Journal of Convex Analysis
Volume 17 (2010), No. 2, 485–507

On Malamud Majorization and the Extreme Points of its Level Sets

Pal Fischer
Department of Mathematics and Statistics,
University of Guelph, Guelph, Ontario, Canada N1G 2W1
pfischer@uoguelph.ca

Hristo Sendov*
Department of Statistical and Actuarial Sciences,
The University of Western Ontario, London, Ontario, Canada N6A 5B7
hssendov@stats.uwo.ca

Received: February 9, 2009
Revised manuscript received: June 19, 2009

We consider two types of majorization relationships between sequences of vectors $y = (y_k)_{k=1}^m$ and $x = (x_k)_{k=1}^{\ell}$ in $\mathbb{R}^n$ with $\ell \leq m$. It is said that x is majorized by y, $x \prec y$, if the sum of any k vectors from x is in the convex hull of all possible sums of k vectors from y. It is said that x is doubly stochastically majorized by y, $x \prec_{\rm ds} y$, if $x_k = \sum_{j=1}^m m_{kj} y_j$, $k = 1, ..., \ell$, for some doubly stochastic matrix $M = (m_{kj})_{k,j=1}^{m,m}$.

In [5], S. M. Malamud formulated the problem of finding a geometric condition guaranteeing that $x \prec y \Leftrightarrow x \prec_{\rm ds} y$. We answer this question in the case when the vectors in y are distinct and are extreme points of their convex hull. In particular, we derive a geometric characterization of the extreme points of the level set $L^2_{\prec}(y) = \{x : x \prec y\}$. Finally, we derive a set of algebraic conditions that characterize the extreme points of $L^{\ell}_{\prec}(y) = \{x : x \prec y\}$ for any $\ell \leq m$ and y.

Keywords: Convex set, doubly stochastic majorization, Malamud majorization, extreme point, convex function, CVS class

2000 Mathematics Subject Classification: Primary 52B11, 42A20

1. Introduction

The notion of majorization between two vectors in $\mathbb{R}^n$ is a classical subject with innumerable applications in areas ranging from matrix analysis to statistics, see for example the monographs [1] and [6]. Given two vectors x_1 and y_1 in $\mathbb{R}^n$ we say that x_1 is *majorized* by y_1 if for every $k = 1, ..., n$ the sum of the k-th largest coordinates of x_1 is not larger than the sum of the k-th largest coordinates of y_1, with the sums equal when $k = n$.

In this work we are interested in the properties of a new type of majorization between two sequences $x = (x_k)_{k=1}^{\ell}$ and $y = (y_k)_{k=1}^m$ of vectors in $\mathbb{R}^n$. This new type of majorization was introduced by S. M. Malamud in [5], who posed several natural

*Research supported by NSERC

ISSN 0944-6532 / \$ 2.50

questions about its properties. For any natural number $k \geq 1$, define the set

$$\mathbb{N}_k := \{1, ..., k\}.$$

Definition 1.1. Let $x = (x_k)_{k=1}^{\ell}$ and $y = (y_k)_{k=1}^{m}$ be two sequences of vectors in $\mathbb{R}^n$ with $\ell \leq m$. We say that x *is (Malamud) majorized by* y and write $x \prec y$ if for all $k \in \mathbb{N}_\ell$ we have

$$\begin{aligned} &\operatorname{conv}\{x_{i_1} + \cdots + x_{i_k} : 1 \leq i_1 < \cdots < i_k \leq \ell\} \\ \subset\ &\operatorname{conv}\{y_{i_1} + \cdots + y_{i_k} : 1 \leq i_1 < \cdots < i_k \leq m\}. \end{aligned}$$

In view of this definition, we introduce the following notation. For a sequence $y = (y_k)_{k=1}^{m}$ of vectors in $\mathbb{R}^n$, we denote by $A_k(y)$, or simply by A_k, the set:

$$A_k(y) := \operatorname{conv}\{y_{i_1} + \cdots + y_{i_k} : 1 \leq i_1 < \cdots < i_k \leq m\} \text{ for } k \in \mathbb{N}_m. \tag{1}$$

The next definition can be found in [5]. Here, we are following its presentation from [4], where it was not introduced as a relation. Recall that a square matrix is *doubly stochastic* if it has non-negative entries whose sums along each row and column is one.

Definition 1.2. Let $x = (x_k)_{k=1}^{\ell}$ and $y = (y_k)_{k=1}^{m}$ be two sequences of vectors in $\mathbb{R}^n$ with $\ell \leq m$. We say that x *is doubly stochastically majorized by* y and write $x \prec_{\mathrm{ds}} y$ if there is a doubly stochastic matrix $M = (m_{ij})_{i,j=1}^{m,m}$ such that

$$x_i = \sum_{j=1}^{m} m_{ij} y_j \text{ for all } i \in \mathbb{N}_\ell. \tag{2}$$

Note that when $n = 1$ (and $\ell = m$) both types of majorization reduce to the classical notion of majorization between two vectors in $\mathbb{R}^m$. Since neither majorization depends on the order of the vectors in the sequences x and y, we often represent the sequence y as an $n \times m$ matrix $(y_1, ..., y_m)$ having column i equal to y_i, and similarly for x. The next easy result, stated in [5], shows that the doubly stochastic majorization is a more restrictive notion.

Lemma 1.3. *If $x = (x_k)_{k=1}^{\ell}$ and $y = (y_k)_{k=1}^{m}$ are two sequences of vectors in $\mathbb{R}^n$ with $\ell \leq m$ then, $x \prec_{\mathrm{ds}} y$ implies that $x \prec y$.*

The converse of Lemma 1.3 does not hold in general. In the following example, given by Malamud (see Example 2.11 in [5]), the relation $x \prec y$ holds but the relation $x \prec_{\mathrm{ds}} y$ does not hold.

Example 1.4. Let $n = 2$ and $m = \ell = 4$. Set

$$\begin{aligned} x = (x_1, x_2, x_3, x_4) &:= \left((12, 12)^T, (12, 12)^T, (5, 3)^T, (3, 5)^T\right), \\ y = (y_1, y_2, y_3, y_4) &:= \left((8, 16)^T, (16, 8)^T, (0, 0)^T, (8, 8)^T\right). \end{aligned}$$

We say that an $\ell \times m$ rectangular matrix is *row stochastic* if it has non-negative entries whose sum along each row is one and along each column is at most one, necessarily

$\ell \leq m$. It is not too difficult to see that every rectangular row stochastic matrix can be completed to a square doubly stochastic matrix. Thus, for two sequences $x = (x_k)_{k=1}^{\ell}$ and $y = (y_k)_{k=1}^{m}$ of vectors in $\mathbb{R}^n$ with $\ell \leq m$ we have $x \prec_{\mathrm{ds}} y$ if and only if there exist a row stochastic matrix $M = (m_{ij})_{i,j=1}^{\ell,m}$ such that (2) holds.

For any sequence $y = (y_k)_{k=1}^{m}$ of vectors in $\mathbb{R}^n$ we consider the *level sets* of the two majorizations

$$L^{\ell}_{\prec_{\mathrm{ds}}}(y) := \{(x_1, ..., x_\ell) \in \mathbb{R}^{n\times\ell} : (x_k)_{k=1}^{\ell} \prec_{\mathrm{ds}} y\}, \quad \text{and}$$
$$L^{\ell}_{\prec}(y) := \{(x_1, ..., x_\ell) \in \mathbb{R}^{n\times\ell} : (x_k)_{k=1}^{\ell} \prec y\}.$$

It is easy to see that both these sets are closed and convex with $L^{\ell}_{\prec_{\mathrm{ds}}}(y) \subset L^{\ell}_{\prec}(y)$.

In this work, we are interested in the three questions Malamud posed in [5, page 4049] regarding the relationships between the two types of matrix majorizations described above. Question 1: formulate geometric conditions which together with $x \prec y$ imply that $x \prec_{\mathrm{ds}} y$. Question 2: characterize the extreme points of the level set $L^{\ell}_{\prec}(y)$. Question 3: under what conditions on the sequence y do we have $L^{\ell}_{\prec}(y) = L^{\ell}_{\prec_{\mathrm{ds}}}(y)$?

Notice that Question 3 is an easy consequence of Question 2. Indeed, since both level sets are polytopes they are equal if and only if they have the same set of extreme points. For similar reasons, a geometric description of the extreme points of $L^{\ell}_{\prec}(y)$ answers Question 1 as well. On the other hand, answering Question 2 turns out to be, as we will see, a more difficult problem.

Answers to Questions 1 and 3 have been known in some special cases. For example, if the vectors $(y_k)_{k=1}^{m}$ are affinely independent and $\ell = m$ then, according to [5, Proposition 2.13] the orders $\prec$ and $\prec_{\mathrm{ds}}$ are equivalent and we have $L^{\ell}_{\prec}(y) = L^{\ell}_{\prec_{\mathrm{ds}}}(y)$. If the vectors $(y_k)_{k=1}^{m}$ are affinely independent and $\ell < m$ then, $x \prec y$, if and only if, there are vectors $x_{\ell+1}$,...,x_m such that $(x_k)_{k=1}^{m} \prec y$ (see Theorem 2.3 below). Thus, again the orders $\prec$ and $\prec_{\mathrm{ds}}$ are equivalent and $L^{\ell}_{\prec_{\mathrm{ds}}}(y) = L^{\ell}_{\prec}(y)$. Finally, when $m = 3$ according to [5, Proposition 2.13] we have $L^{\ell}_{\prec_{\mathrm{ds}}}(y) = L^{\ell}_{\prec}(y)$.

We give a complete answer to Questions 1 and 3 in the case when a) the vectors in y are distinct; and b) every vector in y is an extreme point of A_1, see Theorem 4.22 and Corollary 4.23. We achieve this by deriving a geometric characterization of the extreme points of the level set $L^{2}_{\prec}(y)$ thus, giving a partial answer to Question 2 under the same conditions for y. As a consequence of our results we construct a wide class of convex functions on $\mathbb{R}^n$, see Theorem 4.26, that cannot be approximated by compositions of a convex function on $\mathbb{R}$ with a linear function on $\mathbb{R}^n$, see (5). Finally, in Section 5, see Theorem 5.2, we present an algebraic characterization of the extreme points of $L^{\ell}_{\prec}(y)$ for arbitrary $\ell \in \mathbb{N}_m$ and y.

We conclude this section with a reduction result that plays a crucial role.

