

# A taxonomy of monotonicity properties for the aggregation of multidimensional data

Raúl Pérez-Fernández<sup>a,b,\*</sup>, Bernard De Baets<sup>a</sup>, Marek Gagolewski<sup>c,d</sup>

<sup>a</sup>*KERMIT, Department of Data Analysis and Mathematical Modelling, Ghent University, Coupure links 653, 9000 Gent, Belgium*

<sup>b</sup>*Department of Statistics and O.R. and Mathematics Didactics, University of Oviedo, Calle Federico García Lorca 18, 33007 Oviedo, Spain*

<sup>c</sup>*Faculty of Mathematics and Information Science, Warsaw University of Technology ul. Koszykowa 75, 00-662 Warsaw, Poland*

<sup>d</sup>*Systems Research Institute, Polish Academy of Sciences ul. Newelska 6, 01-447 Warsaw, Poland*

---

## Abstract

The property of monotonicity, which requires a function to preserve a given order, has been considered the standard in the aggregation of real numbers for decades. In this paper, we argue that, for the case of multidimensional data, an order-based definition of monotonicity is far too restrictive. We propose several meaningful alternatives to this property not involving the preservation of a given order by returning to its early origins stemming from the field of calculus. Numerous aggregation methods for multidimensional data commonly used by practitioners are studied within our new framework.

*Keywords:* Monotonicity; Aggregation; Multidimensional data; Centroid; Spatial median.

---

## 1. Introduction

The concept of monotonicity originally arose from calculus but, probably due to its intuitive nature, spread to other fields of science. In particular, the study of monotonicity was later embraced by the field of order theory. Admittedly, the concept of monotonicity is understood differently in each of these domains: whereas in the field of calculus it is a property based on

---

\*Corresponding author; email: raul.perezfernandez@ugent.be

directional derivatives (or directional finite differences in case the function is not differentiable), in the field of order theory it is a property based on the preservation of some order. Due to the fact that both notions are equivalent for functions of the type  $A : \mathbb{R}^n \rightarrow \mathbb{R}$  for some  $n$ , they are most often used interchangeably.

In the field of data aggregation [1–5], monotonicity is a largely venerated property that has been typically understood in the second (order-theoretical) sense, generalized to functions of the type  $A : X^n \rightarrow X$ , for a given poset  $(X, \leq)$ . Although most practitioners admittedly consider the poset to simply be the unit interval  $[0, 1]$  with the usual natural order  $\leq$ , aggregation on different types of poset (e.g., interval-valued data) has also built upon the property of monotonicity [6, 7].

**Definition 1.** Consider a poset  $(X, \leq)$ . A function  $A : X^n \rightarrow X$  is called monotone (increasing)<sup>1</sup> if, for any  $\mathbf{x}, \mathbf{y} \in X^n$ , the fact that<sup>2</sup>  $\mathbf{x} \leq_n \mathbf{y}$  implies that  $A(\mathbf{x}) \leq A(\mathbf{y})$ .

The most common class of aggregation functions is that of averaging functions (often referred to as *means*) [3–5, 8]. Although the first known definition of a mean [9] just required a function  $A : X^n \rightarrow X$  to be idempotent, i.e.,  $A(x, \dots, x) = x$  for any  $x \in X$ , the current understanding adds the property of monotonicity.

Historically considered a standard, the property of monotonicity has received some fierce criticism in recent years. For instance, many scholars point out that the mode, which should arguably be one of the most natural examples of an aggregation function, is not monotone. For this very reason, the property of weak monotonicity, which generalizes both translation equivariance and monotonicity for functions of the type  $A : \mathbb{R}^n \rightarrow \mathbb{R}$ , has been recently proposed [10]. Several studies on the conditions under which some non-monotone functions (e.g., Gini means) are weakly monotone have been addressed in [11, 12].

**Definition 2.** A function  $A : \mathbb{R}^n \rightarrow \mathbb{R}$  is called weakly monotone if, for any  $\mathbf{x} \in \mathbb{R}^n$ , any  $a > 0$  and  $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^n$ , it holds that  $A(\mathbf{x} + a\mathbf{1}) \geq A(\mathbf{x})$ .

---

<sup>1</sup>The term ‘monotone’ actually refers to two types of behaviour: monotone increasing (isotone) and monotone decreasing (antitone). However, for historical reasons, we adhere to the term ‘monotone’ for referring to ‘monotone increasing’ throughout this paper.

<sup>2</sup>Given an order  $\leq$  on a set  $X$ , the product order on  $X^m$  is defined as  $\mathbf{x} \leq_m \mathbf{y}$  if  $x_i \leq y_i$ , for any  $i \in \{1, \dots, m\}$ .

Yet, one can argue that the above definition is much too weak.

Outside the field of aggregation theory, aggregation on different types of mathematical structures has been addressed despite the lack of a natural order on these structures. For instance, the aggregation of rankings has been addressed in the field of social choice theory since (at the latest) the eighteenth century [13–15]. Not only the aggregation of rankings has attracted scientific attention, but also the aggregation of members of different families of binary relations [16–18]. In most of these cases, the inclusion-based order on the set of binary relations does not result in an interesting poset. The aggregation of strings [19] has also been studied despite the lack of a meaningful order on the set of strings. In most cases, these data fusion tasks are performed as the minimization of a ‘distance’ or ‘penalty’ [20–22], as in the prominently-studied case of penalty-based aggregation of real numbers [23–25].

In this paper, we focus on the aggregation of multidimensional data, following the generic direction started in [26] and further discussing different monotonicity properties that could be of interest in this setting. More precisely, given  $n$  points in  $\mathbb{R}^d$ , we would like to study aggregation methods that output a single point in  $\mathbb{R}^d$  as a result. This setting prominently appears in many important disciplines, including data analysis and computational statistics, see, e.g., [27, 28]. This is because most objects in data science are represented in terms of feature vectors.

