathbf{b}_{\pi(i+1)}, i = 1, \ldots, m-1\},$$

where π is the permutation that puts the components of $\mathbf{a}Y$ in decreasing order; that is, $(\mathbf{a}Y)_{\pi(i)} = (\mathbf{a}Y)_{[i]}, i = 1, \dots, m$. Since $g_k(\cdot)$ is a convex function, if it exceeds $f_k(\cdot)$ anywhere in the region $\mathcal{H}_{\mathbf{a}}$, it must also exceed $f_k(\cdot)$ at one of the extreme points of $\mathcal{H}_{\mathbf{a}}$. Each of these extreme points is contained either in T or in the set $\{(1, \dots, 0), \dots, (0, \dots, 1)\}$ of extreme points of $\mathcal{H}$. It thus follows that if $g_k(\cdot)$ exceeds $f_k(\cdot)$ anywhere in $\mathcal{H}$, it must exceed $f_k(\cdot)$ on T. Hence, majorization needs to be checked only for $\mathbf{a} \in T$. □

In the case when $m = 2, T$ is finite, and Theorem 3.2 is all that is needed to check for PC-majorization. For $m > 2$, the sets T_{ij} are either empty or are convex polytopes. In this case, it suffices to check the extreme points of each nonempty T_{ij}. Hence overall only a finite number of $\mathbf{a}$'s need to be checked. A precise algorithm that checks for both marginal and PC-majorization in a finite number of steps is now described.

ALGORITHM 3.3. CHECK FOR PC-MAJORIZATION. Consider the hyperplanes

$$\{\mathbf{a} : \sum_k a_k(y_{ki} - y_{kj}) = 0, \} \qquad i < j.$$

The intersection of these hyperplanes with $\mathcal{H} = \{\mathbf{a} \in \mathbb{R}^m_+ : \sum_j a_j = 1\}$ are the convex polytopes T_{ij}, but notice that these hyperplanes need not intersect $\mathcal{H}$ at all, as happens, for example, when the two column vectors $\mathbf{Y}_i$ and $\mathbf{Y}_j$ are similarly ordered. In general, the above hyperplanes divide A into at most $n!$ convex polytopes (with faces T_{ij}) such that for each $\mathbf{a}$ in a particular polytope, $\mathbf{a}Y$ always has a certain ordering. The vertices of the convex polytopes may be found by taking $m-1$ of the hyperplane equations at a time and solving for a root $\mathbf{a}$ in $\mathcal{H}$ (if any). Since all these vertices are similarly ordered, the following lemma (with $\sum_j \lambda_j = 1$) proves that it is enough to check $\mathbf{a}X \prec \mathbf{a}Y$ for $\mathbf{a}$'s being one of these vertices. □

The next lemma is equivalent to result 5.A.6 on page 121 of Marshall and Olkin (1979).

LEMMA 3.4 *Suppose that the vectors* $\mathbf{z}^{(j)} = (z_1^{(j)}, \dots, z_n^{(j)})$, $j = 1, \dots, L$, *are similarly ordered and that there are corresponding vectors* $\mathbf{x}^{(j)}$ *such that* $\mathbf{x}^{(j)} \prec \mathbf{z}^{(j)}$. *Then* $\sum_j \lambda_j \mathbf{x}^{(j)} \prec \sum_j \lambda_j \mathbf{z}^{(j)}$ *if* $\lambda_j \geq 0$, $j = 1, \dots, L$.

In the next section, we discuss some elementary ways to transform a matrix Y into something smaller with respect to PC-majorization. The simplest way is to post multiply Y by a doubly stochastic matrix D, which leads to something smaller in uniform majorization. For a certain class of matrices Y, which we have not yet been able to specify completely, this is the only way to obtain something that is PC-majorized by Y.

Also, in our study of elementary operations that PC-reduce Y, we are interested in how a targeted X, which is known to be PC-majorized, might be obtained through a series of elementary operations. That is, given $X \prec^{\text{PC}} Y$, we would like to construct a sequence $Y \to Y_1 \to \ldots \to Y_k \to X$ such that each step $Y_{j-1} \to Y_j$ is an elementary operation.

A valuable tool for studying each of these problems is the following check for uniform majorization.

ALGORITHM 3.5. CHECK FOR UNIFORM MAJORIZATION ($\prec^{\text{UM}}$). First note that if $X = YD$ for a doubly stochastic matrix D, and if certain rows of Y are linearly dependent, say $\mathbf{b}Y = 0$ for a $\mathbf{b} \in \mathbb{R}^m$, then $\mathbf{b}X = 0$. However $X = YD$ and $\mathbf{b}X = 0$ need not imply $\mathbf{b}Y = 0$ unless D is invertible.

Given X, Y a procedure to use is the following. First check whether $X \prec^{\text{MM}} Y$. If not, then X, Y are not ordered by $\prec^{\text{UM}}$. Assuming $X \prec^{\text{MM}} Y$, there are a few cases to consider.

(a) If Y has at least n linearly independent rows, let Y_0 be a $n \times n$ submatrix of Y with n linearly independent rows, say rows $i_1 < \cdots < i_n$ of Y. Let X_0 be a $n \times n$ submatrix of X consisting of rows $i_1, \ldots, i_n$ of X. Y_0 is invertible. Let $D = Y_0^{-1} X_0$ and $Z = YD$. If D is doubly stochastic and $Z = X$, then $X \prec^{\text{UM}} Y$. Otherwise X and Y are not ordered by $\prec^{\text{UM}}$.

(b) If the number of linearly independent rows of Y is less than n, check whether each linear dependency in rows of Y implies the same linear dependency in rows of X. If not, then X and Y are not ordered by $\prec^{\text{UM}}$.

(c) Suppose the number of linearly independent rows of Y is exactly $n-1$. Let Y_0 be a (sub)matrix of Y consisting of $n-1$ independent rows, and let X_0 be the corresponding (sub)matrix of X. Then $X \prec^{\text{MM}} Y$ implies that the row sum vectors corresponding to X and Y are the same. Therefore $X_0 = Y_0 D$ reduces to solving for $(n-1)^2$ linear equations in $(n-1)^2$ unknowns d_{jk}, $j = 1, \ldots, n-1$, $k = 1, \ldots, n-1$. Note that the doubly stochastic requirement means that d_{nk} and d_{jn} can be substituted for. If $D = (d_{jk})_{1 \le j,k \le n}$ is nonnegative and $X = YD$, then $X \prec^{\text{UM}} Y$.

(d) If the number of linearly independent rows of Y is less than $n-1$, then potentially more than one D exists. Using $X = YD$ and D doubly stochastic as linear constraints on the d_{ij}, the simplex method of linear programming, for example, can be used to see if a feasible solution exists (any linear function of the d_{ij} can be used as an objective function). If a feasible solution exists, then $X \prec^{\text{UM}} Y$.

4. PC-Reduction of Matrices

The preceding results enable us to recognize when two matrices are related by any of the $\prec^{\text{MM}}, \prec^{\text{PC}}$ or $\prec^{\text{UM}}$ pre-orderings for dispersiveness.

We now consider interventions that reduce dispersiveness according to PC-majorization. These interventions take the mathematical form of special types of linear operations on an initial matrix Y. We describe three different types of operations.

DEFINITION 4.1 *Type (A) (Uniform reduction) operations.*
A type (A) operation is defined as a linear transformation $Y \to YD$ where D is a doubly stochastic matrix.