Theorem 1.5. *Let* $x = (x_k)_{k=1}^{\ell}$ *and* $y = (y_k)_{k=1}^{m}$ *be two sequences of vectors in* $\mathbb{R}^n$ *with* $\ell \leq m$. *Suppose that*

$$\forall j_1, j_2 \in \mathbb{N}_\ell,\ j_1 \neq j_2 \ \ ((x_{j_1}, x_{j_2}) \prec y \implies (x_{j_1}, x_{j_2}) \prec_{\mathrm{ds}} y)\,. \tag{3}$$

Then, we have

$$x \prec y \implies x \prec_{\mathrm{ds}} y.$$

Proof. Suppose that (3) holds and that $x \prec y$, we have to show that $x \prec_{\rm ds} y$. The proof is by induction on ℓ with base case $\ell = 2$ holding by assumption. Suppose the result is true for $\ell - 1 < m$. For any k the sequence $x^k := (x_1, ..., \hat{x}_k, ..., x_\ell)$, with x_k omitted, has length $\ell - 1$ and $x^k \prec y$. Thus, by the induction hypothesis, $x^k \prec_{\rm ds} y$. That is, there is a row stochastic matrix $M^k = (m^k_{ij})^{\ell-1,m}_{i,j=1}$ such that

$$x^k_i := \sum_{j=1}^{m} m^k_{ij} y_j \quad \text{for all } i \in \mathbb{N}_\ell \setminus \{k\}.$$

Define an $\ell \times m$ matrix $\bar{M}^k$ obtained from M^k by adding a row of zeros right after the $(k-1)$-st row and let

$$M := \frac{1}{\ell - 1} \sum_{k=1}^{\ell} \bar{M}^k.$$

It is easy to see that $M = (m_{ij})^{\ell,m}_{i,j=1}$ is row stochastic and satisfies (2). $\square$

Corollary 1.6. *Let $x = (x_k)^\ell_{k=1}$ and $y = (y_k)^m_{k=1}$ be two sequences of vectors in $\mathbb{R}^n$ with $\ell \le m$. Then, $x \prec_{\rm ds} y$ if and only if for any two distinct indexes $k_1, k_2 \in \mathbb{N}_\ell$ there are convex combinations*

$$x_{k_i} = \sum_{j=1}^{m} \alpha_{ij} y_j \quad \text{for } i = 1, 2,$$

with $\alpha_{1j} + \alpha_{2j} \le 1$ for all $j = 1, ..., m$.

In view of Theorem 1.5 it suffices to answer Question 1 only in the case $\ell = 2$. That is, it suffices to find a geometric condition on the vectors in y such that for any $x_1, x_2 \in \mathbb{R}^n$ with $\{x_1, x_2\} \prec y$ we have $\{x_1, x_2\} \prec_{\rm ds} y$. For that reason, in Section 4 our efforts are focused on deriving a geometric characterization of the extreme points of the level set $L^2_\prec(y)$.

2. Characterizations of the majorizations

The following characterization of the relation $x \prec_{\rm ds} y$ was given with the aid of Hardy-Littlewood-Polya-types of inequalities by Fischer and Holbrook in [4]. Here, we are giving a slightly better formulation of the second part of this result.

Theorem 2.1. *Let $x = (x_k)^\ell_{k=1}$ and $y = (y_k)^m_{k=1}$ be two sequences of vectors in $\mathbb{R}^n$ with $\ell \le m$ and let $K = \operatorname{conv}\{y_1, ..., y_m\}$. Then the following are equivalent:*

(1) $x \prec_{\rm ds} y$;

(2) *for any nonnegative continuous convex function f on K we have*

$$\sum_{i=1}^{\ell} f(x_i) \le \sum_{i=1}^{m} f(y_i), \tag{4}$$

provided each $x_i \in K$ for $i \in \mathbb{N}_\ell$.

A similar result for the relation $x \prec y$ was obtained in [5], with the aid of a new class of functions CVS $(\mathbb{R}^n)$. This class is the closed cone (in the pointwise topology of the set of convex functions defined on $\mathbb{R}^n$), generated by the set of convex functions

$$\mathrm{CVS}\,(\mathbb{R}^n) := \{f(\langle x, y\rangle) : f \text{ is a convex function on } \mathbb{R},\ y \in \mathbb{R}^n\}. \tag{5}$$

It is remarked in [5] that CVS $(\mathbb{R}^n)$ is a proper subset of the set of all convex functions defined on $\mathbb{R}^n$. Also, it is stated in the same paper that F. V. Petrov constructed an explicit example of a function that belongs to the difference of these two sets. Immediately after Example 2.4 below, we exhibit another such function that will be considerably generalized in Theorem 4.26. First, we need the following characterization of the majorization relation $\prec$, given by Malamud in [5].

Theorem 2.2. *Let $x = (x_k)_{k=1}^{\ell}$ and $y = (y_k)_{k=1}^{m}$ be two sequences of vectors in $\mathbb{R}^n$ with $\ell \le m$. Then the following are equivalent:*

(1) $x \prec y$;
(2) $(\langle x_1, h\rangle, ..., \langle x_\ell, h\rangle) \prec (\langle y_1, h\rangle, ..., \langle y_m, h\rangle)$ *for all* $h \in \mathbb{R}^n$;
(3) *for any nonnegative* $f \in CVS\,(\mathbb{R}^n)$ *we have*

$$\sum_{i=1}^{\ell} f(x_i) \le \sum_{i=1}^{m} f(y_i).$$

The next characterization, needed for our work, is due to F. V. Petrov, see [5].

Theorem 2.3. *Let $x = (x_k)_{k=1}^{\ell}$ and $y = (y_k)_{k=1}^{m}$ be two sequences of vectors in $\mathbb{R}^n$ with $\ell \le m$. Then the following are equivalent:*

(1) $x \prec y$;
(2) *for any* $k \in \mathbb{N}_\ell$ *and any* k*-tuple of indices* $1 \le i_1 < \cdots < i_k \le \ell$ *there exist numbers* $\{\beta_j\}_{j=1}^m$ *in* $[0,1]$ *with* $\sum_{j=1}^m \beta_j = k$ *such that*

$$x_{i_1} + \cdots + x_{i_k} = \sum_{j=1}^{m} \beta_j y_j;$$

(3) *there exist vectors* $x_{\ell+1}, ..., x_m \in \mathbb{R}^n$ *such that* $(x_k)_{k=1}^m \prec (y_k)_{k=1}^m$.

In the next example, ℓ is equal to the dimension of the linear space spanned by the vectors in y and $x \prec y$ but $x \not\prec_{\mathrm{ds}} y$. Note that not all y's are extreme points of their convex hull.

Example 2.4. Let $n = 2$, $\ell = 2$, and $m = 4$. Set

$$x = (x_1, x_2) := \left((8/3, 16/3)^T, (16/3, 8/3)^T\right);$$
$$y = (y_1, y_2, y_3, y_4) := \left((8, 16)^T, (16, 8)^T, (0, 0)^T, (8, 8)^T\right).$$

We show that $x \prec y$, but $x \not\prec_{\mathrm{ds}} y$. Indeed, there is only one way to write x_1 and x_2 as a convex combination of the y's namely as $x_1 = (1/3)y_1 + (2/3)y_3$ and $x_2 = (1/3)y_2 + (2/3)y_3$. Next, notice that $x_1 + x_2 = y_3 + y_4$. Hence, we conclude by Theorem 2.3 that $x \prec y$. On the other hand, by the uniqueness of the representations of x_1 and x_2 and Corollary 1.6 we conclude $x \not\prec_{\mathrm{ds}} y$.

Given sequences $x = (x_k)_{k=1}^{\ell}$ and $y = (y_k)_{k=1}^{m}$ with $x \prec y$ and $x \not\prec_{\mathrm{ds}} y$, it follows from Theorem 2.1 and Theorem 2.2, that there exists a continuous convex function $f : \mathbb{R}^n \to \mathbb{R}$ violating Inequality (4) and thus, $f \notin \mathrm{CVS}\,(\mathbb{R}^n)$. Furthermore, in view of Theorem 1.5, there are distinct indexes $j_1, j_2 \in \mathbb{N}_\ell$ and a continuous convex function $f : \mathbb{R}^n \to \mathbb{R}$ such that

$$f(x_{j_1}) + f(x_{j_2}) > \sum_{i=1}^{m} f(y_i).$$

For example, this is the case in both Examples 1.4 and 2.4 and the convex function

$$f(w) = \mathrm{dist}\,(w, \mathrm{conv}\,\{y_1, y_2, y_4\}).$$

Imitating this example, in Theorem 4.26, we exhibit a wide family of continuous convex functions $f : \mathbb{R}^n \to \mathbb{R}$ that are not in the class $\mathrm{CVS}\,(\mathbb{R}^n)$.

The next lemma is needed in the sequel. In particular, it implies the equivalency of parts (*1*) and (*2*) in Theorem 2.3. The notion of the support function, $\rho_A(z) = \sup_{x\in A}\langle x, z\rangle$ of a convex set $A \subset \mathbb{R}^n$, where $z \in \mathbb{R}^n$, is employed in the proof.

Lemma 2.5. *Let $y = (y_k)_{k=1}^{m}$ be a sequence of vectors in $\mathbb{R}^n$. For every $k \in \mathbb{N}_m$, the following representation holds*

$$A_k(y) = \left\{ \sum_{j=1}^{m} \beta_j y_j : \beta_j \in [0,1],\ \sum_{j=1}^{m} \beta_j = k,\ j = 1, ..., m \right\}.$$

Proof. Denote the set on the right-hand side by B_k. Clearly both sets A_k and B_k are closed and convex. It is easy to see the inclusion $A_k \subset B_k$. To see the opposite inclusion consider the support functions of A_k and B_k respectively. Fix a vector $z \in \mathbb{R}^n$ and observe that

$$\rho_{B_k}(z) = \max \left\{ \sum_{j=1}^{m} \beta_j \langle z, y_j\rangle : \beta_j \in [0,1],\ \sum_{j=1}^{m} \beta_j = k,\ j \in \mathbb{N}_m \right\}$$

is equal to the sum of the k largest elements in the sequence $\{\langle z, y_j\rangle\}_{j=1}^{m}$. We conclude that $\rho_{B_k}(z) = \langle z, y_{j_1} + \cdots + y_{j_k}\rangle$ for some point $y_{j_1} + \cdots + y_{j_k}$ of A_k thus, $\rho_{B_k}(z) \le \rho_{A_k}(z)$. Hence, $B_k \subset A_k$, by [8, Corollary 13.1.1]. □

Note that the sets $A_k(y)/k$ are a particular case of the *convex interval hull* of y considered in [3]. We show next, that in part (*3*) of Theorem 2.3 the vectors $x_{\ell+1}, ..., x_m$ can be chosen to be all equal to a vector that can be selected independently of $f \in \mathrm{CVS}\,(\mathbb{R}^n)$. A similar result for the relation $x \prec_{\mathrm{ds}} y$ was shown in [4].