Already in [26] it was noted that several popular multidimensional aggregation functions, like the Euclidean median or the Euclidean center, do not fulfill a classical extension of the order-theoretic monotonicity (referred to as componentwise monotonicity). However, this interesting problem was left as a topic of further research, as the main attention of the paper was devoted to the behaviour of different aggregation functions after being subjected to various geometrical transformations, like the study of equivariances to orthogonal, translation, scaling, and other affine transformations. In [29], the present authors pointed out an inherent difficulty in the aggregation of multidimensional data: if a weighted centroid is to be avoided, one must choose *either* the (order-theoretic) monotonicity *or* a desirable behaviour with regard to orthogonal transformations. In this paper, we go much further by introducing new monotonicity properties for the aggregation of multidimensional data that are fulfilled by more general classes than that of weighted centroids. Such properties are especially important due to their interpretability.

ity and the fact that they assure a proper behaviour of the functions' outputs when the inputs are subjected to some basic geometrical transformations.

The remainder of the paper is set out as follows. In Section 2, we point out several drawbacks of the classical order-theoretic monotonicity property when dealing with multidimensional data. Four new potential definitions of monotonicity are introduced in Section 3. These monotonicity properties are compared to the simplistic componentwise monotonicity, resulting in the introduction of the natural property of ultramonotonicity in Section 4. Some prominent examples of 'aggregation functions' for multidimensional data are discussed in Section 5 and classified within this newly introduced taxonomic framework. Finally, in Section 6, we discuss the practical implications of our findings and outline some ideas for further research.

## 2. An inherent difficulty in the aggregation of multidimensional data

Assume we are dealing with  $n$  points  $\mathbf{x}^{(i)} = (x_1^{(i)}, \dots, x_d^{(i)})^T$  in  $\mathbb{R}^d$ , where  $i \in \{1, \dots, n\}$ . Note that, throughout this paper, we use the letter  $i$  for indexing the different points (thus,  $i \in \{1, \dots, n\}$ ) and the letter  $j$  for indexing the different dimensions/features (thus,  $j \in \{1, \dots, d\}$ ). The goal is to aggregate the  $n$  points  $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}$  into a unique point (hereinafter referred to as the aggregate) by means of a function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$ , as follows:

$$A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) = A \begin{pmatrix} x_1^{(1)} & \cdots & x_1^{(n)} \\ \vdots & \ddots & \vdots \\ x_d^{(1)} & \cdots & x_d^{(n)} \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_d \end{pmatrix}.$$

In the unidimensional case ( $d = 1$ ), this setting reduces to the classical aggregation of real numbers and the monotonicity property is determined by the classical order relation on  $\mathbb{R}$ , as in Definition 1. However, in the multidimensional case ( $d \geq 2$ ), it is unclear how to define the monotonicity property. A first idea would be to consider the product order  $\leq_d$  on  $\mathbb{R}^d$  and define monotonicity w.r.t. this order. This is illustrated in the left side of Figure 1.

Unfortunately, this order-based monotonicity might be too restrictive. Let us consider one of the most classical methods for aggregating multidimensional data, the spatial median (also referred to as  $L_1$  median, geometric median, Euclidean median, Fermat–Weber point, among other names),

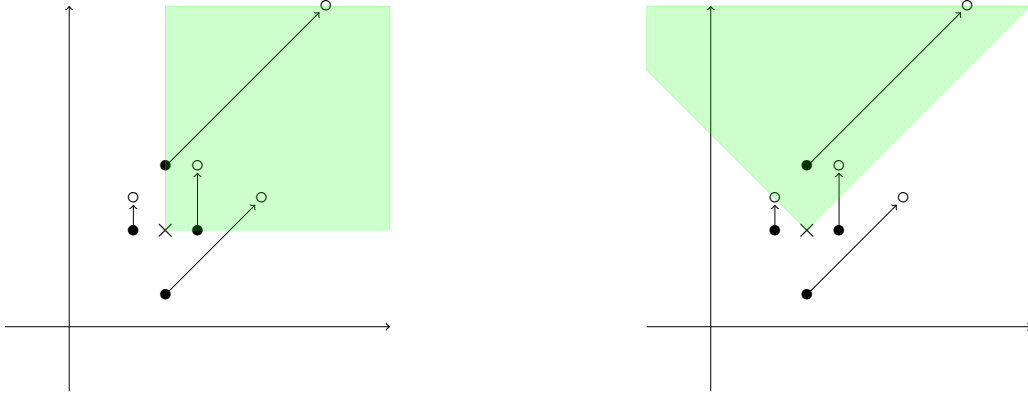


Figure 1: Graphical interpretation of the property of monotonicity w.r.t.  $\leq_d$  (left side) and w.r.t. a rotated  $\leq_d$  (right side). The black circles represent four initial data points, whereas the white circles represent the new positions of the initial data points. The point marked with the symbol  $\times$  represents the initial aggregate and the green area indicates the area in which the new aggregate must be found.

i.e., the point that minimizes the sum of Euclidean distances to the points to be aggregated. This function has an undesirable behaviour. Not only is it not  $\leq_d$ -monotone, but it also fails the weaker condition of  $\leq_d$ -non-decreasingness<sup>3</sup>. For instance, let us recall the example given in [26]. Consider  $\mathbf{x}^{(1)} = (0, 0)^T$ ,  $\mathbf{x}^{(2)} = (1, -5)^T$  and  $\mathbf{x}^{(3)} = (20, 1)^T$ . The spatial median of  $\mathbf{x}^{(1)}$ ,  $\mathbf{x}^{(2)}$  and  $\mathbf{x}^{(3)}$  is given by  $A(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}) \approx (1.961, -2.305)^T$ , whereas for  $\mathbf{x}^{(3)'} = \mathbf{x}^{(3)} + (1980, 1)^T$ , we get  $A(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)'}) \approx (1.946, -3.351)^T \leq_2 (1.961, -2.305)^T \approx A(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)})$ .