That type A operations always produce a PC-smaller matrix is an immediate consequence of Theorem 2.5; namely, if D is doubly stochastic, then $YD \prec^{\mathrm{PC}} Y$. We further conjecture that any doubly stochastic matrix can be written as a product of matrices of the form $\lambda Q + (1-\lambda)I$ where Q is an $n \times n$ permutation matrix, I is the $n \times n$ identity matrix, and $0 \le \lambda \le 1$. This conjectured decomposition would allow an arbitrary uniform reduction to be achieved as a sequence of 'rearrangement transfers' $\lambda Q + (1-\lambda)I$. An exercise in Arnold (1987, p. 75) demonstrates that the conjecture is false if the permutation matrices Q are restricted to transpositions of two coordinates, in which case $\lambda Q + (1-\lambda)I$ is called a 'Robin Hood transfer.'

Recall Example 2.4, where a PC-reduction was obtained by permuting elements in the second row of Y. It was shown there that such a reduction is not obtainable via type A operations. We now devise an elementary operation, based on majorization reduction of individual rows of certain submatrices of Y. Toward this end, consider the following definitions.

DEFINITION 4.2 Let J be a subset of $\{1, \dots, n\}$ with cardinality n_0 where $n_0 \ge 2$. X is said to have *similarly ordered rows in* J, if there is a permutation $j_1, \dots, j_{n_0}$ of the indices in J such that $x_{ij_1} \ge \cdots \ge x_{ij_{n_0}}$, for each $i = 1, \dots, m$. If $J = \{1, \dots, n\}$ then we will say that X has similarly ordered rows. In the case that X has two rows, the notion of *oppositely ordered* rows can be well defined, in the obvious way.

DEFINITION 4.3 *Type (B) (Marginal reduction) operations.*
A type (B) operation $Y \to X$ is described as follows: Let Y^* be an $m \times n_0$ submatrix of Y $(2 \le n_0 \le n)$ with similarly ordered rows; let the i^{th} row of Y^* be denoted by $\mathbf{y}_i^*$. For each i, replace $\mathbf{y}_i^*$ by a vector $\mathbf{x}_i^*$ which is smaller with respect to $\prec$. Assuming that the j^{th} column of Y^* corresponds to the k^{th} column of Y, replace y_{ik} by x_{ij}^*. The result is defined to be X.

We need to show that any X resulting from a type B operation satisfies $X \prec^{\mathrm{PC}} Y$. We begin with a simple lemma whose proof follows directly from elementary properties of vector majorization.

LEMMA 4.4 *Suppose X and Y differ only in the columns $j_1, \dots, j_{n_0}$ where $2 \le n_0 \le n$. Then $X \prec^{\mathrm{PC}} Y$ if and only if $X^* \prec^{\mathrm{PC}} Y^*$, where X^*, Y^* are the* *submatrices obtained from the* $j_1, \ldots, j_{n_0}$ *columns of* X *and* Y *respectively. A similar argument shows that* $\prec^{\mathrm{PC}}$ *can be replaced by* $\prec^{\mathrm{UM}}$, $\prec^{\mathrm{LC}}$ *or* $\prec^{\mathrm{MM}}$.

The key set of conditions under which PC- and marginal majorization become equivalent are described formally as follows:

THEOREM 4.5 *If* Y *has similarly ordered rows and* $X \prec^{\mathrm{MM}} Y$, *then* $X \prec^{\mathrm{PC}} Y$.

PROOF Because $\prec^{\mathrm{PC}}$ and $\prec^{\mathrm{MM}}$ share the same invariance property given in Theorem 2.5, we can assume without loss of generality that $Y = (y_{ij})$ satisfies $y_{i1} \geq \cdots \geq y_{in}$, $i = 1, \ldots, m$. Let $\mathbf{a}$ be an arbitrary vector in $\mathbb{R}_+^m$. Then $\mathbf{z} = \mathbf{a}Y$ satisfies $z_1 \geq \cdots \geq z_n$ and the sum of the k largest elements of $\mathbf{z}$ is $\sum_{i=1}^m a_i \sum_{j=1}^k y_{ij}$. Let j_ℓ be the index of the ℓ^{th} largest element of $\mathbf{a}X$ and let $x_{i[j]}$ be the j^{th} largest element of the i^{th} row of X. Then the sum of the k largest elements of $\mathbf{a}X$ is $\sum_{i=1}^m a_i \sum_{\ell=1}^k x_{ij_\ell}$, and

$$\sum_{i=1}^{m} a_i \sum_{\ell=1}^{k} x_{ij_\ell} \leq \sum_{i=1}^{m} a_i \sum_{j=1}^{k} x_{i[j]} \leq \sum_{i=1}^{m} a_i \sum_{j=1}^{k} y_{ij},$$

where the last inequality follows from the assumption $X \prec^{\mathrm{MM}} Y$. □

Together, Lemma 4.4 and Theorem 4.5 prove that a type B operation on Y leads to something smaller with respect to $\prec^{\mathrm{PC}}$. Reduction via marginal majorization is easy, and we have so far shown that reduction, via marginal majorization, of an $m \times n_0$ submatrix of Y with similarly ordered rows also leads to something smaller with respect to PC-majorization.

We wish to investigate the class of matrices that can be obtained by applying sequences of type (A) and type (B) to a given matrix Y. For this purpose, the following definition is convenient.

DEFINITION 4.6 X is majorized by Y via *simple transfers* ($X \prec^{\mathrm{ST}} Y$) if X can be obtained from Y via a finite number of operations of type (A) or (B) above.

By the above results, $X \prec^{\mathrm{ST}} Y$ implies $X \prec^{\mathrm{PC}} Y$. The next theorem shows that these two orderings are, in fact, equivalent when $n = 2$. Section 5 contains some other conditions under which equivalence is obtained.

THEOREM 4.7 *If* $n = 2$, *then* $X \prec^{\mathrm{ST}} Y$ *if and only if* $X \prec^{\mathrm{PC}} Y$.

PROOF Suppose $X \prec^{\mathrm{PC}} Y$. By Theorem 2.5, $X \prec^{\mathrm{MM}} Y$. If Y has similarly ordered, then by Definition 4.6, X can be obtained from Y via one operation of type (B). Next suppose that Y does not have similarly ordered rows. Since $X \prec^{\mathrm{MM}} Y$, for each $i = 1, \ldots, m$, there exist constants θ_i in the interval [0,1] such that $x_{i1} = \theta_i y_{i1} + (1-\theta_i)y_{i2}$, $x_{i2} = (1-\theta_i)y_{i1} + \theta_i y_{i2}$. Suppose that rows i and i' of Y are oppositely ordered. Let $\mathbf{a}$ be a vector with $a_i = 1$, $a_{i'} = (y_{i1} - y_{i2})/(y_{i'2} - y_{i'1}) > 0$ and $a_k = 0$, $k \neq i, i'$. Then $\mathbf{a}Y$ is a 2–vector of the constant $y_{i1} + a_{i'}y_{i'1}$; and $\mathbf{a}X \prec \mathbf{a}Y$ implies that $\mathbf{a}X = \mathbf{a}Y$, which can happen for this $\mathbf{a}$ only if $\theta_i = \theta_{i'}$. Since each non–constant row must be oppositely ordered to either row i or i' and θ can be anything (in [0,1]) for a constant row, it follows that $X = YD$, where $D = \begin{bmatrix} \theta_i & 1-\theta_i \\ 1-\theta_i & \theta_i \end{bmatrix}$. □

EXAMPLES 4.8 In general, there can be matrices that are $\prec^{\mathrm{PC}}$ ordered but not $\prec^{\mathrm{ST}}$ ordered. The methods of Section 3 can be used to verify that

$$X = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{bmatrix} \prec^{\mathrm{PC}} \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 2 & 3 & 5 \end{bmatrix} = Y.$$

For this example, it is easily checked that the matrices X and Y are not $\prec^{\mathrm{ST}}$ ordered.