Theorem 2.6. *Let $x = (x_k)_{k=1}^{\ell}$ and $y = (y_k)_{k=1}^{m}$ be two sequences of vectors in $\mathbb{R}^n$ with $\ell \le m$. Then the following are equivalent:*

(*1*) $x \prec y$;

(*2*) *there exists a vector $x' \in A_1(y)$ such that with $x_{\ell+1} = ... = x_m := x'$ we have $(x_k)_{k=1}^{m} \prec (y_k)_{k=1}^{m}$. In particular, we may choose*

$$x' := \frac{\sum_{i=1}^{m} y_i - \sum_{i=1}^{\ell} x_i}{m - \ell}; \tag{6}$$

(*3*) *there is an* $x' \in A_1(y)$ *such that for any nonnegative* $f \in CVS(\mathbb{R}^n)$ *we have*

$$\sum_{i=1}^{\ell} f(x_i) + (m-\ell) f(x') \le \sum_{i=1}^{m} f(y_i).$$

Proof. In order to show that (*1*) implies (*2*) note, as in F. V. Petrov's proof of Theorem 2.3, that defining

$$x_{\ell+1} := \frac{\sum_{i=1}^m y_i - \sum_{i=1}^\ell x_i}{m-\ell}, \tag{7}$$

yields $(x_k)_{k=1}^{\ell+1} \prec (y_k)_{k=1}^m$. F. V. Petrov stated that we can continue by induction and we observe that

$$\begin{aligned} x_{\ell+2} := \frac{\sum_{i=1}^m y_i - \sum_{i=1}^{\ell+1} x_i}{m-\ell-1} &= \frac{\sum_{i=1}^m y_i - \sum_{i=1}^\ell x_i - \frac{1}{m-\ell}\left(\sum_{i=1}^m y_i - \sum_{i=1}^\ell x_i\right)}{m-\ell-1} \\ &= \frac{\left(\sum_{i=1}^m y_i - \sum_{i=1}^\ell x_i\right)\left(1 - \frac{1}{m-\ell}\right)}{m-\ell-1} = x_{\ell+1}. \end{aligned}$$

Therefore, following F. V. Petrov's procedure, we see that $x_{\ell+1} = \ldots = x_m$. It follows from part (*3*) of Theorem 2.3 that $x' \in A_1$. In view of Theorem 2.2 and Theorem 2.3, the proof is complete. □

3. Polyhedral set-valued maps

A *set-valued mapping* $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is a correspondence assigning to each $x \in \mathbb{R}^n$ a subset $S(x)$, possibly empty, of $\mathbb{R}^m$. The *graph* of S is the set

$$\operatorname{gph} S = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : y \in S(x)\},$$

and the *domain* of S is the set

$$\operatorname{dom} S = \{x \in \mathbb{R}^n : S(x) \neq \emptyset\}.$$

The mapping S is *closed-valued* if $S(x)$ is a closed set for all x. Denote by $\mathcal{B}_m$ the closed unit ball in $\mathbb{R}^m$ centered at the origin. The mapping S is called *Lipschitz continuous* relative to a (nonempty) set $D \subset \mathbb{R}^n$ if $D \subset \operatorname{dom} S$, S is closed-valued on D, and there exists $\kappa \ge 0$ such that

$$S(x') \subset S(x) + \kappa \|x' - x\| \mathcal{B}_m \quad \text{for all } x', x \in D.$$

The mapping S is called *polyhedral convex* if its graph is a polyhedral convex set. Note that polyhedral convexity of S is equivalent, see [2, Section 3C], to the existence of a positive integer r, matrices $A \in \mathbb{R}^{r\times n}$, $B \in \mathbb{R}^{r\times m}$ and a vector $q \in \mathbb{R}^r$ such that

$$S(x) = \{y \in \mathbb{R}^m : Ax + By \le q\} \quad \text{for all } x \in \mathbb{R}^n. \tag{8}$$

We have the following result.

Theorem 3.1. *Any polyhedral convex mapping $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is Lipschitz continuous relative to its domain.*

The proof of Theorem 3.1 uses the Hoffman lemma regarding approximate solutions of system of linear inequalities, see for example [2, Theorem 3C.3].

Given a sequence of vectors $y = (y_k)_{k=1}^m$ in $\mathbb{R}^n$ consider the set-valued mapping $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, defined by

$$S(x) = \left\{ \alpha \in \mathbb{R}^m : x = \sum_{j=1}^m \alpha_j y_j,\ \sum_{j=1}^m \alpha_j = 1,\ a_j \geq 0,\ j \in \mathbb{N}_m \right\}. \tag{9}$$

It is not difficult to see that S is a polyhedral convex mapping. Indeed, let $r = 2n+m+2$ and define matrices $A \in \mathbb{R}^{r\times n}$, $B \in \mathbb{R}^{r\times m}$ and a vector $q \in \mathbb{R}^r$ by

$$A = \begin{pmatrix} -I_{n\times n} \\ I_{n\times n} \\ 0_{1\times n} \\ 0_{1\times n} \\ 0_{m\times n} \end{pmatrix}, \quad B = \begin{pmatrix} y \\ -y \\ e \\ -e \\ -I_{m\times m} \end{pmatrix}, \quad q = \begin{pmatrix} 0_{n\times 1} \\ 0_{n\times 1} \\ 1 \\ -1 \\ 0_{m\times 1} \end{pmatrix},$$

where I is the square identity matrix of indicated size, and 0 is the zero matrix of indicated size, y is the $n \times m$ matrix with columns $y_1, ..., y_m$, and e is the all-one row-vector of length m. One can see that we have $S(x) = \{\alpha \in \mathbb{R}^m : Ax+B\alpha \leq q\}$ showing that S is a polyhedral convex mapping.

Clearly, $\operatorname{dom} S = \operatorname{conv}\{y_1, ..., y_m\}$ and by Theorem 3.1 it is Lipschitz continuous relative to $\operatorname{dom} S$. Note also that the mapping S is compact-valued. A simple limiting argument leads to the following result.

Corollary 3.2. *For every sequence $\{x'_n\}$ in $\operatorname{conv}\{y_1, ..., y_m\}$ converging to x, there is an $\alpha \in S(x)$ and a sequence $\{\alpha'_{n_s}\}$ converging to α, such that $\alpha'_{n_s} \in S(x'_{n_s})$ for all $s = 1, 2, ...,$ where $\{n_s\}_{s=1}^{\infty}$ is a subsequence of $1, 2, ...$*

4. Answers to Malamud's Questions 1 and 3

In this section, we consider sequences of vectors $y = (y_k)_{k=1}^m$ with the following properties: 1) every vector in y is an extreme point of A_1; and 2) the vectors in y are pairwise distinct.

Characterizing all of the extreme points of $L^{\ell}_{\prec}(y)$ thus, answering Question 2, appears to be a difficult problem even under these assumptions. Still some of the extreme points can be seen immediately. Indeed, every ℓ-tuple in the set

$$\{(y_{i_1}, ..., y_{i_\ell}) : 1 \leq i_j \leq m \text{ for } j \in \mathbb{N}_\ell, \text{ and all } i_1, ..., i_\ell \text{ distinct}\} \tag{10}$$

is an extreme point of both level sets $L^{\ell}_{\prec_{\mathrm{ds}}}(y)$ and $L^{\ell}_{\prec}(y)$. Thus, the problem is to find whether $L^{\ell}_{\prec}(y)$ has other extreme points. Note that the extreme points of the set $L^{\ell}_{\prec_{\mathrm{ds}}}(y)$ are easy to characterize.

Theorem 4.1. *The extreme points of $L^{\ell}_{\prec_{\mathrm{ds}}}(y)$ are precisely given by (10).*

Proof. If $x \in L^{\ell}_{\prec_{\rm ds}}(y)$ then, according to Definition 1.2 there is a doubly stochastic matrix M such that (2) holds. According to Birkhoff's theorem M is a convex combination of permutation matrices. Taking the first ℓ rows of each permutation matrix and mapping it naturally into an ℓ-tuple in (10) shows the result. $\square$

4.1. The extreme points of $L^2_{\prec}(y)$

We completely characterize the extreme points of the set $L^2_{\prec}(y)$ under the two assumptions on y. The answer still depends on the particular configuration of the vectors y. The rather long build-up that follows, culminates in Theorem 4.17, the main result of the section. Its corollaries contain the complete answer to Questions 1 and 3.

Lemma 4.2. *Suppose that $(x_1, x_2) \in L^2_{\prec}(y)$ and that $x_1 \in y$. Then (x_1, x_2) can be represented as a convex combination of $\{(y_i, y_j) : 1 \le i \ne j \le m\}$.*

Proof. Suppose without loss of generality that $x_1 = y_1$. In order to prove the lemma it is sufficient to show that $x_2 \in \operatorname{conv}\{y_2, ..., y_m\}$. Since $x_2 \in A_1$, there is a convex representation

$$x_2 = \sum_{i=1}^{m} \alpha_i y_i. \tag{11}$$

Since $y_1 + x_2 \in A_2$, there are numbers β_1,...,β_m in $[0, 1]$ with $\sum_{i=1}^m \beta_i = 2$ such that $y_1 + x_2 = \sum_{i=1}^m \beta_i y_i$. If $\beta_1 = 1$ then, the proof is complete, since y_1 cancels on both sides and we get the required representation of x_2. Assuming that $\beta_1 \in [0, 1)$, the last equality implies

$$\alpha_1 y_1 + \frac{\alpha_1}{1-\beta_1} x_2 = \sum_{i=2}^{m} \frac{\alpha_1 \beta_i}{1-\beta_1} y_i.$$

Adding this equality to (11), canceling $\alpha_1 y_1$ on both sides, and solving for x_2, we obtain

$$x_2 = \sum_{i=2}^{m} \frac{(1-\beta_1)\alpha_i + \alpha_1\beta_i}{1+\alpha_1-\beta_1} y_i.$$

It is easy to check that the coefficients $\frac{(1-\beta_1)\alpha_i+\alpha_1\beta_i}{1+\alpha_1-\beta_1}$ are in $[0, 1]$ and that they sum up to 1. This concludes the proof. $\square$

Corollary 4.3. *Suppose that $(x_1, x_2) \in L^2_{\prec}(y)$ and suppose that there are convex representations $x_i = \sum_{j=1}^m \alpha_{ij} y_j$, $i = 1, 2$, with $\alpha_{1j_0} + \alpha_{2j_0} > 1$ for some $j_0 \in \mathbb{N}_m$. Then, there is a representation*

$$x_1 + x_2 = \sum_{i=1}^{m} \beta_i y_i \quad \textit{with } \beta_1, ..., \beta_m \in [0, 1], \ \sum_{i=1}^{m} \beta_i = 2,$$

such that $\beta_{j_0} = 1$.