Even if too restrictive, one could consider this order-based monotonicity to be perfectly fine for certain types of multidimensional data for which all components are independent and expected to be increasing all at the same time. Yet, when the components are not independent (and this is often the case when dealing with multidimensional data), it might be carrying the wrong semantics – since there may be no associated notion of order. Note that with this property it is possible that if all data points move ‘right’, the aggregate moves ‘up’. As another illustrative example, let us consider

<sup>3</sup>Consider a poset  $(X, \leq)$  and  $n \in \mathbb{N}$ . A function  $A : X^n \rightarrow X$  is called non-decreasing if, for any  $\mathbf{x}, \mathbf{y} \in X^n$ , the fact that  $\mathbf{x} \leq_n \mathbf{y}$  implies that  $A(\mathbf{y}) \not\leq A(\mathbf{x})$ .

spatial data coming from satellite imagery, represented as points on a two-dimensional rectangular map. There, we would not only expect the aggregate to move to the top-right corner of the map whenever all data points move towards this top-right corner, but we would also expect the same in all possible directions. This is illustrated in the right side of Figure 1. In [29], we proposed the notion of orthomonotonicity for solving this (potential) semantic issue.

**Definition 3.** A function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is called orthomonotone if, for any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}), (\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(n)}) \in \mathbb{R}^{d \times n}$  and any orthogonal matrix<sup>4</sup>  $\mathbf{O} \in \mathbb{R}^{d \times d}$ , the fact that  $\mathbf{O} \mathbf{x}^{(i)} \leq_d \mathbf{O} \mathbf{y}^{(i)}$  for any  $i \in \{1, \dots, n\}$ , implies that  $\mathbf{O} A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \leq_d \mathbf{O} A(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(n)})$ .

Unfortunately, the only idempotent functions satisfying such orthomonotonicity turned out to be weighted centroids [29]. In the upcoming section, we will discuss new monotonicity properties that are satisfied by a broader class of idempotent functions than that of weighted centroids.

### 3. A taxonomy of monotonicity properties for multidimensional data

#### 3.1. Definitions

In this section, we will introduce and analyze four natural monotonicity properties (we refer to Section 5 for illustrative examples of functions satisfying each of these monotonicity properties). First, following the rationale behind the introduction of weak monotonicity, we distinguish two different settings concerning whether we allow one single point to be moved (S) or we allow multiple points to be moved simultaneously (M). Second, following the rationale behind the inner product-based property of monotonicity in functional analysis<sup>5</sup>, we distinguish two different settings concerning whether we allow the aggregate to preserve the direction in which the data points are moving (P) or we allow the direction in which the aggregate moves to be consistent with the direction in which the data points are moving (C). Thus, we have four possible combinations of these settings that we will denote by

<sup>4</sup>An orthogonal matrix is a matrix  $\mathbf{O} \in \mathbb{R}^d$  such that  $\mathbf{O}^T = \mathbf{O}^{-1}$ .

<sup>5</sup>In the field of functional analysis, a function  $A : X \rightarrow X$  on a Hilbert space  $X$  is called monotone if  $\langle y - x, A(y) - A(x) \rangle \geq 0$ , for any  $x, y \in X$ .

SP-, MP-, SC- and MC-monotonicity. In Figure 2, we illustrate all of these properties, which will be formalized right after.

First, we consider SP-monotonicity, i.e., we allow one single point to move in any direction, and we require the aggregate to move in this same direction. For a graphical interpretation of this property, see the top left part of Figure 2. We use the notation  $\vec{\mathbf{0}} = (0, \dots, 0)^T$ .

**Definition 4.** For a given direction  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\vec{\mathbf{0}}\}$ , a function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is called  $\vec{\mathbf{u}}$ -SP-monotone if, for any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$ , any  $i \in \{1, \dots, n\}$  and any step size  $t \geq 0$ , there exists  $k \geq 0$  such that:

$$A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)} + t\vec{\mathbf{u}}, \dots, \mathbf{x}^{(n)}) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)}, \dots, \mathbf{x}^{(n)}) = k\vec{\mathbf{u}}.$$

**Definition 5.** A function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is called SP-monotone if it is  $\vec{\mathbf{u}}$ -SP-monotone with respect to all directions  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\vec{\mathbf{0}}\}$ .

Second, we consider MP-monotonicity, i.e., we allow all points to move in the same direction with the same step size, and we require the aggregate to move in this same direction (possibly with a different step size). For a graphical interpretation of this property, see the top right part of Figure 2.

**Definition 6.** For a given direction  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\vec{\mathbf{0}}\}$ , a function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is called  $\vec{\mathbf{u}}$ -MP-monotone if, for any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$  and any step size  $t \geq 0$ , there exists  $k \geq 0$  such that:

$$A(\mathbf{x}^{(1)} + t\vec{\mathbf{u}}, \dots, \mathbf{x}^{(n)} + t\vec{\mathbf{u}}) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) = k\vec{\mathbf{u}}.$$

**Definition 7.** A function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is called MP-monotone if it is  $\vec{\mathbf{u}}$ -MP-monotone with respect to all directions  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\vec{\mathbf{0}}\}$ .

Third, we consider SC-monotonicity, i.e., we allow one single point to move in any direction, and we require the direction in which the aggregate moves to be consistent with the direction in which the data point is moving (i.e., the angle between both directions belongs to  $[-\frac{\pi}{2}, \frac{\pi}{2}]$ ). For a graphical interpretation of this property, see the bottom left part of Figure 2. Note that the fact that the angle between two vectors  $\vec{\mathbf{u}}$  and  $\vec{\mathbf{v}}$  belongs to  $[-\frac{\pi}{2}, \frac{\pi}{2}]$  is equivalent to the dot product of  $\vec{\mathbf{u}}$  and  $\vec{\mathbf{v}}$  being positive, since

$$\vec{\mathbf{u}} \cdot \vec{\mathbf{v}} = \vec{\mathbf{u}}^T \vec{\mathbf{v}} = \sum_{j=1}^d u_j v_j = \|\vec{\mathbf{u}}\| \|\vec{\mathbf{v}}\| \cos(\theta),$$

where  $\theta$  represents the angle between  $\vec{\mathbf{u}}$  and  $\vec{\mathbf{v}}$ .