The example on page 59 of Arnold (1987) with $Y = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 4 & 6 \end{bmatrix}$, $X = \begin{bmatrix} 1 & .5 & .5 \\ 3 & 4 & 5 \end{bmatrix}$, and $X = YD$, with $D = \begin{bmatrix} .5 & .5 & 0 \\ .5 & 0 & .5 \\ 0 & .5 & .5 \end{bmatrix}$ shows that Definition 4.6 cannot be simplified to simple transfers of type (A) and (B) that operate on only two columns at a time. This is one difference in going from vector majorization to matrix majorization.

The next, and last, example shows that with $X \prec^{\mathrm{ST}} Y$, it is possible for one column of X to be dominant even if there is no dominant column in Y.

$$X = \begin{bmatrix} 2.1 & 2.5 & 2.4 \\ 2.9 & 4 & 3.1 \end{bmatrix} \prec^{\mathrm{ST}} \begin{bmatrix} 2.1 & 3.0 & 1.9 \\ 2.9 & 4 & 3.1 \end{bmatrix} \prec^{\mathrm{ST}} \begin{bmatrix} 4 & 3 & 0 \\ 1 & 4 & 5 \end{bmatrix} = Y$$

This cannot happen if there is a nonnegative vector $\mathbf{a}$ such that $\mathbf{a}Y$ is a constant vector.

The following development suggests a simple operation other than (A) and (B) that also preserves the $\prec^{\mathrm{PC}}$ ordering, and which allows the transformation from Y to X in the first of Examples 4.8.

DEFINITIONS 4.9 Let $\mathbf{x}, \mathbf{y}$, and $\mathbf{z}$ be n–dimensional row vectors. Then $\mathbf{x}$ is $\mathbf{z}$–majorized by $\mathbf{y}$ (written $\mathbf{x} \prec_{\mathbf{z}} \mathbf{y}$) if $c\mathbf{z} + \mathbf{x} \prec c\mathbf{z} + \mathbf{y}$ for every $c > 0$.

In the case where $x \prec \mathbf{y}$, we define $S_{\mathbf{xy}} \equiv \{\mathbf{z} : \mathbf{x} \prec_{\mathbf{z}} \mathbf{y}\}$.

DEFINITION 4.10 *Type (C) (Directed reduction) operations.*

A type (C) operation is described as follows. Let Y have rows $\mathbf{y}_1, \ldots, \mathbf{y}_m$, and let $\mathbf{x} \prec \mathbf{y}_i$. Then $\mathbf{y}_i$ may be replaced by $\mathbf{x}$ if $S_{\mathbf{xy}_i}$ is convex and $\mathbf{y}_j \in S_{\mathbf{xy}_i}$ for each $j \neq i$.

To prove that a type C operation produces a PC-reduction, it is necessary first to develop properties of $\mathbf{z}$–majorization and the set $S_{\mathbf{xy}}$. Some immediate consequences of the definiton of $\mathbf{z}$–majorization are listed in the following lemma.

LEMMA 4.11
(a) If $\mathbf{x} \prec_{\mathbf{z}} \mathbf{y}$*, then* $\mathbf{x} \prec_{c\mathbf{z}} \mathbf{y}$ *for every* $c > 0$.
(b) $\mathbf{x} \prec_{\mathbf{z}} \mathbf{y}$ *if and only if* $\mathbf{z} + c\mathbf{x} \prec \mathbf{z} + c\mathbf{y}$ *for every* $c > 0$.
(c) If $\mathbf{x} \prec \mathbf{y}$*, then* $\mathbf{x} \prec_{\mathbf{e}} \mathbf{y}$*, where* $\mathbf{e}$ *is the* $1 \times n$ *vector of ones.*
(d) $\mathbf{x} \prec_{\mathbf{z}} \mathbf{y}$ *if and only if*

$$X = \begin{bmatrix} \mathbf{z} \\ \mathbf{x} \end{bmatrix} \prec^{\mathrm{PC}} \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix} = Y.$$

Part (a) of the lemma shows that the $\prec_{\mathbf{z}}$ ordering depends only on the direction, and not the actual value, of z. Part (b) is an alternate definition useful in proofs. Part (c) shows that the definition is not vacuous. Part (d) makes the key association with PC-majorization.

The role of $\mathbf{z}$ in $\mathbf{x} \prec_{\mathbf{z}} \mathbf{y}$ is to determine a set of permutation pairs $P_{\mathbf{xy}} \equiv \{(\pi, \nu)\}$ under which

$$\sum_{i=1}^{k} \mathbf{x}_{\pi(i)} \leq \sum_{i=1}^{k} \mathbf{y}_{\nu(i)}$$

for each $k = 1, \ldots, n$. Suppose that X and Y have the form given in Lemma 4.11d, with $X \prec^{\mathrm{PC}} Y$. Then

$$(1-\lambda)\mathbf{z} + \lambda\mathbf{x} \prec (1-\lambda)\mathbf{z} + \lambda\mathbf{y} \quad \text{for} \quad 0 \leq \lambda \leq 1.$$

If λ is sufficiently close to 0, Lemma 4.11d implies that $(\pi_{\mathbf{z}}, \pi_{\mathbf{z}})$ must be in $P_{\mathbf{xy}}$, where $\pi_{\mathbf{z}}$ is the permutation that puts the components of $\mathbf{z}$ in decreasing order. Similarly the pair $(\pi_{\mathbf{x}}, \pi_{\mathbf{y}})$ must be in $P_{\mathbf{xy}}$. As λ moves from 0 to 1, the ordering of $(1-\lambda)\mathbf{z} + \lambda\mathbf{x}$ changes from $\pi_{\mathbf{z}}$ to $\pi_{\mathbf{x}}$, and the ordering of $(1-\lambda)\mathbf{z} + \lambda\mathbf{y}$ changes from $\pi_{\mathbf{z}}$ to $\pi_{\mathbf{y}}$. In fact, the number of transpositions (Hamming distance) between either $(1-\lambda)\mathbf{z} + \lambda\mathbf{x}$ or $(1-\lambda)\mathbf{z} + \lambda\mathbf{y}$ and $\pi_{\mathbf{z}}$ is an increasing function of λ (cf. Theorem 5.2).

Lemma 4.11d describes, for $m = 2$, when one of the rows $\mathbf{y}$ of Y can be replaced by a new row $\mathbf{x}$ to achieve a PC-reduction; namely, when $\mathbf{x} \prec_{\mathbf{z}} \mathbf{y}$. We are thus interested in the question: Given $Y = \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix}$, what vectors $\mathbf{x}$ satisfy $\mathbf{x} \prec_{\mathbf{z}} \mathbf{y}$? It appears easiest to approach this problem indirectly, by first fixing $\mathbf{x}$ and $\mathbf{y}$ and then relating $P_{\mathbf{xy}}$ to the key construct of Definition 4.9, $S_{\mathbf{xy}}$.

LEMMA 4.12 *The set* $S_{\mathbf{xy}}$ *has the following properties:*

(a) If $\mathbf{z} \in S_{\mathbf{xy}}$*, then* $c\mathbf{z} \in S_{\mathbf{xy}}$*, for every* $c > 0$.
(b) If $\mathbf{z}$ *is similarly ordered to* $\mathbf{y}$*, then* $\mathbf{z} \in S_{\mathbf{xy}}$.
(c) If $\mathbf{z}$ *is oppositely ordered to* $\mathbf{x}$*, then* $\mathbf{z} \in S_{\mathbf{xy}}$.