Proof. Suppose without loss of generality that $j_0 = 1$. Define the vectors

$$x'_1 := y_1,$$

$$x'_2 := (\alpha_{11} + \alpha_{21} - 1) y_1 + \sum_{j=2}^{m} (\alpha_{1j} + \alpha_{2j}) y_j.$$

Note first, that $\alpha_{11} + \alpha_{21} < 2$. Then, since $\alpha_{1j} + \alpha_{2j} < 1$ for all indexes $j \neq j_0$, we conclude that $x_1', x_2' \in A_1$. In addition, it is clear that $x_1' + x_2' = x_1 + x_2 \in A_2$. Thus, by Lemma 4.2, (x_1', x_2') is a convex combination of $\{(y_i, y_j) : 1 \leq i \neq j \leq m\}$. Since y_1 is an extreme point of A_1 we see that (x_1', x_2') is a convex combination of $\{(y_1, y_j) : 1 < j \leq m\}$ showing that $x_2' \in \operatorname{conv}\{y_2, ..., y_m\}$. Hence, the proof is complete. □

Recall the mapping S defined by (9). For any $x \in \operatorname{dom} S$ and any $\alpha \in S(x)$, we define the *support of* α by

$$\operatorname{supp}(\alpha) := \{j : \alpha_j \neq 0\}.$$

Lemma 4.4. *If $(x_1, x_2) \in L^2_{\prec}(y)$ is an extreme point then, for any $\alpha_i \in S(x_i)$, $i = 1, 2$, we have*

$$\operatorname{card}(\operatorname{supp}(\alpha_1) \cap \operatorname{supp}(\alpha_2)) \leq 1.$$

Proof. Suppose $\operatorname{card}(\operatorname{supp}(\alpha_1) \cap \operatorname{supp}(\alpha_2)) \geq 2$ for some $\alpha_i \in S(x_i)$, $i = 1, 2$. Without loss of generality assume that $\{1, 2\} \subset \operatorname{supp}(\alpha_1) \cap \operatorname{supp}(\alpha_2)$, that is, the elements in the set $\{\alpha_{ij} : i, j = 1, 2\}$ are strictly positive and strictly less than one. (If one of these numbers is equal to one then, $\operatorname{card}(\operatorname{supp}(\alpha_1) \cap \operatorname{supp}(\alpha_2)) \leq 1$.) For a real number ϵ define

$$x_1' := (\alpha_{11} + \epsilon) y_1 + (\alpha_{12} - \epsilon) y_2 + \sum_{j=3}^{m} \alpha_{1j} y_j,$$

$$x_2' := (\alpha_{21} - \epsilon) y_1 + (\alpha_{22} + \epsilon) y_2 + \sum_{j=3}^{m} \alpha_{2j} y_j,$$

and

$$x_1'' := (\alpha_{11} - \epsilon) y_1 + (\alpha_{12} + \epsilon) y_2 + \sum_{j=3}^{m} \alpha_{1j} y_j,$$

$$x_2'' := (\alpha_{21} + \epsilon) y_1 + (\alpha_{22} - \epsilon) y_2 + \sum_{j=3}^{m} \alpha_{2j} y_j.$$

Clearly, for ϵ close enough to zero we have that $(x_1', x_2'), (x_1'', x_2'') \in L^2_{\prec}(y)$ and $(x_1, x_2) = \frac{1}{2}((x_1', x_2') + (x_1'', x_2''))$, contradicting the assumption that (x_1, x_2) is an extreme point. □

Corollary 4.5. *If $(x_1, x_2) \in L^2_{\prec}(y)$ is an extreme point then, for any $\alpha_i \in S(x_i)$, $i = 1, 2$, we have*

$$\operatorname{card}(\operatorname{supp}(\alpha_1) \setminus \operatorname{supp}(\alpha_2)) \geq 1 \quad \textit{and} \quad \operatorname{card}(\operatorname{supp}(\alpha_2) \setminus \operatorname{supp}(\alpha_1)) \geq 1.$$

Proof. It is enough to show only the first inequality. Suppose on the contrary that $\operatorname{supp}(\alpha_1) \subseteq \operatorname{supp}(\alpha_2)$ for some $\alpha_i \in S(x_i)$, $i = 1, 2$. By Lemma 4.4 we have $\operatorname{card}(\operatorname{supp}(\alpha_1)) = 1$ and thus, $x_1 \in y$. Then, by Lemma 4.2, (x_1, x_2) is a convex combination of $\{(y_i, y_j) : 1 \leq i \neq j \leq m\}$. Since (x_1, x_2) is an extreme point of $L^2_{\prec}(y)$, we see, using Theorem 4.1, that $(x_1, x_2) \in \{(y_i, y_j) : 1 \leq i \neq j \leq m\}$, a contradiction. □

A simple observation using Corollary 1.6 and Theorem 4.1 yields the following lemma.

Lemma 4.6. *If $(x_1, x_2) \in L^2_{\prec}(y)$ and if for some $\alpha_i \in S(x_i)$, $i = 1, 2$, we have*

$$\alpha_{1j} + \alpha_{2j} \leq 1 \quad \textit{for all } j \in \mathbb{N}_m$$

then, (x_1, x_2) is in the convex hull of $\{(y_i, y_j) : 1 \leq i \neq j \leq m\}$.

Thus, to find extreme points of $L^2_{\prec}(y)$ that are not in (10), it needs to be assumed that (x_1, x_2) does not have the property described in Lemma 4.6.

Assumption 4.7. Assume $(x_1, x_2) \in L^2_{\prec}(y)$ is such that for any $\alpha_i \in S(x_i)$, $i = 1, 2$, we have

$$\alpha_{1j_0} + \alpha_{2j_0} > 1 \ \text{ for some } j_0 \in \mathbb{N}_m. \tag{12}$$

For any fixed choice of $\alpha_i \in S(x_i)$, $i = 1, 2$, there can be only one index j_0 for which (12) holds and (without further assumptions) it depends on the particular choice of the vectors $\alpha_i \in S(x_i)$, $i = 1, 2$. As we will see in Lemma 4.9, this is no longer the case if we assume that (x_1, x_2) is an extreme point of $L^2_{\prec}(y)$.

For a set $C \subset \mathbb{R}^n$, denote by $\operatorname{aff} C$ the *affine space* spanned by C:

$$\operatorname{aff} C = \left\{ x \in \mathbb{R}^n : x = \sum_{j=1}^{k} \alpha_j c_j, \sum_{j=1}^{k} \alpha_j = 1, c_j \in C, j \in \mathbb{N}_k, k = 1, 2, ... \right\}.$$

For a convex set $C \in \mathbb{R}^n$ denote by $\operatorname{ri} C$ the *relative interior* of C:

$$\operatorname{ri} C = \{x \in C : \exists \epsilon > 0 \text{ such that } (x + \epsilon \mathcal{B}_n) \cap \operatorname{aff} C \subset C\}.$$

Note that if $C = \{x\}$ is a singleton then, $\operatorname{aff} C = C$, $\operatorname{ri} C = C$, and $\dim \operatorname{aff} C = 0$.

Lemma 4.8. *Let $c_1, ..., c_k \in \mathbb{R}^n$ and suppose that each c_i is an extreme point of* $\operatorname{conv}\{c_i : i \in \mathbb{N}_k\}$. *Then*

$$\operatorname{ri}\operatorname{conv}\{c_i : i \in \mathbb{N}_k\} = \left\{ \sum_{i=1}^{k} \alpha_i c_i : \sum_{i=1}^{k} \alpha_i = 1, \alpha_i \in (0, 1), i \in \mathbb{N}_k \right\}.$$

Proof. Let $L := \operatorname{aff}\{c_i : i \in \mathbb{N}_k\}$, $d := \dim L$, $C := \operatorname{conv}\{c_i : i \in \mathbb{N}_k\}$ and let the right-hand side set, displayed in the lemma, be R. It is not difficult to see that $\operatorname{ri} C \supseteq R$ and that C is equal to the closure of R. We need to show the opposite inclusion. Suppose there is a point $x \in (\operatorname{ri} C) \setminus R$. Since R is a convex set, there is an affine subspace M of L, of dimension $d - 1$ such that $x \in M$ and R is on one side of M. Since $x \in \operatorname{ri} C$, there is a point $z \in C$ that is strictly on the other side of M. Thus, $z \notin \operatorname{cl} R$, a contradiction. □

A nonempty set F is an *extreme face* of A_1 if there is a vector $a \in \mathbb{R}^n$ and a real number α such that $a^T y \leq \alpha$ for all $y \in A_1$ and $F = A_1 \cap \{x \in \mathbb{R}^n : a^T x = \alpha\}$. Furthermore, F is a *proper extreme face* if $F \neq A_1$ and F is not an extreme point of A_1. For every $x \in A_1$ there is a unique (not necessarily proper) extreme face F of A_1, such that $x \in \operatorname{ri} F$. Moreover, for any $\alpha \in S(x)$, the set $\{y_j : j \in \operatorname{supp}(\alpha)\}$ is a subset of the set of extreme points of F. In other words, x can not be expressed as a convex combination of the extreme points of A_1 with a positive weight given to any vector that is not an extreme point of F.

Let F_1 (resp. F_2) be the extreme face of A_1 containing x_1 (resp. x_2) in its relative interior.

Lemma 4.9. *Let (x_1, x_2) be an extreme point of $L^2_{\prec}(y)$ for which Assumption 4.7 holds. Then*

(*1*) F_1 *and* F_2 *are proper extreme faces of* A_1*;*
(*2*) $F_1 \cap F_2 = \{y_{j_0}\}$ *for some* $j_0 \in \mathbb{N}_m$*;*
(*3*) $\operatorname{supp}(\alpha_1) \cap \operatorname{supp}(\alpha_2) = \{j_0\}$ *for any* $\alpha_i \in S(x_i)$, $i = 1, 2$.

Proof. (*1*). By Lemma 4.8, there is $\alpha_i \in S(x_i)$ such that $\{y_j : j \in \operatorname{supp}(\alpha_i)\}$ is precisely the set of extreme points of F_i, $i = 1, 2$. In addition, by Lemma 4.4, we have $\operatorname{card}(\operatorname{supp}(\alpha_1) \cap \operatorname{supp}(\alpha_2)) \le 1$. Since Assumption 4.7 holds, there is an index $j_0 \in \mathbb{N}_m$ such that $\alpha_{1j_0} + \alpha_{2j_0} > 1$, that is $\operatorname{supp}(\alpha_1) \cap \operatorname{supp}(\alpha_2) = \{j_0\}$. Corollary 4.5 now shows that neither F_1 nor F_2 is an extreme point and that both F_1 and F_2 are proper extreme faces of A_1.

(*2*). Since the intersection of the two extreme faces F_1 and F_2 is an extreme face, if $F_1 \cap F_2$ contains more than one point, it has to contain more than one extreme point, contradicting the fact that $\operatorname{card}(\operatorname{supp}(\alpha_1) \cap \operatorname{supp}(\alpha_2)) \le 1$.

(*3*). Now, take any $\alpha'_i \in S(x_i)$, $i = 1, 2$. By Assumption 4.7, there is an index $j' \in \mathbb{N}_m$ such that $\alpha'_{1j'} + \alpha'_{2j'} > 1$, that is $\{j'\} = \operatorname{supp}(\alpha'_1) \cap \operatorname{supp}(\alpha'_2)$. Since $\{y_j : j \in \operatorname{supp}(\alpha'_i)\}$ is a subset of the extreme points of F_i we obtain $y_{j'} \in F_1 \cap F_2$. That is, $j' = j_0$, since the vectors in y are distinct. □

For a set $F \subset \mathbb{R}^n$, denote by $\operatorname{lin} F$ the unique linear space parallel to $\operatorname{aff} F$:

$$\operatorname{lin} F := \operatorname{aff} F - a \quad \text{for any } a \in \operatorname{aff} F.$$

Lemma 4.10. *Let (x_1, x_2) be an extreme point of $L^2_{\prec}(y)$ for which Assumption 4.7 holds. Then*

$$\operatorname{lin} F_1 \cap \operatorname{lin} F_2 = \{0\}. \tag{13}$$

Proof. By Lemma 4.8, there is $\alpha_i \in S(x_i)$ such that $\{y_j : j \in \operatorname{supp}(\alpha_i)\}$ is precisely the set of extreme points of F_i, $i = 1, 2$. Suppose on the contrary that $v \neq 0$ belongs to the set on the left-hand side of (13). In particular that means

$$v = \sum_{j \in \operatorname{supp}(\alpha_i)} \beta_{ij} y_j \quad \text{for } i = 1, 2,$$

for some coefficients with $\sum_{j \in \operatorname{supp}(\alpha_i)} \beta_{ij} = 0$ where $i = 1, 2$. For a number ϵ define

$$x'_1 := \sum_{j \in \operatorname{supp}(\alpha_1)} \alpha_{1j} y_j + \epsilon v \quad \text{and} \quad x'_2 := \sum_{j \in \operatorname{supp}(\alpha_2)} \alpha_{2j} y_j - \epsilon v,$$

together with

$$x_1'' := \sum_{j \in \operatorname{supp}(\alpha_1)} \alpha_{1j} y_j - \epsilon v \quad \text{and} \quad x_2'' := \sum_{j \in \operatorname{supp}(\alpha_2)} \alpha_{2j} y_j + \epsilon v.$$

It is clear that for all ϵ close enough to zero we have $x_1', x_2', x_1'', x_2'' \in A_1$ and that $x_1' + x_2' = x_1'' + x_2'' = x_1 + x_2 \in A_2$. Then, we have $(x_1, x_2) = \frac{1}{2}((x_1', x_2') + (x_1'', x_2''))$, which contradicts the assumption that (x_1, x_2) is an extreme point. □

Assume that (x_1, x_2) is an extreme point of $L^2_{\prec}(y)$ for which Assumption 4.7 holds. Define

$$z := x_1 + x_2 - y_{j_0} \tag{14}$$

and note that by Corollary 4.3 we have $z \in \operatorname{conv}\{y_1, ..., \hat{y}_{j_0}, ..., y_m\}$, where the hat indicates that vector y_{j_0} is omitted from the list. Let

$$A_1' := \operatorname{conv}\{y_1, ..., \hat{y}_{j_0}, ..., y_m\}.$$

Let F_3 be the unique extreme face of A_1' that contains z in its relative interior.