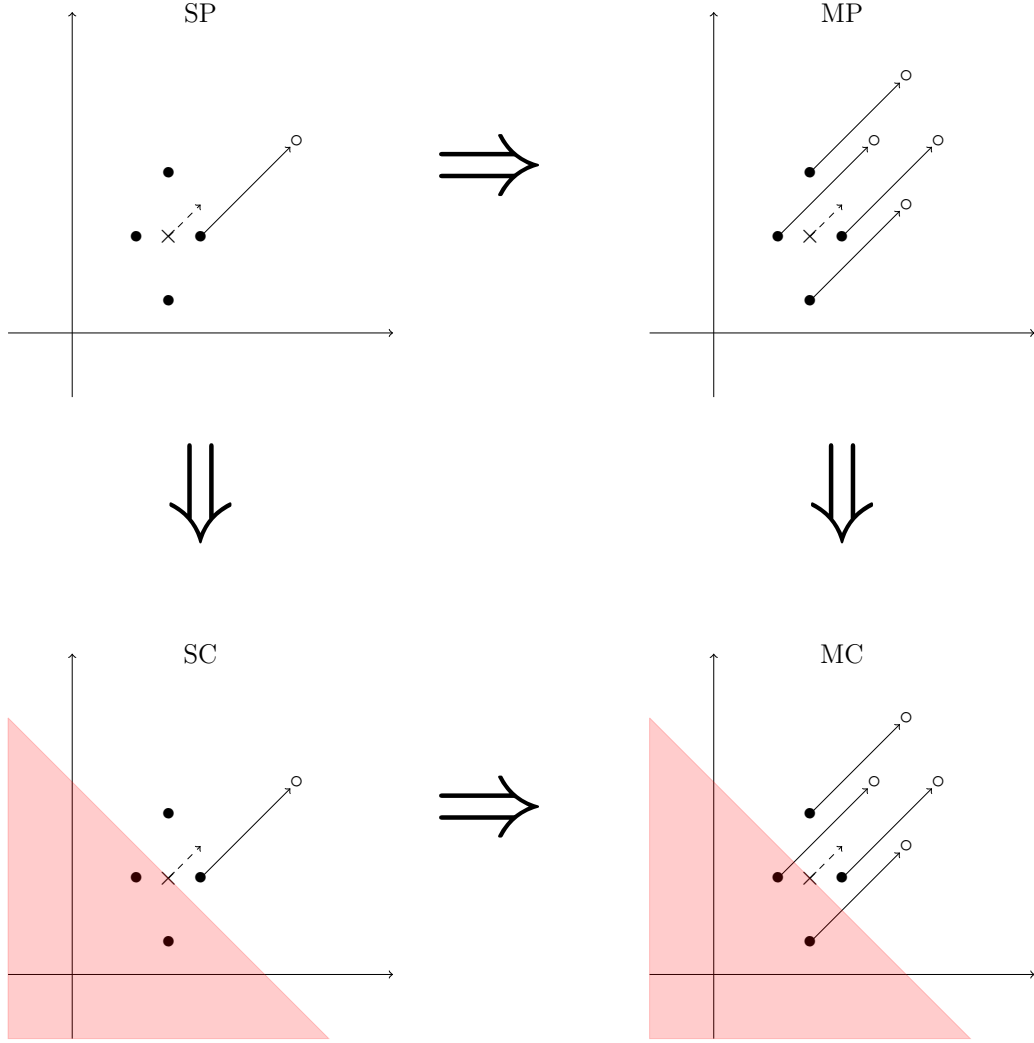


Figure 2: Graphical comparison of the different monotonicity properties defined in this section. The black circles represent four initial data points, whereas the white circles represent the new positions of (some of) the initial data points. The point marked with the symbol  $\times$  represents the initial aggregate and the dashed arrow represents the direction in which we are considering the monotonicity. The red area indicates the area to which the aggregate must *not* go.



**Definition 8.** For a given direction  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\vec{\mathbf{0}}\}$ , a function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is called  $\vec{\mathbf{u}}$ -SC-monotone if, for any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$ , any  $i \in \{1, \dots, n\}$  and any step size  $t \geq 0$ , it holds that

$$\left( A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)} + t\vec{\mathbf{u}}, \dots, \mathbf{x}^{(n)}) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)}, \dots, \mathbf{x}^{(n)}) \right) \cdot \vec{\mathbf{u}} \geq 0.$$

**Definition 9.** A function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is called SC-monotone if it is  $\vec{\mathbf{u}}$ -SC-monotone with respect to all directions  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\vec{\mathbf{0}}\}$ .

Fourth, we consider MC-monotonicity, i.e., we allow all points to move in the same direction with the same step size, and we require the direction in which the aggregate moves to be consistent with the direction in which the data points are moving. For a graphical interpretation of this property, see the bottom right part of Figure 2.

**Definition 10.** For a given direction  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\vec{\mathbf{0}}\}$ , a function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is called  $\vec{\mathbf{u}}$ -MC-monotone if, for any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$  and any step size  $t \geq 0$ , it holds that

$$\left( A(\mathbf{x}^{(1)} + t\vec{\mathbf{u}}, \dots, \mathbf{x}^{(n)} + t\vec{\mathbf{u}}) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \right) \cdot \vec{\mathbf{u}} \geq 0.$$

**Definition 11.** A function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is called MC-monotone if it is  $\vec{\mathbf{u}}$ -MC-monotone with respect to all directions  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\vec{\mathbf{0}}\}$ .

*Remark 12.* Note that both the magnitude and the sign of the direction  $\vec{\mathbf{u}}$  do not matter for defining any of the above monotonicity properties w.r.t.  $\vec{\mathbf{u}}$ . Formally, for any function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  and any direction  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\vec{\mathbf{0}}\}$ , the following results hold:

- (i)  $A$  is  $\vec{\mathbf{u}}$ -SP-monotone if and only if it is  $(c\vec{\mathbf{u}})$ -SP-monotone for any  $c \neq 0$ .
- (ii)  $A$  is  $\vec{\mathbf{u}}$ -MP-monotone if and only if it is  $(c\vec{\mathbf{u}})$ -MP-monotone for any  $c \neq 0$ .
- (iii)  $A$  is  $\vec{\mathbf{u}}$ -CP-monotone if and only if it is  $(c\vec{\mathbf{u}})$ -SC-monotone for any  $c \neq 0$ .
- (iv)  $A$  is  $\vec{\mathbf{u}}$ -MC-monotone if and only if it is  $(c\vec{\mathbf{u}})$ -MC-monotone for any  $c \neq 0$ .