PROOF Part (a) is an immediate consequence of Lemma 4.11a. Part (b) follows from

$$\sum_{i=1}^{k} (\mathbf{x} + c\mathbf{z})_{[i]} \leq \sum_{i=1}^{k} \mathbf{x}_{[i]} + c\mathbf{z}_{[i]} \leq \sum_{i=1}^{k} \mathbf{y}_{[i]} + c\mathbf{z}_{[i]} = \sum_{i=1}^{k} (\mathbf{y} + c\mathbf{z})_{[i]}.$$

Part (c) follows from the existence of some permutation π such that

$$\sum_{i=1}^{k} (\mathbf{y} + c\mathbf{z})_{[i]} = \sum_{i=1}^{k} \mathbf{y}_{[i]} + c\mathbf{z}_{\pi(i)} \geq \sum_{i=1}^{k} \mathbf{x}_{[i]} + c\mathbf{z}_{\pi(i)} \geq \sum_{i=1}^{k} (\mathbf{x} + c\mathbf{z})_{[i]}. \qquad \square$$

EXAMPLE 4.13 Let $\mathbf{y} = (1, 2, 3)$. We will show that, for each $\mathbf{x} = (x_1, x_2, x_3) \prec \mathbf{y}$, $S_{\mathbf{xy}}$ can be described as the intersection of half-spaces.

First note that for $\mathbf{x} = \mathbf{y}$, $S_{\mathbf{xy}}$ is all of $\mathbb{R}^n$. Otherwise,

a) If $x_1 \neq 1$ and $x_3 \neq 3$, then $S_{\mathbf{xy}}$ is constrained by $z_3 \geq z_1$.

b1) If $x_2 < 2$, then $S_{\mathbf{xy}}$ is constrained by $z_2 \geq z_1$.

b2) If $x_2 > 2$, then $S_{\mathbf{xy}}$ is constrained by $z_3 \geq z_2$.

Thus there are six possibilities for $S_{\mathbf{xy}}$. Either it is all of $\mathbb{R}^n$; is one of the three half spaces $\{z_2 \geq z_1\}, \{z_3 \geq z_1\}$ or $\{z_3 \geq z_2\}$; or is the intersection of two half spaces $\{z_3 \geq z_1\} \cap \{z_2 \geq z_1\}$ or $\{z_3 \geq z_1\} \cap \{z_3 \geq z_2\}$.

More generally, it appears that $S_{\mathbf{xy}}$ is the union of convex cones that are subsets of sets of the form $\{\mathbf{z} : z_{\pi(1)} \geq \cdots \geq z_{\pi(n)}\}$, where π ranges over a subset of permutations. As the following example illustrates, $S_{\mathbf{xy}}$ need not always be convex.

EXAMPLE 4.14 Let $\mathbf{x} = (3, 5, 3, 1)$ and $\mathbf{y} = (0, 2, 4, 6)$. Then $S_{\mathbf{xy}}$ is the union of 6 regions:

A. $z_4 \geq z_3 \geq z_2 \geq z_1$
B. $z_4 \geq z_2 \geq z_3 \geq z_1$
C. $z_4 \geq z_3 \geq z_1 \geq z_2$
D. $z_4 \geq z_1 \geq z_3 \geq z_2$
E. $z_3 \geq z_4 \geq z_1 \geq z_2$
F. $z_3 \geq z_4 \geq z_2 \geq z_1$.

The point $(0, 1, .1, 3)$ is in B, the point $(1, 0, .1, 3)$ is in D, but the average of these points, $(.5, .5, .1, 3)$ is not in the union $S_{\mathbf{xy}}$ of the 6 regions.

Although the convexity of $S_{\mathbf{xy}}$ is not always guaranteed, the following theorem, which presumes the convexity of $S_{\mathbf{xy}}$ for a particular pair of vectors $(\mathbf{x}, \mathbf{y})$, can be useful for reducing a matrix Y when redistribution is restricted to a single resource.

THEOREM 4.15 *Let Y have row vectors $\mathbf{y}_1, \ldots, \mathbf{y}_m$. Suppose that X differs from Y only in the i^{th} row, and that $S_{\mathbf{x}_i \mathbf{y}_i}$ is convex. Then $X \prec^{\mathrm{PC}} Y$ if and only if $\mathbf{x}_i \prec_{\mathbf{y}_j} \mathbf{y}_i$ for each $j \neq i$.*

PROOF First suppose that $X \prec^{\mathrm{PC}} Y$. For fixed $j \neq i$, and $k \in 1, 2, \ldots, m$, let

$$a_k = \begin{cases} c, & \text{if } k = i; \\ 1, & \text{if } k = j; \\ 0, & \text{otherwise.} \end{cases}$$

Then $\sum_{k=1}^m a_k \mathbf{x}_k \prec \sum_{k=1}^m a_k \mathbf{y}_k$ implies that $\mathbf{x}_j + c\mathbf{y}_i \prec \mathbf{x}_j + c\mathbf{x}_i$, and so $\mathbf{y}_i \prec_{\mathbf{x}_j} \mathbf{x}_i$.

We prove the converse by induction on m. By Lemma 4.11(d), the converse is true for $m = 2$. Suppose that it holds for some $m - 1$ where $m \geq 3$. Then there are at least two indices in $1, \ldots, m$ that are distinct from i. Without loss of generality, call these $m - 1$ and m. Then for any $\mathbf{a} \in \mathbb{R}_+^m$ define

$$\mathbf{z}_k = \begin{cases} \mathbf{y}_k, & \text{for } k = 1, \ldots, m-2; \\ a_{m-1}\mathbf{y}_{m-1} + a_m \mathbf{y}_m, & \text{for } k = m-1, \end{cases}$$

and

$$\mathbf{z}_k^* = \begin{cases} \mathbf{z}_k, & \text{for } k \neq i; \\ \mathbf{x}_i, & \text{for } k = i. \end{cases}$$

To invoke induction, we need to verify that

$$\mathbf{z}_i^* \prec_{\mathbf{z}_j} \mathbf{z}_i \quad \text{for each } j \in \{1, \ldots, m-1\} \backslash \{i\}. \tag{4.1}$$

For $j < m - 1$, (4.1) holds by assumption. For $j = m - 1$, (4.1) follows from the convexity and scale invariance (Lemma 4.12a) of $S_{\mathbf{z}_i^* \mathbf{z}_i}$. Now define

$$b_k = \begin{cases} a_k, & \text{for } k = 1, \ldots, m-2; \\ 1, & \text{for } k = m-1. \end{cases}$$

Then

$$\sum_{k=1}^{m} a_k \mathbf{x}_k = \sum_{k=1}^{m-1} b_k \mathbf{z}_k^* \prec \sum_{k=1}^{m-1} b_k \mathbf{z}_k = \sum_{k=1}^{m} a_k \mathbf{y}_k,$$

where the majorization follows from the induction assumption. Thus $X \prec^{\mathrm{PC}} Y$. □

Theorem 4.15 proves that type (C) operations preserve the $\prec^{\mathrm{PC}}$ ordering: It should be clear that a type (C) operation is distinct from types (A) and (B). Nevertheless, even with this additional operation, repeated application of operations (A), (B), and (C) do not generally produce all the matrices that are $\prec^{\mathrm{PC}}$ smaller than an arbitrary initial matrix, as the following counterexample shows.

EXAMPLE 4.16 Let $X = \begin{bmatrix} 1.5 & 1.5 & 3 \\ 2.5 & 3.5 & 1 \end{bmatrix}$ and $Y = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 1 & 2 \end{bmatrix}$. Using Theorem 3.2, it can be shown that $X \prec^{\mathrm{PC}} Y$. It is easy to argue that X and Y are not $\prec^{\mathrm{ST}}$ ordered. Also, it can be shown that X is not attainable from Y using type (C) operations. This example is a specific case of a more detailed study of 2×3 matrices in Theorem 5.6.

We do not know if there is yet another simple operation, which, in combination with (A), (B), and (C), will produce all matrices that are $\prec^{\mathrm{PC}}$ smaller than an arbitrary initial matrix.