Lemma 4.11. *Let (x_1, x_2) be an extreme point of $L^2_{\prec}(y)$ for which Assumption 4.7 holds. Then*

$$\operatorname{lin} F_3 \cap \operatorname{lin}(F_1 \cup F_2) = \{0\}. \tag{15}$$

Proof. By Lemma 4.8, there is $\alpha_i \in S(x_i)$ such that $\{y_j : j \in \operatorname{supp}(\alpha_i)\}$ is precisely the set of extreme points of F_i, $i = 1, 2$. Note that $\operatorname{supp}(\alpha_1) \cap \operatorname{supp}(\alpha_2) = \{j_0\}$. Suppose on the contrary that $v \neq 0$ belongs to the left-hand side of (15). In particular, this means that

$$v = \beta_{j_0} y_{j_0} + \sum_{\substack{j \in \operatorname{supp}(\alpha_1) \\ j \neq j_0}} \beta_j y_j + \sum_{\substack{j \in \operatorname{supp}(\alpha_2) \\ j \neq j_0}} \beta_j y_j,$$

for some coefficients with $\beta_{j_0} + \sum_{\substack{j \in \operatorname{supp}(\alpha_1) \\ j \neq j_0}} \beta_j + \sum_{\substack{j \in \operatorname{supp}(\alpha_2) \\ j \neq j_0}} \beta_j = 0$. Define the vectors

$$v_1 := \beta_{j_0} y_{j_0} + \sum_{\substack{j \in \operatorname{supp}(\alpha_1) \\ j \neq j_0}} \beta_j y_j \quad \text{and} \quad v_2 := \sum_{\substack{j \in \operatorname{supp}(\alpha_2) \\ j \neq j_0}} \beta_j y_j$$

and the scalars

$$t_1 := \beta_{j_0} + \sum_{\substack{j \in \operatorname{supp}(\alpha_1) \\ j \neq j_0}} \beta_j \alpha_i \quad \text{and} \quad t_2 := \sum_{\substack{j \in \operatorname{supp}(\alpha_2) \\ j \neq j_0}} \beta_j.$$

Clearly, we have $t_1 + t_2 = 0$ and $v_1 + v_2 = v$. For a number ϵ define

$$x_1' = (\alpha_{1j_0} - \epsilon t_1) y_{j_0} + \sum_{\substack{j \in \operatorname{supp}(\alpha_1) \\ j \neq j_0}} \alpha_{1j} y_j + \epsilon v_1,$$

$$x_2' = (\alpha_{2j_0} - \epsilon t_2) y_{j_0} + \sum_{\substack{j \in \operatorname{supp}(\alpha_2) \\ j \neq j_0}} \alpha_{2j} y_j + \epsilon v_2,$$

and

$$x_1'' = (\alpha_{1j_0} + \epsilon t_1) y_{j_0} + \sum_{\substack{j \in \operatorname{supp}(\alpha_1) \\ j \neq j_0}} \alpha_{1j} y_j - \epsilon v_1,$$

$$x_2'' = (\alpha_{2j_0} + \epsilon t_2) y_{j_0} + \sum_{\substack{j \in \operatorname{supp}(\alpha_2) \\ j \neq j_0}} \alpha_{2j} y_j - \epsilon v_2.$$

For all ϵ close enough to zero, we have $x_1', x_2' \in A_1$ and

$$x_1' + x_2' = x_1 + x_2 + \epsilon(v_1 + v_2) = z + y_{j_0} + \epsilon v \in y_{j_0} + F_3 \subset A_2,$$

since z is in the relative interior of F_3 and $v \in \operatorname{lin} F_3$. This shows that $(x_1', x_2') \in L^2_\prec(y)$ and similarly $(x_1'', x_2'') \in L^2_\prec(y)$. Finally,

$$(x_1, x_2) = \frac{1}{2}\left((x_1', x_2') + (x_1'', x_2'')\right)$$

contradicts the fact that (x_1, x_2) is an extreme point. □

Lemma 4.12. *Let (x_1, x_2) be an extreme point of $L^2_\prec(y)$ for which Assumption* 4.7 *holds. Then $z \notin \{y_1, ..., \hat{y}_{j_0}, ..., y_m\}$.*

Proof. Fix any $\alpha_i \in S(x_i)$ and note that by Lemma 4.9, part (*3*), and by Assumption 4.7 we have $\alpha_{1j_0} + \alpha_{2j_0} > 1$. Suppose on the contrary that $z = y_{j'}$ (recall (14)) for some $j' \in \mathbb{N}_m \setminus \{j_0\}$. If $j' \in \operatorname{supp}(\alpha_1)$ then, by (14), we obtain

$$y_{j'} = z = x_1 + x_2 - y_1 = (\alpha_{1j_0} + \alpha_{2j_0} - 1) y_{j_0} + \sum_{\substack{j \in \operatorname{supp}(\alpha_1) \\ j \neq j_0}} \alpha_{1j} y_j + \sum_{\substack{j \in \operatorname{supp}(\alpha_2) \\ j \neq j_0}} \alpha_{2j} y_j.$$

Solving for $y_{j'}$ gives

$$y_{j'} = \frac{\alpha_{1j_0} + \alpha_{2j_0} - 1}{1 - \alpha_{1j'}} y_{j_0} + \sum_{\substack{j \in \operatorname{supp}(\alpha_1) \\ j \neq j_0, j'}} \frac{\alpha_{1j}}{1 - \alpha_{1j'}} y_j + \sum_{\substack{j \in \operatorname{supp}(\alpha_2) \\ j \neq j_0}} \frac{\alpha_{2j}}{1 - \alpha_{1j'}} y_j.$$

The right-hand side is a convex combination not involving $y_{j'}$, contradicting the fact that $y_{j'}$ is an extreme point. The case $j' \in \operatorname{supp}(\alpha_2)$ is analogous, while the case $j' \notin \operatorname{supp}(\alpha_1) \cup \operatorname{supp}(\alpha_2)$ is elementary. □

Corollary 4.13. *Assume that (x_1, x_2) is an extreme point of $L^2_\prec(y)$ for which Assumption* 4.7 *holds. Then*

(*1*) *F_3 cannot intersect $F_1 \cup F_2$ at more than one point. If such a point exists, it is necessarily different from y_{j_0};*

(*2*) *F_3 is a proper extreme face of A_1', that is $z \notin \operatorname{ri} A_1'$.*

Definition 4.14. We say that the quadruple (y_{j_0}, F_1, F_2, F_3), consisting of a vector y_{j_0} from y, two proper extreme faces F_1, F_2 of the polytope A_1, and a proper extreme face F_3 of the polytope A_1', has the property $\mathcal{P}$ if the following conditions are satisfied:

(1) $F_1 \cap F_2 = \{y_{j_0}\}$;
(2) $\operatorname{lin} F_1 \cap \operatorname{lin} F_2 = \{0\}$ and $\operatorname{lin}(F_1 \cup F_2) \cap \operatorname{lin} F_3 = \{0\}$;
(3) Defining the set

$$B := \{(x_1, x_2) : x_i \in \operatorname{ri} F_i \text{ and } \alpha_{1j_0} + \alpha_{2j_0} > 1 \text{ for all } \alpha_i \in S(x_i), i = 1, 2\}$$

we have that $\{x_1 + x_2 - y_{j_0} : (x_1, x_2) \in B\} \cap \operatorname{ri} F_3 \neq \emptyset$.

Lemma 4.15. *If the quadruple (y_{j_0}, F_1, F_2, F_3) has the property $\mathcal{P}$ then, there is precisely one pair $(x_1, x_2) \in B$ such that $x_1 + x_2 - y_{j_0} \in F_3$.*

Proof. Suppose on the contrary, that there are two such pairs (x_1, x_2) and (x_1', x_2'). That is, $u := x_1 + x_2 - y_{j_0} \in F_3$ and $u' := x_1' + x_2' - y_{j_0} \in F_3$. Then, $u - u' = (x_1 - x_1') + (x_2 - x_2') \in \operatorname{lin} F_3 \cap \operatorname{lin}(F_1 \cup F_2)$, implies that $u = u'$. Thus, $x_1 - x_1' = x_2' - x_2 \in \operatorname{lin} F_1 \cap \operatorname{lin} F_2$, showing that $(x_1, x_2) = (x_1', x_2')$. □

Definition 4.16. We say that the pair of vectors (x_1, x_2) in $\mathbb{R}^n$ has the property $\mathcal{P}$ if there is a quadruple (y_{j_0}, F_1, F_2, F_3) with the property $\mathcal{P}$, such that $(x_1, x_2) \in B$ and $x_1 + x_2 - y_{j_0} \in F_3$.

It is easy to see from the definition that, if a pair of vectors (x_1, x_2) in $\mathbb{R}^n$ has the property $\mathcal{P}$ then, $(x_1, x_2) \in L^2_{\prec}(y)$, but $(x_1, x_2) \notin L^2_{\prec_{\mathrm{ds}}}(y)$. We now state the main result of the section.