### 3.2. Relations between the new notions

An important (and straightforward) conclusion concerning the relationships between the above monotonicity properties is provided in the following proposition.

**Proposition 13.** *For any function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  and any direction  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\vec{\mathbf{0}}\}$ , the following results hold:*

- (i) *If  $A$  is  $\vec{\mathbf{u}}$ -SP-monotone, then it is  $\vec{\mathbf{u}}$ -MP-monotone.*
- (ii) *If  $A$  is  $\vec{\mathbf{u}}$ -SP-monotone, then it is  $\vec{\mathbf{u}}$ -SC-monotone.*
- (iii) *If  $A$  is  $\vec{\mathbf{u}}$ -MP-monotone, then it is  $\vec{\mathbf{u}}$ -MC-monotone.*
- (iv) *If  $A$  is  $\vec{\mathbf{u}}$ -SC-monotone, then it is  $\vec{\mathbf{u}}$ -MC-monotone.*

The following theorem follows as a corollary from the proposition above. This result was already anticipated in Figure 2.

**Theorem 14.** *For any function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$ , the following results hold:*

- (i) *If  $A$  is SP-monotone, then it is MP-monotone.*
- (ii) *If  $A$  is SP-monotone, then it is SC-monotone.*
- (iii) *If  $A$  is MP-monotone, then it is MC-monotone.*
- (iv) *If  $A$  is SC-monotone, then it is MC-monotone.*

### 3.3. Interaction with rotations

When dealing with multidimensional data, one often applies orthogonal transformations to an input data set (for example, when performing Principal Component Analysis). Let us recall that the dot product is invariant under orthogonal transformations, i.e., for any  $\vec{\mathbf{u}}$  and  $\vec{\mathbf{v}}$  and any orthogonal matrix  $\mathbf{O}$ , it holds that  $\mathbf{O}\vec{\mathbf{u}} \cdot \mathbf{O}\vec{\mathbf{v}} = (\mathbf{O}\vec{\mathbf{u}})^T(\mathbf{O}\vec{\mathbf{v}}) = \vec{\mathbf{u}}^T \mathbf{O}^T \mathbf{O} \vec{\mathbf{v}} = \vec{\mathbf{u}}^T \mathbf{O}^{-1} \mathbf{O} \vec{\mathbf{v}} = \vec{\mathbf{u}} \cdot \vec{\mathbf{v}}$ . Typical examples of orthogonal transformations are rotations and reflections.

A function  $A$  is called orthogonal equivariant if  $A(\mathbf{O}\mathbf{X}) = \mathbf{O}A(\mathbf{X})$  for every orthogonal matrix  $\mathbf{O}$  and, in particular, it is called rotation equivariant if  $A(\mathbf{R}\mathbf{X}) = \mathbf{R}A(\mathbf{X})$  for every rotation matrix  $\mathbf{R}$ . Please note that all functions obtained as minimizers of penalty functions based on the Euclidean distance (e.g. the spatial median, the centroid and the Euclidean center from the functions in Section 5) trivially are orthogonal equivariant [26].

If a function is rotation equivariant, then one of the above monotonicity properties holding with respect to one direction implies that this same monotonicity property holds with respect to all directions.

**Proposition 15.** *For any rotation equivariant function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$ , the following results hold:*

- (i)  *$A$  is SP-monotone if and only if it is  $\vec{\mathbf{u}}$ -SP-monotone with respect to at least one direction  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ .*
- (ii)  *$A$  is MP-monotone if and only if it is  $\vec{\mathbf{u}}$ -MP-monotone with respect to at least one direction  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ .*
- (iii)  *$A$  is SC-monotone if and only if it is  $\vec{\mathbf{u}}$ -SC-monotone with respect to at least one direction  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ .*
- (iv)  *$A$  is MC-monotone if and only if it is  $\vec{\mathbf{u}}$ -MC-monotone with respect to at least one direction  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ .*

*Proof.* (i) Consider  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$  and assume that  $A$  is  $\vec{\mathbf{u}}$ -SP-monotone. Consider any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$ , any  $i \in \{1, \dots, n\}$  and any  $t \geq 0$ , and  $\vec{\mathbf{v}} = \mathbf{R}\vec{\mathbf{u}}$  for some rotation matrix  $\mathbf{R}$ . It holds that

$$\begin{aligned} & A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)} + t\vec{\mathbf{v}}, \dots, \mathbf{x}^{(n)}) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)}, \dots, \mathbf{x}^{(n)}) \\ &= \mathbf{R} \left( A(\mathbf{R}^{-1}\mathbf{x}^{(1)}, \dots, \mathbf{R}^{-1}\mathbf{x}^{(i)} + t\mathbf{R}^{-1}\mathbf{R}\vec{\mathbf{u}}, \dots, \mathbf{R}^{-1}\mathbf{x}^{(n)}) \right. \\ & \quad \left. - A(\mathbf{R}^{-1}\mathbf{x}^{(1)}, \dots, \mathbf{R}^{-1}\mathbf{x}^{(i)}, \dots, \mathbf{R}^{-1}\mathbf{x}^{(n)}) \right) \\ &= \mathbf{R} k \vec{\mathbf{u}} = k \vec{\mathbf{v}}, \end{aligned}$$

for some  $k \geq 0$  due to the fact that  $A$  is  $\vec{\mathbf{u}}$ -SP-monotone. Thus,  $A$  is SP-monotone.