5. Further Properties and PC-Reductions

In this section, we obtain additional properties for the $\prec^{\mathrm{PC}}$ and $\prec^{\mathrm{ST}}$ orderings. We show that the set $\mathcal{A}_Y = \{X : X \prec^{\mathrm{PC}} Y\}$ is convex and is contained in the convex set $\{X : X \prec^{\mathrm{MM}} Y\}$. The extreme points of $\{X : X \prec^{\mathrm{MM}} Y\}$ constitute the set, denoted by $\mathcal{C}_Y$, of matrices whose rows are permutations of the corresponding rows of Y. Methods for identifying the extreme points of $\mathcal{A}_Y$ are described and illustrated in the case $m = 2$, $n = 3$. One of the methods suggests a (non-simple) PC-reduction operation to be studied in further research.

In the course of this development, we show that, when restricted to $\mathcal{C}_Y$, the $\prec^{\mathrm{ST}}$ ordering is the same as the multivariate arrangement ordering of Boland and Proschan (1988).

THEOREM 5.1 *For fixed Y, the set $\mathcal{A}_Y = \{X : X \prec^{\mathrm{PC}} Y\}$ is a convex set.*

PROOF Suppose $X_1 \prec^{\mathrm{PC}} Y$ and $X_2 \prec^{\mathrm{PC}} Y$. Let λ be in the interval $[0,1]$. For $\mathbf{a} \in \mathbb{R}^m_+$, $\mathbf{a}(\lambda X_1 + (1-\lambda)X_2) = \lambda(\mathbf{a}X_1) + (1-\lambda)(\mathbf{a}X_2) \prec \mathbf{a}Y$ since $\mathbf{a}X_1 \prec \mathbf{a}Y$, $\mathbf{a}X_2 \prec \mathbf{a}Y$ and the set $\{\mathbf{x} : \mathbf{x} \prec \mathbf{a}Y\}$ is convex. Therefore $\lambda X_1 + (1-\lambda)X_2 \prec^{\mathrm{PC}} Y$. □

A natural question now concerns the extreme points of $\mathcal{A}_Y$ and relationships with convex hull results for vector majorization. First recall that $\mathcal{C}_Y$ is the set of the extreme points of $\{X : X \prec^{\mathrm{MM}} Y\}$, and that the extreme points of $\mathcal{A}_Y$ include some, but not necessary all, of the points in $\mathcal{C}_Y$.

Consider the $\prec^{\mathrm{ST}}$ and the $\prec^{\mathrm{PC}}$ orderings on $\mathcal{C}_Y$. The maximal matrices in $\mathcal{C}_Y$ with respect to $\prec^{\mathrm{ST}}$ or $\prec^{\mathrm{PC}}$ have similarly ordered rows. If Y has similarly ordered rows, then by Birkhoff's theorem, $\mathcal{A}_Y$ is the convex hull of the matrices in $\mathcal{C}_Y$. By Theorem 4.6, the $\prec^{\mathrm{ST}}$ and $\prec^{\mathrm{PC}}$ orderings on $\mathcal{C}_Y$ are the same if $n = 2$. For $m = 2$ and $n = 3$ or 4, the $\prec^{\mathrm{PC}}$ and $\prec^{\mathrm{ST}}$ orderings on $\mathcal{C}_Y$ can be shown to be the same by complete pairwise comparisons. Example 4.8 shows that the equivalence does not extend to $m = 2$, $n \geq 5$.

From results of Section 4, if two matrices in $\mathcal{C}_Y$ are $\prec^{\mathrm{ST}}$ ordered, then they are $\prec^{\mathrm{PC}}$ ordered.

We next note that the $\prec^{\mathrm{ST}}$ ordering on $\mathcal{C}_Y$ is the same as the multivariate arrangement increasing ordering in Boland and Proschan (1988). This is the arrangement increasing ordering or decreasing in transposition ordering of Hollander, Proschan and Sethuraman (1977) in the case $m = 2$. Boland and Proschan's definition involve only transfers that operate on two columns at a time. The equivalence follows from the proposition below.

PROPOSITION 5.2 *Let Y be a $m \times n$ matrix with similarly ordered rows and let X be a matrix such that its i^{th} row is a permutation of the i^{th} row of Y. Then X can be obtained from Y using a sequence of type (B) operations that operate on two columns.*

PROOF Without loss of generality, suppose that each row of Y is increasing. X can be obtained from Y by operating in columns $(n-1, n)$ followed by $(n-2, n), \dots, (1, n), (n-2, n-1), \dots, (1, n-1), \dots, (1, 2)$. In the first stage, each row involves switches of the j^{th} column with the n^{th} column, $j = n-1, \dots, 1$, until the n^{th} column of X is obtained. In the next stage switches are made until the $(n-1)^{st}$ column is correct, etc. Note that the appropriate columns are ordered correctly after each transfer of type (B) on two columns in order that the later type (B) operations can be made. □

We now go to a further study of $\mathcal{A}_Y$. Consider the set $\mathcal{D}_Y = \{Z \in \mathcal{C}_Y : Z \prec^{\mathrm{ST}} Y\}$ and let $\mathcal{B}_Y$ be the convex hull of the points in $\mathcal{D}_Y$. Then clearly $\mathcal{B}_Y \subset \mathcal{A}_Y$. By Theorem 4.6, $\mathcal{B}_Y = \mathcal{A}_Y$ if $n = 2$, and it can be shown that this is also valid sometimes when $n > 2$. One problem for future research is to deduce all conditions for which $\mathcal{B}_Y = \mathcal{A}_Y$.

From Theorem 3.2, the region $\mathcal{A}_Y$ can be specified precisely through a finite number of linear inequalities. The extreme points of $\mathcal{A}_Y$ can be enumerated by using some theory from linear programming and the simplex method. An alternative approach makes use of separating hyperplanes. Both approaches are not difficult to implement on a computer for a given Y but the combinatorial enumeration grows rapidly as m and n increase. However, both approaches led to the examples showing that the $\prec^{\mathrm{PC}}$ and $\prec^{\mathrm{ST}}$ orderings are not equivalent, and partly with the help of symbolic manipulation software, have been used to prove general results for $m = 2$, $n = 3$. We illustrate both approaches below. The separating hyperplane approach is shown first. Its advantage is that it suggests a more general reduction operation which is mentioned at the end of this section.

Since each matrix in $\mathcal{A}_Y$ has the same row sum vector as Y, $\mathcal{A}_Y$ is a set in $m(n-1)$ dimensional space. Hence $\mathcal{B}_Y$ can be partitioned into simplices, each with $m(n-1)+1$ vertices, where each vertex is in $\mathcal{D}_Y$. To see if $\mathcal{A}_Y$

extends beyond $\mathcal{B}_Y$, for each point Y^* in $\mathcal{C}_Y \backslash \mathcal{D}_Y$, simplices extending from $\mathcal{B}_Y$ to Y^* can be defined and then it can be checked whether any point in these simplices are in $\mathcal{A}_Y$. All such simplices can be found by enumerating sets of $m(n-1)$ matrices in $\mathcal{D}_Y$ and obtaining the hyperplane that these matrices lie on; the hyperplane can be found by solving a linear system or using a singular value decomposition depending on whether the matrix consisting of the column vectors formed from the first $n-1$ columns of each of the $m(n-1)$ matrices has full or less than full rank. The simplices involving Y^* that need to be checked are those resulting from hyperplanes that separate Y^* from $\mathcal{B}_Y$. The ideas will be clearer from the next result below for $m=2$, $n=3$ where Y has two oppositely ordered rows.