Theorem 4.17 (The extreme points of $L^2_{\prec}(y)$). *The pair of vectors (x_1, x_2) in $\mathbb{R}^n$ is an extreme point of $L^2_{\prec}(y)$ if and only if $(x_1, x_2) = (y_i, y_j)$ for some $1 \leq i \neq j \leq m$ or has the property $\mathcal{P}$.*

Proof. We only need to prove the sufficiency of the property $\mathcal{P}$ for (x_1, x_2) to be an extreme point. Thus, let $(x_1, x_2) \in L^2_{\prec}(y)$ have the property $\mathcal{P}$ with corresponding quadruple (y_{j_0}, F_1, F_2, F_3). Suppose that

$$(x_1, x_2) = t(x_1', x_2') + (1 - t)(x_1'', x_2'') \tag{16}$$

for some distinct $(x_1', x_2'), (x_1'', x_2'') \in L^2_{\prec}(y)$ and $t \in (0, 1)$. Since the whole segment $[(x_1', x_2'), (x_1'', x_2'')]$ belongs to $L^2_{\prec}(y)$ we can assume that (x_1', x_2') and (x_1'', x_2'') are arbitrary close to (x_1, x_2). By applying Corollary 3.2 four times, there are vectors $\alpha_i^k \in S(x_i)$ for $i, k = 1, 2$ and vectors $\alpha_i' \in S(x_i')$, $\alpha_i'' \in S(x_i'')$ for $i = 1, 2$ such that α_i^1 is arbitrarily close to α_i' and α_i^2 is arbitrarily close to α_i'', for $i = 1, 2$. Using the fact that $(x_1, x_2) \in B$, defined in part (3) of Definition 4.14, we obtain

$$\alpha_{1j_0}' + \alpha_{2j_0}' > 1 \quad \text{and} \quad \alpha_{1j_0}'' + \alpha_{2j_0}'' > 1.$$

Corollary 4.3 now implies that $x_1' + x_2' - y_{j_0} \in A_1'$. Since F_3 is an extreme face of A_1', $x_1 + x_2 - y_{j_0} \in \operatorname{ri} F_3$, and since (x_1', x_2') is arbitrarily close to (x_1, x_2), we conclude that $x_1' + x_2' - y_{j_0} \in \operatorname{ri} F_3$. Analogously, we obtain $x_1'' + x_2'' - y_{j_0} \in \operatorname{ri} F_3$. Subtracting the last two inclusions, we obtain

$$x_1' + x_2' - (x_1'' + x_2'') \in \operatorname{lin} F_3.$$

The last difference can also be written as

$$(x_1' - x_1'') + (x_2' - x_2'') \in \operatorname{lin}(F_1 \cup F_2).$$

By the fact that $\operatorname{lin} F_3 \cap \operatorname{lin}(F_1 \cup F_2) = \{0\}$, we conclude that

$$x_1' + x_2' = x_1'' + x_2'' = x_1 + x_2,$$

where the last equality follows from (16). Therefore, we have $x_1 - x_1' = x_2 - x_2' \in \operatorname{lin} F_1 \cap \operatorname{lin} F_2$. This, by the fact that $\operatorname{lin} F_1 \cap \operatorname{lin} F_2 = \{0\}$, implies $x_1 = x_1'$ and $x_2 = x_2'$. After analogous argument for (x_1'', x_2'') we conclude

$$(x_1, x_2) = (x_1', x_2') = (x_1'', x_2'').$$

This contradicts the fact that the points (x_1', x_2') and (x_1'', x_2'') are distinct and proves the result. □

Example 4.18. Consider the vectors $y = (y_1, ..., y_5)$ in R^3, where $y_1 = (0,0,0)^T$, $y_2 = (0,-1,2)^T$, $y_3 = (0,1,2)^T$, $y_4 = (1,0,1)^T$, and $y_5 = (-1,0,1)^T$. Let $x_1 = (0,-1/4,2/4)^T$ and $x_2 = (0,1/4,2/4)^T$. Then, we have $x_1 = (3/4)y_1 + (1/4)y_2$ and $x_2 = (3/4)y_1 + (1/4)y_3$, showing that $x_1, x_2 \in A_1$ and $x_1 + x_2 = y_1 + (1/2)y_4 + (1/2)y_5$ showing that $(x_1, x_2) \in L^2_{\prec}(y)$. On the other hand (x_1, x_2) has the property $\mathcal{P}$ with quadruple (y_1, F_1, F_2, F_3), where $F_1 = \operatorname{conv}\{y_1, y_2\}$, $F_2 = \operatorname{conv}\{y_1, y_3\}$, and $F_3 = \operatorname{conv}\{y_4, y_5\}$. Thus, (x_1, x_2) is an extreme point of $L^2_{\prec}(y)$ not in the set $\{(y_i, y_j) : 1 \le i \ne j \le 4\}$.

It is possible to prove the next corollary directly, without relying on Theorem 4.17, but now it comes as a simple exercise in plane geometry.

Corollary 4.19. *Suppose that vectors $y = (y_k)_{k=1}^m$ belong to a two dimensional affine subspace of $\mathbb{R}^n$. The extreme point of $L^2_{\prec}(y)$ are $\{(y_i, y_j) : 1 \le i \ne j \le m\}$, that is, there are no extreme points with the property $\mathcal{P}$.*

Proof. It is not difficult to see that there are no quadruples (y_{j_0}, F_1, F_2, F_3) with the property $\mathcal{P}$. □

The next corollary recovers an observation made by Malamud in [5] and mentioned in the introduction.

Corollary 4.20. *Suppose that vectors $y = (y_k)_{k=1}^m$ in $\mathbb{R}^n$ are affinely independent. The extreme point of $L^2_{\prec}(y)$ are $\{(y_i, y_j) : 1 \le i \ne j \le m\}$, that is, there are no extreme points with the property $\mathcal{P}$.*

Definition 4.21. We say that the polytope $A_1 \subset \mathbb{R}^n$ with extreme points $y = (y_k)_{k=1}^m$ has the property $\mathcal{P}$ if there is a quadruple (y_{j_0}, F_1, F_2, F_3) having the property $\mathcal{P}$.

The next theorem is the answer to Question 1. We repeat the standing assumptions that we imposed on the vectors in y throughout this subsection.

Theorem 4.22 (Answer to Question 1). *Let $y = (y_k)_{k=1}^m$ be a sequence of distinct vectors in $\mathbb{R}^n$ such that every y_k is an extreme point of the polytope $A_1(y)$. Then, for any sequence $x = (x_k)_{k=1}^{\ell}$ we have ($x \prec y \Leftrightarrow x \prec_{\mathrm{ds}} y$) if and only if the polytope $A_1(y)$ does not have the property $\mathcal{P}$.*

Proof. If the polytope $A_1(y)$ does not have the property $\mathcal{P}$ then, the extreme points of $L^2_{\prec}(y)$ are the same as the extreme points of $L^2_{\prec_{\rm ds}}(y)$. In other words $(x_1, x_2) \prec y$ is equivalent to $(x_1, x_2) \prec_{\rm ds} y$ for any $x_1, x_2 \in \mathbb{R}^n$. By Theorem 1.5 together with Lemma 1.3 we conclude.

Suppose that $A_1(y)$ has the property $\mathcal{P}$, that is, there is a quadruple (y_{j_0}, F_1, F_2, F_3) having the property $\mathcal{P}$. Let $(x_1, x_2) \in B$ be such that $x_1 + x_2 - y_{j_0} \in F_3$, that is $x_1 + x_2 \in A_2$ showing that $(x_1, x_2) \prec y$. Since $\alpha_{1j_0} + \alpha_{2j_0} > 1$ for all $\alpha_i \in S(x_i)$, $i = 1, 2$ we see that $(x_1, x_2) \prec_{\rm ds} y$ does not hold. □

Corollary 4.23 (Answer to Question 3). *Let $y = (y_k)_{k=1}^m$ be a sequence of distinct vectors in $\mathbb{R}^n$ such that every y_k is an extreme point of the polytope $A_1(y)$. Then, for $\ell \le m$ we have $L^\ell_{\prec}(y) = L^\ell_{\prec_{\rm ds}}(y)$ if and only if the polytope $A_1(y)$ does not have the property $\mathcal{P}$.*

Corollary 4.24. *Let $y = (y_k)_{k=1}^m$ be a sequence of distinct vectors in $\mathbb{R}^n$ such that every y_k is an extreme point of the polytope $A_1(y)$. Suppose that the vectors in y belong to a two dimensional affine subspace of $\mathbb{R}^n$. Then, for any sequence $x = (x_k)_{k=1}^\ell$, $\ell \le m$, we have $x \prec y \Leftrightarrow x \prec_{\rm ds} y$.*

Corollary 4.25. *Let $y = (y_k)_{k=1}^m$ be a sequence of affinely independent vectors in $\mathbb{R}^n$. Then, for any sequence $x = (x_k)_{k=1}^\ell$, $\ell \le m$, we have $x \prec y \Leftrightarrow x \prec_{\rm ds} y$.*

We conclude this section with a result providing a geometric description of a wide family of convex functions $f : \mathbb{R}^n \to \mathbb{R}$ that are not in the class $\mathrm{CVS}\,(\mathbb{R}^n)$.

Theorem 4.26. *Let $y = (y_k)_{k=1}^m$ be a sequence of distinct vectors in $\mathbb{R}^n$ such that every y_k is an extreme point of the polytope $A_1(y)$. Suppose the polytope $A_1(y)$ has the property $\mathcal{P}$, that is, there is a quadruple (y_{j_0}, F_1, F_2, F_3) having the property $\mathcal{P}$. Then, the function*

$$f(w) := \mathrm{dist}\,(w, A'_1(y))$$

is convex on $\mathbb{R}^n$ but is not in the class $\mathrm{CVS}\,(\mathbb{R}^n)$, where

$$A'_1(y) = \mathrm{conv}\,\{y_1, ..., \hat{y}_{j_0}, ..., y_m\}.$$

Proof. Let B be the set given in Definition 4.14 and let (x_1, x_2) be the unique pair in B (see Lemma 4.15) such that $z := x_1 + x_2 - y_{j_0} \in \mathrm{ri}\, F_3$. By Definition 4.14, we have $(x_1, x_2) \in L^2_{\prec}(y)$, but $(x_1, x_2) \notin L^2_{\prec_{\rm ds}}(y)$. Thus, according to Theorem 2.1 and Theorem 2.2, in order to show that $f \notin \mathrm{CVS}\,(\mathbb{R}^n)$, it is sufficient to see that Inequality (4) is violated, that is

$$\mathrm{dist}\,(x_1, A'_1(y)) + \mathrm{dist}\,(x_2, A'_1(y)) > \mathrm{dist}\,(y_{j_0}, A'_1(y)). \tag{17}$$

Since $z \in F_3 \subset A'_1(y)$, we conclude that $\mathrm{dist}\,(y_{j_0}, A'_1(y)) \le \|z - y_{j_0}\|$, that is, the right-hand side of (17) is always at most $\|z - y_{j_0}\|$. Without loss of generality we may assume that

$$\mathrm{dist}\,(y_{j_0}, A'_1(y)) = \|z - y_{j_0}\|. \tag{18}$$

Since $z \in \operatorname{ri} F_3$, it is not difficult to see that $z - y_{j_0}$ is orthogonal to $\operatorname{aff} F_3$. Let H be the hyperplane in $\mathbb{R}^n$ with normal $z - y_{j_0}$ and containing $\operatorname{aff} F_3$, that is

$$H := \{x \in \mathbb{R}^n : \langle z - x, z - y_{j_0} \rangle = 0\}.$$

Define the closed and open half-space

$$\begin{aligned} H^+ &:= \{x \in \mathbb{R}^n : \langle z - x, z - y_{j_0} \rangle \geq 0\}, \quad \text{and} \\ H^{++} &:= \{x \in \mathbb{R}^n : \langle z - x, z - y_{j_0} \rangle > 0\}, \end{aligned}$$

and similarly H^- and H^{--}. Note that $z \in H$. Since $y_{j_0} \notin F_3$ we have $y_{j_0} \neq z$ and thus, $y_{j_0} \in H^{++}$.