(iii) Consider  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$  and assume that  $A$  is  $\vec{\mathbf{u}}$ -SC-monotone. Consider any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$ , any  $i \in \{1, \dots, n\}$  and any  $t \geq 0$ , and  $\vec{\mathbf{v}} = \mathbf{R}\vec{\mathbf{u}}$  for some rotation matrix  $\mathbf{R}$ . It holds that

$$\begin{aligned} & \left( A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)} + t\vec{\mathbf{v}}, \dots, \mathbf{x}^{(n)}) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)}, \dots, \mathbf{x}^{(n)}) \right) \cdot \vec{\mathbf{v}} \\ &= \mathbf{R}^{-1} \left( A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)} + t\vec{\mathbf{v}}, \dots, \mathbf{x}^{(n)}) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)}, \dots, \mathbf{x}^{(n)}) \right) \cdot \mathbf{R}^{-1}\vec{\mathbf{v}} \\ &= \left( A(\mathbf{R}^{-1}\mathbf{x}^{(1)}, \dots, \mathbf{R}^{-1}\mathbf{x}^{(i)} + t\vec{\mathbf{u}}, \dots, \mathbf{R}^{-1}\mathbf{x}^{(n)}) \right. \\ & \quad \left. - A(\mathbf{R}^{-1}\mathbf{x}^{(1)}, \dots, \mathbf{R}^{-1}\mathbf{x}^{(i)}, \dots, \mathbf{R}^{-1}\mathbf{x}^{(n)}) \right) \cdot \vec{\mathbf{u}} \geq 0, \end{aligned}$$

due to the fact that  $A$  is  $\vec{u}$ -SC-monotone and the dot product is rotation invariant. Thus,  $A$  is SC-monotone.

Statements (ii) and (iv) can be proven in a very similar manner.  $\square$

### 3.4. New monotonicity properties as generalizations of the unidimensional (weak) monotonicity

Note that SP-monotonicity and SC-monotonicity coincide if  $d = 1$  and reduce to the classical monotonicity in the setting of real numbers.

**Proposition 16.** *For any unidimensional function  $A : \mathbb{R}^n \rightarrow \mathbb{R}$ , the following results are equivalent:*

- (i)  $A$  is monotone in the sense of Definition 1.
- (ii)  $A$  is SP-monotone.
- (iii)  $A$  is SC-monotone.

*Proof.* (i) $\Rightarrow$ (ii) Consider any direction  $\vec{u} \in \mathbb{R} \setminus \{0\}$ , any  $(x^{(1)}, \dots, x^{(n)}) \in \mathbb{R}^n$ , any  $i \in \{1, \dots, n\}$  and any  $t \geq 0$ . If  $\vec{u} > 0$ , then it follows that

$$A(x^{(1)}, \dots, x^{(i)} + t\vec{u}, \dots, x^{(n)}) - A(x^{(1)}, \dots, x^{(n)}) \geq 0,$$

or, equivalently,

$$A(x^{(1)}, \dots, x^{(i)} + t\vec{u}, \dots, x^{(n)}) - A(x^{(1)}, \dots, x^{(n)}) = k\vec{u},$$

for some constant  $k \geq 0$ .

Similarly, if  $\vec{u} < 0$ , then it follows that

$$A(x^{(1)}, \dots, x^{(i)} + t\vec{u}, \dots, x^{(n)}) - A(x^{(1)}, \dots, x^{(n)}) \leq 0,$$

or, equivalently,

$$A(x^{(1)}, \dots, x^{(i)} + t\vec{u}, \dots, x^{(n)}) - A(x^{(1)}, \dots, x^{(n)}) = k\vec{u},$$

for some constant  $k \geq 0$ .

(ii) $\Rightarrow$ (iii) Follows from Theorem 14.

(iii) $\Rightarrow$ (i) Consider  $(x^{(1)}, \dots, x^{(n)}), (y^{(1)}, \dots, y^{(n)}) \in \mathbb{R}^n$  such that  $x^{(i)} \leq y^{(i)}$  for any  $i \in \{1, \dots, n\}$ . Thus, for  $\vec{u} = 1$  and any  $i \in \{1, \dots, n\}$ , we can

find  $t_i \geq 0$  such that  $x^{(i)} + t_i \vec{u} = y^{(i)}$ . From the SC-monotonicity of  $A$ , it follows that

$$\begin{aligned} & \left( A(y^{(1)}, \dots, y^{(n)}) - A(x^{(1)}, \dots, x^{(n)}) \right) \cdot \vec{u} \\ &= \left( A(y^{(1)}, \dots, y^{(n)}) - A(y^{(1)}, \dots, y^{(n-1)}, x^{(n)}) \right. \\ & \quad \left. + \dots + A(y^{(1)}, x^{(2)}, \dots, x^{(n)}) - A(x^{(1)}, \dots, x^{(n)}) \right) \cdot \vec{u} \\ & \geq 0, \end{aligned}$$

due to the linearity of the dot product.  $\square$

Additionally, MP-monotonicity and MC-monotonicity coincide if  $d = 1$  and reduce to the classical weak monotonicity in the setting of real numbers.

**Proposition 17.** *For any unidimensional function  $A : \mathbb{R}^n \rightarrow \mathbb{R}$ , the following results are equivalent:*

- (i)  *$A$  is weakly monotone in the sense of Definition 2.*
- (ii)  *$A$  is MP-monotone.*
- (iii)  *$A$  is MC-monotone.*

*Proof.* The proof is similar to that of Proposition 16.  $\square$

### 3.5. Further results

Another typical transformation when dealing with multidimensional data is that of translation. A function  $A$  is called translation equivariant if  $A(\mathbf{X} + \mathbf{t}) = A(\mathbf{X}) + \mathbf{t}$  for any  $\mathbf{t} \in \mathbb{R}^d$ . Note that MP-monotonicity obviously is a weaker condition than translation equivariance.

**Proposition 18.** *Any translation equivariant function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is MP-monotone.*

Note that SP-monotonicity w.r.t.  $d$  pairwise orthogonal directions  $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_d \in \mathbb{R}^d \setminus \{\mathbf{0}\}$  is enough to assure that SC-monotonicity holds (with respect to all directions). We refer to  $d$  pairwise orthogonal directions  $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_d \in \mathbb{R}^d \setminus \{\mathbf{0}\}$  as an orthogonal basis (of  $\mathbb{R}^d$ ) and, additionally, as an orthonormal basis (of  $\mathbb{R}^d$ ) in case all directions are of unit length. Note that an orthonormal basis of  $\mathbb{R}^d$  can be written in terms of an orthogonal matrix  $\mathbf{O}$  in which the columns of  $\mathbf{O}$  represent the basis vectors. We recall that, given

any direction  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ , it is always possible to construct an orthonormal basis  $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_d$  such that  $\vec{\mathbf{v}}_1 = \vec{\mathbf{u}}$ . A typical way of doing this is by means of the Gram–Schmidt process (see [30, 31] for the original works or [32] for a historical review).