THEOREM 5.3 *Let $u_1 < u_2 < u_3$ and $v_1 < v_2 < v_3$. For* $Y = \begin{bmatrix} u_1 & u_2 & u_3 \\ v_3 & v_2 & v_1 \end{bmatrix}$, *$\mathcal{A}_Y = \mathcal{B}_Y$. Hence for this Y, if $X \prec^{\mathrm{PC}} Y$, then $X \prec^{\mathrm{ST}} Y$ and $X \prec^{\mathrm{UM}} Y$.*

PROOF Let $Y_1 = Y$, $Y_2 = \begin{bmatrix} u_1 & u_3 & u_2 \\ v_3 & v_1 & v_2 \end{bmatrix}$, $Y_3 = \begin{bmatrix} u_2 & u_1 & u_3 \\ v_2 & v_3 & v_1 \end{bmatrix}$, $Y_4 = \begin{bmatrix} u_2 & u_3 & u_1 \\ v_2 & v_1 & v_3 \end{bmatrix}$, $Y_5 = \begin{bmatrix} u_3 & u_1 & u_2 \\ v_1 & v_3 & v_2 \end{bmatrix}$, $Y_6 = \begin{bmatrix} u_3 & u_2 & u_1 \\ v_1 & v_2 & v_3 \end{bmatrix}$. These are the 6 points in $\mathcal{D}_Y$ and $\mathcal{B}_Y$ lies in a four-dimensional space determined by the row sums remaining constant. Let $X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{bmatrix}$ be a generic point in this four-dimensional space.

Let $b_1 = (u_2 - u_1)/(v_3 - v_2)$, $b_2 = (u_3 - u_1)/(v_3 - v_1)$ and $b_3 = (u_3 - u_2)/(v_2 - v_1)$. These are the values of b such that $(1, b)Y$ is a vector with at least two components equal. Either b_1, b_2, b_3 are distinct or they are all the same. In the latter case, the conclusion follows from Proposition 5.4 below. Hence we now assume the former case. Note that b_2 is a convex combination of b_1 and b_3 so that it cannot be largest or smallest of the 3 b_j's. The 15 subsets of size 4 from $\mathcal{D}_Y$ and the equations of the corresponding hyperplanes for the each set of 4 are:

(a) Y_1, Y_2, Y_3, Y_4: $x_{11} + b_1 x_{21}$
(b) Y_1, Y_2, Y_3, Y_5: $x_{13} + b_3 x_{23}$
(c) Y_1, Y_2, Y_3, Y_6: $(v_3 - v_1)x_{11} + (u_3 - u_1)x_{21} + (v_2 - v_1)x_{12} + (u_3 - u_2)x_{22}$
(d) Y_1, Y_2, Y_4, Y_5: $(v_3 - v_2)x_{11} + (u_2 - u_1)x_{21} - (v_2 - v_1)x_{12} - (u_3 - u_2)x_{22}$
(e) Y_1, Y_2, Y_4, Y_6: $x_{12} + b_3 x_{22}$
(f) Y_1, Y_2, Y_5, Y_6: $x_{11} + b_2 x_{21}$
(g) Y_1, Y_3, Y_4, Y_5: $(v_2 - v_1)x_{11} + (u_3 - u_2)x_{21} + (v_3 - v_1)x_{12} + (u_3 - u_1)x_{22}$
(h) Y_1, Y_3, Y_4, Y_6: $x_{13} + b_2 x_{23}$
(i) Y_1, Y_3, Y_5, Y_6: $x_{12} + b_1 x_{22}$
(j) Y_1, Y_4, Y_5, Y_6: $(v_3 - v_1)x_{11} + (u_3 - u_1)x_{21} + (v_3 - v_2)x_{12} + (u_2 - u_1)x_{22}$
(k) Y_2, Y_3, Y_4, Y_5: $x_{12} + b_2 x_{22}$
(l) Y_2, Y_3, Y_4, Y_6: $(v_3 - v_2)x_{11} + (u_2 - u_1)x_{21} + (v_3 - v_1)x_{12} + (u_3 - u_1)x_{22}$

(m) Y_2, Y_3, Y_5, Y_6: $-(v_2-v_1)x_{11}-(u_3-u_2)x_{21}+(v_3-v_2)x_{12}+(u_2-u_1)x_{22}$
(n) Y_2, Y_4, Y_5, Y_6: $x_{13}+b_1x_{23}$
(o) Y_3, Y_4, Y_5, Y_6: $x_{11}+b_3x_{21}$.

For cases (c), (d), (g), (j), (l), (m), it is straightforward to check that the remaining 2 points in $\mathcal{D}_Y$ are on opposite sides of the hyperplane (by making use of the fact the b_2 is neither largest or smallest). The argument for the other cases is analogous to case (a) which is given next.

The equation $x_{11}+b_1x_{21}$ applied to Y_1, Y_2, Y_3, Y_4 leads to the constant $\alpha = u_1+b_1v_3 = u_2+b_2v_2$, and when applied to Y_5, Y_6 it leads to the constant $\beta = u_3+b_1v_1$. For a Y_7 in $\mathcal{C}_Y\backslash\mathcal{D}_Y$, suppose the equation leads to γ. The hyperplane $x_{11}+b_1x_{21}=\alpha$ separates Y_7 from Y_5, Y_6 only if $\beta<\alpha<\gamma$ or $\beta>\alpha>\gamma$. In either case, if X is any convex combination of Y_1, Y_2, Y_3, Y_4, Y_7 with a positive weight for Y_7, then $(1, b_1)X$ is not majorized by $(1, b_1)Y_1$. Hence it is not possible to extend $\mathcal{B}_Y$ towards Y_7.

The above argument works for the nine hyperplanes of the form (a) and any Y_7 not in $\mathcal{D}_Y$ so that the conclusion of the theorem follows. □

PROPOSITION 5.4 *Let Y be a $2\times n$ matrix. If, for a nonnegative nonzero vector (a_1, a_2), $(a_1, a_2)Y$ is a constant times a vector of ones, then $X \prec^{\mathrm{PC}} Y$ is equivalent to $X \prec^{\mathrm{UM}} Y$.*

PROOF Let $\mathbf{y}_1, \mathbf{y}_2$ be the first and second rows of Y. Then $\mathbf{y}_2 = c(1,\ldots,1) - a_1\mathbf{y}_1$ for a constant c. Suppose $X \prec^{\mathrm{PC}} Y$. Then $(a_1, a_2)X = c(1,\ldots,1)$ and $\mathbf{x}_2 = c(1,\ldots,1) - a_1\mathbf{x}_1$, where $\mathbf{x}_1, \mathbf{x}_2$ are the first and second rows of X. $X \prec^{\mathrm{PC}} Y$ implies $\mathbf{x}_1 \prec \mathbf{y}_1$ so that there is a doubly stochastic matrix D such that $\mathbf{x}_1 = \mathbf{y}_1 D$. The preceding equalities imply then that $\mathbf{x}_2 = \mathbf{y}_2 D$ and hence $X \prec^{\mathrm{UM}} Y$. □

We next illustrate the "linear inequalities" approach. Because of the second invariance property in Theorem 2.5, we can assume that the minimum component of each row of Y is 0 so that all components of Y are nonnegative. In the general case, one can subtract the i^{th} row minimum c_i from the i^{th} row, find the extreme points of $\mathcal{A}_{Y^*}$ for the resulting Y^*, and then add c_i to the i^{th} row of each extreme point in $\mathcal{A}_{Y^*}$ to get $\mathcal{A}_Y$. Note that the ordering $\mathbf{a}X \prec \mathbf{a}Y = \mathbf{y}^*$ is equivalent to the following set of inequalities:

$$\sum_{i=1}^{m} a_i(x_{ij_1}+\cdots+x_{ij_k}) \le y^*_{[1]}+\cdots y^*_{[k]}, \quad j_1<\cdots<j_k, k=1,\ldots,n-1,$$

together with the sum constraint. The set $\mathcal{A}_Y$ can be represented by a finite number of inequalities of the above type together with the m row sum constraints. As in the simplex method for linear programming (see for example Gass (1985)), nonnegative slack variables can be introduced for each inequality. The total number of variables, N, is now mn plus the number of nonredundant inequalities, and the number of equations, M, is m plus the number of nonredundant inequalities. The extreme points can now be enumerated by setting in turn $N - M$ of the variables to zero and solving the resulting $M \times M$ linear system if the linear system is nonsingular; nonnegative solutions generated in this way correspond to the extreme points. However there are too many inequalities to prove a general theorem except for in a couple of "small" cases, given in the next two results.