Claim. *$A_1'(y) \subset H^-$.*

Proof. Assume that $z_1 \in A_1'(y)$ is different from z and $z_1 \in H^{++}$. Then, by the accessibility lemma, see [7, page 164], the segment $(z, z_1]$ is contained in $A_1'(y) \cap H^{++}$. Since $y_{j_0} + \|z - y_{j_0}\| \operatorname{int} \mathcal{B}_n \subset H^{++}$ and H is tangent to the closed ball $y_{j_0} + \|z - y_{j_0}\| \mathcal{B}_n$, we obtain that $(z, z_1] \cap (y_{j_0} + \|z - y_{j_0}\| \operatorname{int} \mathcal{B}_n) \neq \emptyset$. This contradicts the fact that $\operatorname{dist}(y_{j_0}, A_1'(y)) = \|z - y_{j_0}\|$. Thus, $A_1'(y) \subset H^-$. □

Claim. *We have that $x_1, x_2 \in H^{++}$.*

Proof. By the supporting hyperplane theorem, since x_1 and x_2 belong to the convex set $A_1(y)$ with extreme point y_{j_0}, we conclude that $\langle x_2 - y_{j_0}, x_1 - y_{j_0} \rangle \geq 0$. Thus,

$$\begin{aligned} \langle z - x_1, z - y_{j_0} \rangle &= \langle x_2 - y_{j_0}, x_1 + x_2 - 2y_{j_0} \rangle \\ &= \|x_2 - y_{j_0}\|^2 + \langle x_2 - y_{j_0}, x_1 - y_{j_0} \rangle \\ &> 0, \end{aligned}$$

since $x_2 \in \operatorname{ri} F_2$ and y_{j_0} is an extreme point of F_2. Analogously, we see that $\langle z - x_2, z - y_{j_0} \rangle \geq 0$. □

To summarize, we showed that $x_1, x_2, y_{j_0} \in H^{++}$ and $A_1'(y) \subset H^-$. This implies

$$\operatorname{dist}(x_i, H) \leq \operatorname{dist}(x_i, A_1'(y)), \quad i = 1, 2. \tag{19}$$

Next, since $x_i \in \operatorname{ri} F_i$ is a convex combination of y_{j_0} and a point from $A_1'(y)$, the ray $\{y_{j_0} + t(x_i - y_{j_0}) : t \geq 0\}$ intersects H in a unique point, say P_i, $i = 1, 2$. Moreover, since $x_1, x_2, y_{j_0} \in H^{++}$ we have $P_i = y_{j_0} + t_i(x_i - y_{j_0})$ for some $t_i > 1$, $i = 1, 2$.

Claim. *z belongs to the open segment between P_1 and P_2.*

Proof. By the fact that $P_i \in H$ we obtain

$$0 = \langle z - P_i, z - y_{j_0} \rangle = \|z - y_{j_0}\|^2 - t_i \langle x_i - y_{j_0}, z - y_{j_0} \rangle$$

or equivalently

$$\frac{1}{t_i} = \frac{\langle x_i - y_{j_0}, z - y_{j_0} \rangle}{\|z - y_{j_0}\|^2}, \quad i = 1, 2.$$

By the definition of z and the observation $1/t_1 + 1/t_2 = 1$, we obtain

$$
\begin{aligned}
z = (x_1 - y_{j_0}) + (x_2 - y_{j_0}) + y_{j_0} &= \frac{P_1 - y_{j_0}}{t_1} + \frac{P_2 - y_{j_0}}{t_2} + y_{j_0} \\
&= \frac{1}{t_1}P_1 + \frac{1}{t_2}P_2.
\end{aligned}
$$

The fact that $1/t_i \in (0,1)$, $i = 1, 2$, completes the proof of the claim. $\square$

We now focus our attention on the triangle with vertices y_{j_0}, P_1, and P_2. We have shown that x_i belongs to the open segment between y_{j_0} and P_i, $i = 1, 2$ and that z belongs to the open segment between P_1 and P_2. Let z_i be the projection of x_i onto H, that is

$$
\operatorname{dist}(x_i, H) = \|z_i - x_i\|, \quad i = 1, 2. \tag{20}
$$

Since vector $z - y_{j_0}$ is orthogonal to H we conclude that z_i belongs to the open segment between P_i and z, $i = 1, 2$. Thus, z is a convex combination of z_1 and z_2. Observe that $z_1 \neq z_2$ or else $z_1 = z_2 = z$ and x_1, x_2 must be on the open segment between y_{j_0} and z, showing that the extreme faces F_1 and F_2 have more than one point in common contradicting Definition 4.14. Finally, since the triangle with vertices y_{j_0}, P_1, P_2 is in $\operatorname{lin}(F_1 \cup F_2)$, we conclude that

$$
z_1 - z_2 \in \operatorname{lin}(F_1 \cup F_2). \tag{21}
$$

Claim. *We have*

$$
\operatorname{dist}(x_1, H) + \operatorname{dist}(x_2, H) < \operatorname{dist}(x_1, A_1'(y)) + \operatorname{dist}(x_2, A_1'(y)). \tag{22}
$$

Proof. Suppose on the contrary that the claim is not true, that is by (19) we must have $\operatorname{dist}(x_i, H) = \operatorname{dist}(x_i, z_i) = \operatorname{dist}(x_i, A_1'(y))$ implying that $z_i \in A_1'(y)$, $i = 1, 2$. Since $z \in \operatorname{ri} F_3$ and F_3 is an extreme face of $A_1'(y)$, we see that $z_1, z_2 \in F_3$ and thus,

$$
z_1 - z_2 \in \operatorname{lin} F_3.
$$

This, together with (21) contradicts the fact that $\operatorname{lin}(F_1 \cup F_2) \cap \operatorname{lin} F_3 = \{0\}$ in Definition 4.14. $\square$

Inequality (22) together with (18) and (20) shows that in order to prove (17) it is enough to see that

$$
\|z_1 - x_1\| + \|z_2 - x_2\| = \|z - y_{j_0}\|.
$$

Indeed, the triangle with vertices x_1, z_1, z is similar to the triangle with vertices y_{j_0}, z, P_2, implying that

$$
\frac{\|z_1 - x_1\|}{\|z - y_{j_0}\|} = \frac{\|z - x_1\|}{\|P_2 - y_{j_0}\|} = \frac{\|x_2 - y_{j_0}\|}{\|P_2 - y_{j_0}\|}, \tag{23}
$$

where we used that $z - x_1 = x_2 - y_{j_0}$. Analogously, since the triangle with vertices x_2, z_2, P_2 is similar to the triangle with vertices y_{j_0}, z, P_2, we obtain

$$\frac{\|z_2 - x_2\|}{\|z - y_{j_0}\|} = \frac{\|P_2 - x_2\|}{\|P_2 - y_{j_0}\|}. \tag{24}$$

Combining (23) and (24) we finally obtain

$$\begin{aligned}\|z_1 - x_1\| + \|z_2 - x_2\| &= \|z - y_{j_0}\| \frac{\|x_2 - y_{j_0}\|}{\|P_2 - y_{j_0}\|} + \|z - y_{j_0}\| \frac{\|P_2 - x_2\|}{\|P_2 - y_{j_0}\|} \\ &= \|z - y_{j_0}\|,\end{aligned}$$

where we used that x_2 is on the segment between y_{j_0} and P_2. $\square$

5. The extreme points of $L^\ell_\prec(y)$

Let $x = (x_k)_{k=1}^\ell$ and $y = (y_k)_{k=1}^m$ be two sequences of vectors in $\mathbb{R}^n$ with $\ell \le m$ and $x \in L^\ell_\prec(y)$. We do not impose any restrictions on the vectors y. In this section we derive an algebraic characterization of the extreme points of $L^\ell_\prec(y)$ in this general case. For any subset $\pi \subset \mathbb{N}_\ell$ denote

$$x_\pi := \sum_{k \in \pi} x_k.$$

According to Theorem 2.3, for every subset $\pi \subseteq \mathbb{N}_\ell$ there are numbers $\{\alpha_{\pi,j}\}_{j=1}^m$ in $[0,1]$ with $\sum_{j=1}^m \alpha_{\pi,j} = \operatorname{card}(\pi)$ and such that

$$x_\pi = \sum_{j=1}^m \alpha_{\pi,j} y_j.$$

The m-tuple $\alpha_\pi := (\alpha_{\pi,1}, ..., \alpha_{\pi,m})$ will be called a *Petrov representation of* x_π. If the set π contains just one element, say $\pi = \{k\}$ then, we just write x_k and $\alpha_k = (\alpha_{k,1}, ..., \alpha_{k,m})$. Slight care is required with this notation. While x_π denotes the sum of the vectors $\{x_k : k \in \pi\}$, the coefficient $\alpha_{\pi,j}$ is not necessarily the sum of $\{\alpha_{k,j} : k \in \pi\}$. Given a Petrov representation α_π, consider the set

$$\operatorname{supp}_*(\alpha_\pi) := \{j \in \mathbb{N}_m : \alpha_{\pi,j} \in (0,1)\}.$$

If α'_π and α''_π are two Petrov representations of x_π then, we say that $\alpha'_\pi \le \alpha''_\pi$ if $\operatorname{supp}_*(\alpha'_\pi) \subset \operatorname{supp}_*(\alpha''_\pi)$. This relation is reflexive and transitive and thus, for a fixed π, it defines a preorder on all Petrov representations of x_π. A Petrov representation α^*_π of x_π will be called *maximal* if there is no Petrov representation α_π of x_π with $\operatorname{supp}_*(\alpha^*_\pi) \subsetneqq \operatorname{supp}_*(\alpha_\pi)$. It is easy to see that for any Petrov representation α_π of x_π there is a maximal Petrov representation α^*_π with $\alpha_\pi \le \alpha^*_\pi$.

Lemma 5.1. *Let $\pi \subset \mathbb{N}_\ell$ with card$(\pi) = k$, and let α^*_π be a maximal Petrov representation of x_π. Let F be the unique extreme face of A_k such that $x_\pi \in \operatorname{ri} F$. Furthermore, let $\mathcal{S}$ be the set*

$$\mathcal{S} = \{\{i_1, ..., i_k\} : y_{i_1} + \cdots + y_{i_k} \text{ is an extreme point of } F\}.$$

Then, for any $j \in \mathbb{N}_m$ we have:

(*1*) *If* $\alpha^*_{\pi,j} = 1$ *then* $j \in s$ *for all* $s \in \mathcal{S}$*;*
(*2*) *If* $\alpha^*_{\pi,j} = 0$ *then* $j \notin s$ *for any* $s \in \mathcal{S}$*;*
(*3*) *If* $x \in \operatorname{ri} F$ *has a maximal Petrov representation* α^*_x *then,*

$$\{j \in \mathbb{N}_m : \alpha^*_{\pi,j} = i\} = \{j \in \mathbb{N}_m : \alpha^*_{x,j} = i\} \quad \textit{for } i = 0, 1.$$

Proof. (*1*). Since $x_\pi \in \operatorname{ri} F$, by Lemma 4.8 there are numbers $\{\gamma_s \in (0,1) : s \in \mathcal{S}\}$ with $\sum_{s\in\mathcal{S}} \gamma_s = 1$ such that

$$x_\pi = \sum_{s\in\mathcal{S}} \gamma_s \left(\sum_{j\in s} y_j\right) = \sum_{j=1}^{m} \left(\sum_{s\in\mathcal{S}: j\in s} \gamma_s\right) y_j. \tag{25}$$

Define a vector $\alpha \in \mathbb{R}^m$ by

$$\alpha_j := \sum_{s\in\mathcal{S}: j\in s} \gamma_s \quad \text{for } j \in \mathbb{N}_m.$$

It is easy to see that $\alpha_j \in [0,1]$ and $\sum_{j=1}^m \alpha_j = k$, since every γ_s is counted k times. Thus, α is a Petrov representation of x_π.