**Proposition 19.** *For any function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  and any pairwise orthogonal directions  $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_d \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ , if  $A$  is  $\vec{\mathbf{v}}_j$ -SP-monotone for any  $j \in \{1, \dots, d\}$ , then it is SC-monotone.*

*Proof.* Assume without loss of generality that all  $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_d$  are unitary. Consider any direction  $\vec{\mathbf{u}} = \sum_{j=1}^d \lambda_j \vec{\mathbf{v}}_j$ , any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$ , any  $i \in \{1, \dots, n\}$  and any step size  $t \geq 0$ . We define  $\vec{\mathbf{w}}_j = \sum_{\ell=1}^j \lambda_\ell \vec{\mathbf{v}}_\ell$  for any  $j \in \{1, \dots, d\}$ . It holds that

$$\begin{aligned} & \left( A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)} + t\vec{\mathbf{u}}, \dots, \mathbf{x}^{(n)}) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)}, \dots, \mathbf{x}^{(n)}) \right) \cdot \vec{\mathbf{u}} \\ &= \left( A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)} + t\vec{\mathbf{w}}_d, \dots, \mathbf{x}^{(n)}) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)} + t\vec{\mathbf{w}}_{d-1}, \dots, \mathbf{x}^{(n)}) \right. \\ & \quad \left. + \dots + A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)} + t\vec{\mathbf{w}}_1, \dots, \mathbf{x}^{(n)}) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)}, \dots, \mathbf{x}^{(n)}) \right) \cdot \vec{\mathbf{u}} \\ &= \left( k_1 \operatorname{sgn}(\lambda_1) \vec{\mathbf{v}}_1 + \dots + k_d \operatorname{sgn}(\lambda_d) \vec{\mathbf{v}}_d \right) \cdot \vec{\mathbf{u}} \\ &\stackrel{(*)}{=} \sum_{j=1}^d k_j \operatorname{sgn}(\lambda_j) \lambda_j = \sum_{j=1}^d k_j |\lambda_j| \geq 0, \end{aligned}$$

since all  $k_j$  are nonnegative constants arising from the  $\vec{\mathbf{v}}_j$ -SP-monotonicity of  $A$  and in  $(*)$  we use that the dot product is invariant under orthogonal transformations.  $\square$

*Remark 20.* Note that the above result might not be true if the directions  $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_d$  are not pairwise orthogonal. For instance, consider  $A(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}) = (\operatorname{Med}(\mathbf{x}_1 - \mathbf{x}_2) + \operatorname{Med}(\mathbf{x}_2), \operatorname{Med}(\mathbf{x}_2))^T$ , which is SP-monotone w.r.t. the directions  $\vec{\mathbf{v}}_1 = (1, 0)^T$  and  $\vec{\mathbf{v}}_2 = (1, 1)^T$ . If we consider  $\mathbf{x}^{(1)} = (3, 4)^T$ ,  $\mathbf{x}^{(2)} = (6, 3)^T$ ,  $\mathbf{x}^{(3)} = (8, 5)^T$  and  $\vec{\mathbf{u}} = (-2, 1)^T$  and  $t = 1$ , then one can easily verify that  $A(\mathbf{x}^{(1)} + t\vec{\mathbf{u}}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}) = (7, 4)^T$  whereas  $A(\mathbf{x}^{(1)} + t\vec{\mathbf{u}}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}) = (8, 5)^T$ . Thus,  $A$  is not  $\vec{\mathbf{u}}$ -SC-monotone (and, thus, not SC-monotone).

The above remark serves as the source of inspiration for the following lemma.

**Lemma 21.** For any function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$ , any invertible matrix  $\mathbf{P} \in \mathbb{R}^{d \times d}$  and any direction  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ , if  $A$  is  $\vec{\mathbf{u}}$ -SP-monotone, then the function  $F : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  defined by  $F(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) = \mathbf{P} A(\mathbf{P}^{-1}\mathbf{x}^{(1)}, \dots, \mathbf{P}^{-1}\mathbf{x}^{(n)})$  is  $\mathbf{P}\vec{\mathbf{u}}$ -SP-monotone.

*Proof.* Consider any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$ , any  $i \in \{1, \dots, n\}$  and any step size  $t \geq 0$ . It holds that

$$\begin{aligned} & F(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)} + t\mathbf{P}\vec{\mathbf{u}}, \dots, \mathbf{x}^{(n)}) - F(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)}, \dots, \mathbf{x}^{(n)}) \\ &= \mathbf{P} A(\mathbf{P}^{-1}\mathbf{x}^{(1)}, \dots, \mathbf{P}^{-1}\mathbf{x}^{(i)} + t\mathbf{P}^{-1}\mathbf{P}\vec{\mathbf{u}}, \dots, \mathbf{P}^{-1}\mathbf{x}^{(n)}) \\ &\quad - \mathbf{P} A(\mathbf{P}^{-1}\mathbf{x}^{(1)}, \dots, \mathbf{P}^{-1}\mathbf{x}^{(i)}, \dots, \mathbf{P}^{-1}\mathbf{x}^{(n)}) \\ &= k\mathbf{P}\vec{\mathbf{u}}, \end{aligned}$$

for some  $k \geq 0$  due to the fact that  $A$  is  $\vec{\mathbf{u}}$ -SP-monotone.  $\square$

*Remark 22.* Note that the above result also holds for all other monotonicity properties.

Interestingly, any function defined as the componentwise extension<sup>6</sup> of some  $d$  monotone unidimensional functions is SP-monotone with respect to the directions given by the canonical basis. A more general result follows as a corollary by additionally considering Lemma 21.