These results involve one or two inversions of the second row relative to the first. For a $2 \times n$ matrix with a strictly increasing first row vector and with a second row vector $\mathbf{z} = (z_1, \ldots, z_n)$, the number of inversions is the cardinality of the set $\{(j_1, j_2) : j_1 < j_2, z_{j_1} > z_{j_2}\}$.

THEOREM 5.5 *Let* $u_1 < u_2 < u_3$, $v_1 < v_2 < v_3$. *For* $Y = \begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_3 & v_2 \end{bmatrix}$, $\mathcal{A}_Y = \mathcal{B}_Y$.

PROOF Without loss of generality, we can assume that $u_1 = v_1 = 0$. From Theorem 3.2, $X \prec^{\mathrm{PC}} Y$ if $\mathbf{a}X \prec \mathbf{a}Y$ for $\mathbf{a} = (1,0)$, $(0,1)$, and $(1,b)$ where $b = (u_3 - u_2)/(v_3 - v_2)$. In addition to nonnegativity constraints, the linear inequalities and equalities imposed by the three majorization orderings are:

$$x_{1j} \le u_3, \qquad x_{2j} \le v_3, \qquad j = 1,2,3,$$

$$0 = u_1 + bv_1 \le x_{1j} + bx_{2j} \le u_2 + bv_3 = u_3 + bv_2, \qquad j = 1,2,3,$$

$$x_{11} + x_{12} + x_{13} = u_2 + u_3, \qquad x_{21} + x_{22} + x_{23} = v_2 + v_3.$$

Note that the lower bound of the second set of inequalities are redundant given the nonnegativity constraints. By adding 9 nonnegative slack variables $s_1, \ldots, s_9$, the 11 resulting equations are:

$$x_{1j} + s_j = u_3, \qquad x_{2j} + s_{3+j} = v_3, \qquad j = 1,2,3,$$

$$x_{1j} + bx_{2j} + s_{6+j} = u_2 + bv_3, \qquad j = 1,2,3,$$

$$x_{11} + x_{12} + x_{13} = u_2 + u_3, \qquad x_{21} + x_{22} + x_{23} = v_2 + v_3.$$

The extreme points can be found by setting in turn 4 of the 15 variables to zero and solving the 11×11 linear system. At most one of x_{1j} can be set to zero and at most one of x_{2j} can be set to zero and at most two of s_7, s_8, s_9 can be set to zero. If four zeros are chosen in this way, then either no nonnegative solution exists or YQ is the solution where Q is a permutation matrix. For other cases, one of $s_1, \ldots, s_6$ is set to zero. By symmetry, assume that $s_1 = 0$ and $x_{11} = u_3$. Substitute $x_{13} = u_2 - x_{12}$, $x_{23} = v_2 + v_3 - x_{21} - x_{22}$. The linear system simplifies to:

$$x_{12} + t_1 = u_2, \quad x_{21} + t_2 = v_2, \quad x_{22} + t_3 = v_3, \quad x_{21} + x_{22} - t_4 \ge v_2,$$

where t_1, t_2, t_3, t_4 are nonnegative slack variables. Now set in turn 3 of the 7 variables to zero. 16 of the resulting systems are nonsingular, of which 4 lead to nonnegative solutions and the other 12 lead to solutions of the form YQ where Q is a permutation matrix. □

The last case to consider for $m = 2, n = 3$ is when there are two inversions of the second row relative to the first. For example, $Y = \begin{bmatrix} u_1 & u_2 & u_3 \\ v_3 & v_1 & v_2 \end{bmatrix}$, where $u_1 < u_2 < u_3$ and $v_1 < v_2 < v_3$. This is interesting because it is the simplest case where $\mathcal{A}_Y$ is strictly larger than $\mathcal{B}_Y$. To further illustrate the two approaches, a combination of them are used to find the extreme points of $\mathcal{A}_Y$ in this case. Not all details are provided.

THEOREM 5.6 *Let $u_1 < u_2 < u_3$ and $v_1 < v_2 < v_3$. Let $Y = \begin{bmatrix} u_1 & u_2 & u_3 \\ v_3 & v_1 & v_2 \end{bmatrix}$. Define $u_{ij} = u_i - u_j$, $v_{ij} = v_i - v_j$, for $i > j$, and $d = u_{31}v_{31} - u_{21}v_{32}$. In addition to the matrices in $\mathcal{D}_Y$, the extreme points of $\mathcal{A}_Y$ include*

$$(5.1) \quad Z = \begin{bmatrix} d^{-1}(u_{31}v_{21}u_1 + u_{32}v_{32}u_2) & d^{-1}(u_{32}v_{32}u_1 + u_{31}v_{21}u_2) & u_3 \\ d^{-1}(u_{32}v_{31}v_2 + u_{21}v_{21}v_3) & d^{-1}(u_{21}v_{21}v_2 + u_{32}v_{31}v_3) & v_1 \end{bmatrix},$$

and the matrices obtained by column permutations of this matrix.

OUTLINE OF DETAILS. Let $Y_1, \ldots, Y_6$ be as in the proof of Theorem 5.3. Let $Y_7 = Y$,

$$Y_8 = \begin{bmatrix} u_1 & u_3 & u_2 \\ v_3 & v_2 & v_1 \end{bmatrix}, Y_9 = \begin{bmatrix} u_2 & u_1 & u_3 \\ v_1 & v_3 & v_2 \end{bmatrix}, Y_{10} = \begin{bmatrix} u_2 & u_3 & u_1 \\ v_1 & v_2 & v_3 \end{bmatrix},$$

$$Y_{11} = \begin{bmatrix} u_3 & u_1 & u_2 \\ v_2 & v_3 & v_1 \end{bmatrix}, Y_{12} = \begin{bmatrix} u_3 & u_2 & u_1 \\ v_2 & v_1 & v_3 \end{bmatrix}.$$

These are the 12 points in $\mathcal{D}_Y$. Furthermore, let $Y_{13} = \begin{bmatrix} u_1 & u_2 & u_3 \\ v_2 & v_3 & v_1 \end{bmatrix}$.

Let $X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{bmatrix}$ be a generic point in the four-dimensional space containing $\mathcal{B}_Y$. It is straightforward to show that the hyperplane $-v_{32}x_{11} + u_{21}x_{22} = u_2v_2 - u_1v_3$ contains Y_1, Y_3, Y_8, Y_9 and separates the other Y's in $\mathcal{D}_Y$ from Y_{13}.