We now show the necessity. Suppose there is an index $j_0 \in \mathbb{N}_m$ such that $\alpha^*_{\pi,j_0} = 1$ but $j_0 \notin s_0$ for some $s_0 \in \mathcal{S}$, that is, $0 \le \alpha_{j_0} < 1$. Recalling that we also have $x_\pi = \sum_{j=1}^m \alpha^*_{\pi,j} y_j$ and combining it with (25), we find that for any real number t we have

$$x_\pi = \sum_{j=1}^{m} \left(\alpha^*_{\pi,j} + t(\alpha_j - \alpha^*_{\pi,j})\right) y_j. \tag{26}$$

Note that if $\alpha^*_{\pi,j} = 1$ then, $\alpha_j - \alpha^*_{\pi,j} \le 0$ and if $\alpha^*_{\pi,j} = 0$ then, $\alpha_j - \alpha^*_{\pi,j} \ge 0$. In addition, we have $\alpha_{j_0} - \alpha^*_{\pi,j_0} < 0$. Thus, for all positive values of t close enough to zero, (26) defines a Petrov representation of x_π strictly bigger than α^*_π, which is a contradiction. The proof of part (*2*) is analogous.

(*3*). Let $x \in \operatorname{ri} F$ have a maximal Petrov representation α^*_x. Suppose that there is an index $j_0 \in \{j \in \mathbb{N}_m : \alpha^*_{\pi,j} = 1\} \setminus \{j \in \mathbb{N}_m : \alpha^*_{x,j} = 1\}$. Note that, if $\alpha^*_{\pi,j} = 1$ then, $\alpha^*_{x,j} - \alpha^*_{\pi,j} \le 0$ and if $\alpha^*_{\pi,j} = 0$ then, $\alpha^*_{x,j} - \alpha^*_{\pi,j} \ge 0$, while in addition

$$\alpha^*_{x,j_0} - \alpha^*_{\pi,j_0} < 0.$$

Thus, for all positive t close enough to zero,

$$x_\pi = \sum_{j=1}^{m} (\alpha^*_{\pi,j} + t(\alpha^*_{x,j} - \alpha^*_{\pi,j})) y_j$$

defines a Petrov representation of x_π strictly bigger than α^*_π. This contradiction shows that $\{j \in \mathbb{N}_m : \alpha^*_{\pi,j} = 1\} \subset \{j \in \mathbb{N}_m : \alpha^*_{x,j} = 1\}$. The verification of the opposite inclusion is analogous, after exchanging the roles of α^*_π and α^*_x, as well as the proof of the equality $\{j \in \mathbb{N}_m : \alpha^*_{\pi,j} = 0\} = \{j \in \mathbb{N}_m : \alpha^*_{x,j} = 0\}$. $\square$

For any subset $\pi \subset \mathbb{N}_\ell$, fix a maximal Petrov representation α^*_π of x_π, and define the following system of linear equations with variables $\mathcal{E} := (\epsilon_{k,j})_{k,j=1,1}^{\ell,m}$ and $\Gamma := \{\gamma_{\pi,j} : \pi \subset \mathbb{N}_\ell, j \in \mathbb{N}_m\}$:

$$\sum_{j=1}^{m}\left(\sum_{k\in\pi}\epsilon_{k,j}\right)y_j = \sum_{j=1}^{m}\gamma_{\pi,j}y_j \quad \text{for all } \pi \subset \mathbb{N}_\ell;$$
$$\sum_{j=1}^{m}\gamma_{\pi,j} = 0 \quad \text{for all } \pi \subset \mathbb{N}_\ell; \tag{27}$$
$$\gamma_{\pi,j} = 0 \quad \text{whenever } \alpha^*_{\pi,j} \in \{0,1\}.$$

By Lemma 5.1, the system is independent of the maximal Petrov representations chosen.

Theorem 5.2. *The sequence $x \in L^\ell_\prec(y)$ is an extreme point of $L^\ell_\prec(y)$ if and only if every solution of the system* (27) *satisfies $\sum_{j=1}^m \epsilon_{k,j}y_j = 0$ for all $k \in \mathbb{N}_\ell$.*

Proof. Suppose that the linear system (27) has a solution $(\mathcal{E},\Gamma)$ with $\sum_{j=1}^m \epsilon_{k,j}y_j \neq 0$ for some $k \in \mathbb{N}_\ell$. By homogeneity, for any number t, $(t\mathcal{E}, t\Gamma)$ is also a solution to (27). Define the vectors

$$x'_k := \sum_{j=1}^{m}(\alpha^*_{k,j} + t\epsilon_{k,j})y_j \quad \text{and} \quad x''_k := \sum_{j=1}^{m}(\alpha^*_{k,j} - t\epsilon_{k,j})y_j, \quad k \in \mathbb{N}_\ell,$$

and let $x' := (x'_k)_{k=1}^\ell$ and $x'' := (x''_k)_{k=1}^\ell$. By assumption, we have that x', x, and x'' are distinct sequences of vectors. We need to show now that $x', x'' \in L^\ell_\prec(y)$. Fix a subset $\pi \subset \mathbb{N}_\ell$ and observe that

$$\begin{aligned} x'_\pi &= \sum_{k\in\pi}x'_k = \sum_{k\in\pi}\sum_{j=1}^{m}(\alpha^*_{k,j} + t\epsilon_{k,j})y_j = \sum_{j=1}^{m}\alpha^*_{\pi,j}y_j + t\sum_{j=1}^{m}\left(\sum_{k\in\pi}\epsilon_{k,j}\right)y_j \\ &= \sum_{j=1}^{m}\alpha^*_{\pi,j}y_j + t\sum_{j=1}^{m}\gamma_{\pi,j}y_j = \sum_{j=1}^{m}(\alpha^*_{\pi,j} + t\gamma_{\pi,j})y_j. \end{aligned}$$

For all t close to zero, the last expression is a Petrov representation of x'_π showing that $x'_\pi \in A_{\mathrm{card}\,(\pi)}$. That is $x' \in L^\ell_\prec(y)$. The argument for x'' is analogous. The fact that $x' \neq x''$ and $x = (x' + x'')/2$ implies that x is not an extreme point.

Suppose now, that $x \in L^\ell_\prec(y)$ is not an extreme point. Then, there are distinct $x', x'' \in L^\ell_\prec(y)$ with $x = (x' + x'')/2$. Since $x'_k \in A_1$, there are convex representations

$$x'_k := \sum_{j=1}^{m}\alpha'_{k,j}y_j \quad \text{for } k = 1, ..., \ell,$$

and we define $\epsilon_{k,j} := \alpha'_{k,j} - \alpha^*_{k,j}$. We show now that $\mathcal{E} := (\epsilon_{k,j})_{k,j=1,1}^{\ell,m}$ together with some values $\Gamma := \{\gamma_{\pi,j} : \pi \subset \mathbb{N}_\ell, j \in \mathbb{N}_m\}$ is a solution of (27). That solution, using the fact that $x \neq x'$, trivially satisfies $\sum_{j=1}^m \epsilon_{k,j}y_j \neq 0$ for some $k \in \mathbb{N}_\ell$. Fix a set $\pi \subset \mathbb{N}_\ell$

and let F be the unique extreme face of $A_{\mathrm{card}\,(\pi)}$ with the property $x_\pi \in \mathrm{ri}\, F$. Since $x_\pi = (x'_\pi + x''_\pi)/2$, we may assume, without loss of generality, that x' is close enough to x so that $x'_\pi \in \mathrm{ri}\, F$. Let the maximal Petrov representation of $x'_\pi \in \mathrm{ri}\, F \subset A_{\mathrm{card}\,(\pi)}$ be $x'_\pi = \sum_{j=1}^m \alpha'^*_{\pi,j} y_j$. Then,

$$\begin{aligned}\sum_{j=1}^m \alpha'^*_{\pi,j} y_j = &\sum_{k\in\pi} x'_k = \sum_{k\in\pi}\sum_{j=1}^m \alpha'_{k,j} y_j = \sum_{k\in\pi}\sum_{j=1}^m (\alpha^*_{k,j} + \epsilon_{k,j}) y_j \\ = &\sum_{k\in\pi} x_k + \sum_{j=1}^m \left(\sum_{k\in\pi} \epsilon_{k,j}\right) y_j = \sum_{j=1}^m \alpha^*_{\pi,j} y_j + \sum_{j=1}^m \left(\sum_{k\in\pi} \epsilon_{k,j}\right) y_j,\end{aligned}$$

and consequently

$$\sum_{j=1}^m \left(\sum_{k\in\pi} \epsilon_{k,j}\right) y_j = \sum_{j=1}^m (\alpha'^*_{\pi,j} - \alpha^*_{\pi,j}) y_j.$$

Using Lemma 5.1, part (*3*), for the quantities $\gamma_{\pi,j} := \alpha'^*_{\pi,j} - \alpha^*_{\pi,j}$ for all $j \in \mathbb{N}_m$ we have $\sum_{j=1}^m \gamma_{\pi,j} = 0$ with $\gamma_{\pi,j} = 0$ whenever $\alpha^*_{\pi,j} \in \{0, 1\}$. This shows that the linear system (27) has a solution with the required properties. $\square$

Acknowledgements. We thank an anonymous referee for his/her valuable comments and for pointing out reference [3].

References

[1] R. Bhatia: Matrix Analysis, Springer, New York (1997).

[2] A. L. Dontchev, R. T. Rockafellar: Implicit Functions and Solution Mappings. A View from Variational Analysis, Springer, New York (2008).

[3] B. Ćurgus, K. Kołodziejczyk: On a convex operator for finite sets, Discrete Appl. Math. 155 (2007) 1774–1792.

[4] P. Fischer, J. A. R. Holbrook: Balayage defined by the nonnegative convex functions, Proc. Amer. Math. Soc. 79 (1980) 445–448.

[5] S. M. Malamud: Inverse spectral problem for normal matrices and the Gauss-Lucas theorem, Trans. Amer. Math. Soc 357(10) (2004) 4043–4064.

[6] A. W. Marshall, I. Olkin: Inequalities: Theory of Majorization and its Applications, Mathematics in Science and Engineering 143, Academic Press, New York (1979).

[7] A. L. Peressini, F. E. Sullivan, J. J. Uhl, Jr.: The Mathematics of Nonlinear Programming, Undergraduate Texts in Mathematics, Springer, New York (1988).

[8] R. T. Rockefellar: Convex Analysis, Princeton University Press, Princeton (1970).

[9] F. A. Valentine: Convex Sets, McGraw-Hill, New York (1964).