Consider the simplex with vertices $Y_1, Y_3, Y_8, Y_9, Y_{13}$. Let X be the convex combination $\lambda_1 Y_1 + \lambda_3 Y_3 + \lambda_8 Y_8 + \lambda_9 Y_9 + \lambda_{13} Y_{13}$. Let $c_1 = u_{21}/v_{31}$ and $c_2 = u_{31}/v_{32}$. Since $X \prec^{\mathrm{MM}} Y$, by Theorem 3.2, $X \prec^{\mathrm{PC}} Y$ if $(1, c_j)X \prec (1, c_j)Y$, $j = 1, 2$. These two majorization orderings impose the constraints

$$\begin{aligned}(5.2)\ (u_1 + c_1v_3)\lambda_1 \;&+\; (u_2 + c_1v_2)\lambda_3 + (u_1 + c_1v_3)\lambda_8 + (u_2 + c_1v_1)\lambda_9 \\ &+\; (u_1 + c_1v_2)\lambda_{13} \geq u_1 + c_1v_3 = u_2 + c_1v_1,\end{aligned}$$

$$\begin{aligned}(5.3)\ (u_2+c_1v_2)\lambda_1 &+ (u_1+c_1v_3)\lambda_3+(u_3+c_1v_2)\lambda_8+(u_1+c_1v_3)\lambda_9 \\ &+ (u_2+c_1v_3)\lambda_{13} \le u_3+c_1v_2,\end{aligned}$$

$$\begin{aligned}(5.4)\ (u_1+c_2v_3)\lambda_1 &+ (u_2+c_2v_2)\lambda_3+(u_1+c_2v_3)\lambda_8+(u_2+c_2v_1)\lambda_9 \\ &+ (u_1+c_2v_2)\lambda_{13} \ge u_2+c_2v_1,\end{aligned}$$

$$\begin{aligned}(5.5)\ (u_2+c_2v_2)\lambda_1 &+ (u_1+c_2v_3)\lambda_3+(u_3+c_2v_2)\lambda_8+(u_1+c_2v_3)\lambda_9 \\ &+ (u_2+c_2v_3)\lambda_{13} \le u_1+c_2v_3 = u_3+c_2v_2\end{aligned}$$

on $\lambda_1, \lambda_3, \lambda_8, \lambda_9, \lambda_{13}$. If the inequality (5.3) is multiplied by $(u_3+c_2v_2)/(u_3+c_1v_2)$, then $(u_i+c_1v_j)(u_3+c_2v_2)/(u_3+c_1v_2) \le u_i+c_2v_j$, for $(i,j) =$ (2,2), (1,3), (3,2), (2,3) or (3,3). Comparison with (5.5) then makes the inequality (5.3) redundant. Similarly if the inequality (5.4) is multiplied by $(u_2+c_1v_1)/(u_2+c_2v_1)$, then $(u_i+c_2v_j)(u_2+c_1v_1)/(u_2+c_2v_1) \ge u_i+c_1v_j$, for $(i,j) =$ (2,2), (1,3), (2,1), (1,2) or (1,1). Comparison with (5.2) then makes the inequality (5.4) redundant.

For the remaining two inequalities (5.2) and (5.5), substituting $\lambda_{13} = 1-\lambda_1-\lambda_3-\lambda_8-\lambda_9$ leads to

$$c_1v_{32}\lambda_1+u_{21}\lambda_3+c_1v_{32}\lambda_8+c_1v_{32}\lambda_9 \ge c_1v_{32}$$

and

$$c_2v_{32}\lambda_1+u_{21}\lambda_3+u_{21}\lambda_8+u_{21}\lambda_9 \ge u_{21}.$$

Other than the extreme points with $\lambda_1 = 1$, $\lambda_3 = 1$, $\lambda_8 = 1$, $\lambda_9 = 1$, the single nontrivial extreme point from these inequalities is when $\lambda_1 = u_{21}v_{21}/d$, $\lambda_3 = u_{32}v_{32}/d$, $\lambda_8 = \lambda_9 = 0$ (and $\lambda_{13} = u_{32}v_{21}/d$). This leads to the matrix given in (5.1).

Hence the separating hyperplane approach has extended $\mathcal{A}_Y$ beyond $\mathcal{B}_Y$. However in this case, it cannot show that ZQ, where Q is a permutation matrix, are the only other extreme points of $\mathcal{A}_Y$. The linear inequalities approach can be used to complete this last step. The details are more tedious than in Theorem 5.5.

Let $u_1 = v_1 = 0$ now without loss of generality. The linear inequalities and equalities imposed by $X \prec^{\mathrm{PC}} Y$ are:

$$x_{1j} \le u_3, \qquad x_{2j} \le v_3, \qquad j = 1,2,3,$$

$$u_2 \le x_{1j}+c_1x_{2j} \le u_3+c_1v_2, \qquad j = 1,2,3,$$

$$u_2 \le x_{1j}+c_2x_{2j} \le c_2v_3, \qquad j = 1,2,3,$$

$$x_{11}+x_{12}+x_{13} = u_2+u_3, \quad x_{21}+x_{22}+x_{23} = v_2+v_3.$$

Slack variables $s_1, \ldots, s_{18}$ can be added (or subtracted) corresponding to the 18 inequalities. This leads to 20 linear equations in 24 variables. At most one of x_{1j} can be set to zero and at most one of x_{2j} can be set to zero, so that at least two of the slack variables must to set to zero to find the extreme points. By symmetries and constraints, there are 21 pairs $(s_{i_1}, s_{i_2}) = (0,0)$ that have to be considered. Once a pair of slack variables is set to zero, the variables x_{ij} and the remaining s_k's can be expressed in terms of two slack variables, say s_{j_1}, s_{j_2}. Symbolic manipulation software helps for the algebra in this reduction. Extreme points from inequalities in two variables (in symbols) can be solved by hand. The outcome of all this is that the only extreme points are of the form YQ, Y_1Q, ZQ, where Q is a permutation matrix. □

REMARKS The two techniques for identifying the structure of $\mathcal{A}_Y$ can in general be implemented in computer programs. Numerical examples studied in this way may lead to more general theorems. Further research will attempt to determine when $\mathcal{A}_Y = \mathcal{B}_Y$ and when $\mathcal{A}_Y$ is strictly larger than $\mathcal{B}_Y$ for general (m, n) not covered by theorems in this paper. In addition, it would be useful to discover an approach that does not require enumeration.

The derivation in the first part of the proof of Theorem 5.6 suggest the following (non–simple) operation for extending to points in $\mathcal{A}_Y \backslash \mathcal{B}_Y$. Choose a matrix Y_0 in $\mathcal{C}_Y \backslash \mathcal{D}_Y$ satisfying some properties. Find a separating hyperplane that separates Y_0 from some points in $\mathcal{C}_Y$. Take a convex combination of Y_0 and matrices in $\mathcal{C}_Y$ that on the hyperplane. The linear inequalities imposed by the PC-ordering will put constraints on the possible convex combinations. A goal is to identify those Y_0 where the coefficient of Y_0 in convex combinations can be definitely positive.

Acknowledgements. The authors thank the referee for a careful reading and many detailed comments.

REFERENCES

ARNOLD, B. C. (1987). *Majorization and the Lorenz Order.* Springer-Verlag, New York.

BOLAND, P. J. AND PROSCHAN, F. (1988). Multivariate arrangement increasing functions with applications in probability and statistics. *J. Multivariate Anal.* **25** 286–298.

GASS, S. I. (1985). *Linear Programming – Methods and Applications.* Fifth edition. McGraw-Hill, New York.

HOLLANDER, M., PROSCHAN, F. AND SETHURAMAN, J. (1977). Functions decreasing in transposition and their applications in ranking problems. *Ann. Statist.* **5** 722–733.

JOE, H. (1985). An ordering of dependence for contingency tables. *Linear Alg. Appl.* **70** 89–103.

MARSHALL, A. W. AND OLKIN, I. (1979). *Inequalities: Theory of Majorization and its Applications.* Academic Press, New York.

TONG, Y. L. (1989). Probability inequalities for n–dimensional rectangles via multivariate majorization. In *Contributions to Probability and Statistics: Essays in Honor of Ingram Olkin.* L. J. Gleser, M. D. Perlman, S. J. Press, and A. R. Sampson, eds. Springer-Verlag, New York, 146-159.

DEPARTMENT OF STATISTICS
UNIVERSITY OF BRITISH COLUMBIA
VANCOUVER, B.C. CANADA V6T 1Z2

DEPARTMENT OF STATISTICS
OHIO STATE UNIVERSITY
COLUMBUS, OH 43210