**Theorem 23.** For any function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$ , it holds that  $A$  is  $\vec{\mathbf{e}}_j$ -SP-monotone for any  $j \in \{1, \dots, d\}$  (where  $\{\vec{\mathbf{e}}_j\}_{j=1}^d$  represents the canonical basis) if and only if  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})_j = A_j(\mathbf{x}_j^{(1)}, \dots, \mathbf{x}_j^{(n)})$  for some  $d$  monotone unidimensional functions  $A_j : \mathbb{R}^n \rightarrow \mathbb{R}$  (with  $j \in \{1, \dots, d\}$ ).

*Proof.* We take part of the proof of Theorem 5 in [29] and make use of the following characterization provided in [26] (Proposition 15):  $A$  is a componentwise extension of  $d$  monotone unidimensional functions if and only if for any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}), (\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(n)}) \in \mathbb{R}^{d \times n}$  and any  $j \in \{1, \dots, d\}$ , it holds that  $x_j^{(i)} \leq y_j^{(i)}$  for any  $i \in \{1, \dots, n\}$  implies  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})_j \leq A(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(n)})_j$ . Consider  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}), (\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(n)}) \in \mathbb{R}^{d \times n}$  and  $j \in$

---

<sup>6</sup>A function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is called the componentwise extension of a unidimensional function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  if  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})_j = f(\mathbf{x}_j^{(1)}, \dots, \mathbf{x}_j^{(n)})$ , for any  $j \in \{1, \dots, d\}$ . Similarly, we can consider a componentwise extension of  $d$  different unidimensional functions  $f_j$ , i.e.,  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})_j = f_j(\mathbf{x}_j^{(1)}, \dots, \mathbf{x}_j^{(n)})$ .

$\{1, \dots, d\}$  such that  $x_j^{(i)} \leq y_j^{(i)}$  for any  $i \in \{1, \dots, n\}$ . For any  $i \in \{1, \dots, n\}$ , we define  $z_j^{(i)} = x_j^{(i)}$  and  $z_\ell^{(i)} = y_\ell^{(i)}$  for any  $\ell \neq j$ . From the  $\vec{e}_j$ -SP-monotonicity, it follows that  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})_j = A(\mathbf{z}^{(1)}, \dots, \mathbf{z}^{(n)})_j$  and  $A(\mathbf{z}^{(1)}, \dots, \mathbf{z}^{(n)})_j \leq A(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(n)})_j$ . We combine both facts and apply the aforementioned characterization to conclude that  $A$  is a componentwise extension of  $d$  monotone unidimensional functions. The converse implication is straightforward.  $\square$

**Corollary 24.** *For any function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  and any linearly independent directions  $\vec{v}_1, \dots, \vec{v}_d \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ , it holds that  $A$  is  $\vec{v}_j$ -SP-monotone for any  $j \in \{1, \dots, d\}$  if and only if  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) = \mathbf{P}^{-1} \left( A_1((\mathbf{P}\mathbf{x}^{(1)})_1, \dots, (\mathbf{P}\mathbf{x}^{(n)})_1), \dots, A_d((\mathbf{P}\mathbf{x}^{(1)})_d, \dots, (\mathbf{P}\mathbf{x}^{(n)})_d) \right)^T$  for some  $d$  monotone unidimensional functions  $A_j : \mathbb{R}^n \rightarrow \mathbb{R}$  (with  $j \in \{1, \dots, d\}$ ) and  $\mathbf{P}$  being the matrix with  $\vec{v}_j$  as  $j$ -th column.*

Thus, two propositions follow as corollaries of the preceding results.

**Proposition 25.** *If a function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is such that  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})_j = A_j(\mathbf{x}_j^{(1)}, \dots, \mathbf{x}_j^{(n)})$  for some  $d$  monotone unidimensional functions  $A_j : \mathbb{R}^n \rightarrow \mathbb{R}$  (with  $j \in \{1, \dots, d\}$ ), then it is SC-monotone.*

**Proposition 26.** *If a function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is such that  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})_j = A_j(\mathbf{x}_j^{(1)}, \dots, \mathbf{x}_j^{(n)})$  for some  $d$  monotone unidimensional functions  $A_j : \mathbb{R}^n \rightarrow \mathbb{R}$  (with  $j \in \{1, \dots, d\}$ ) and it is rotation equivariant, then it is SP-monotone.*

The orthogonalization of multivariate location estimators has been studied in the field of applied statistics. For instance, the orthomedian [33] is obtained by averaging the componentwise median over all orthogonal transformations. In the following, we prove that any function obtained by averaging a componentwise extension of some  $d$  monotone unidimensional functions over all orthogonal transformations turns out to be SC-monotone.

**Proposition 27.** *If a function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is such that  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})_j = A_j(\mathbf{x}_j^{(1)}, \dots, \mathbf{x}_j^{(n)})$  for some  $d$  monotone unidimensional functions  $A_j : \mathbb{R}^n \rightarrow \mathbb{R}$  (with  $j \in \{1, \dots, d\}$ ), then the function  $F : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$ , defined as*

$$F(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) = \int_{\mathcal{O}} \mathbf{O}^T A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(n)}) d\mathbf{O},$$



where  $\mathcal{O}$  denotes the group of orthogonal  $d \times d$  matrices and  $\int \dots d\mathbf{O}$  refers to the unique Haar measure on  $\mathcal{O}$  with total mass 1, is SC-monotone.

*Proof.*  $F$  is orthogonal equivariant by construction, thus, according to Proposition 15, it suffices to prove that  $F$  is SC-monotone w.r.t. just one direction. For instance, we consider the direction  $\vec{\mathbf{e}}_1$  of the canonical basis. We partition  $\mathcal{O}$  in two sets

$$\begin{aligned}\mathcal{O}^+ &= \{\mathbf{O} \in \mathcal{O} \mid \vec{\mathbf{e}}_1 \cdot \mathbf{O}\vec{\mathbf{e}}_1 \geq 0\}, \\ \mathcal{O}^- &= \{\mathbf{O} \in \mathcal{O} \mid \vec{\mathbf{e}}_1 \cdot \mathbf{O}\vec{\mathbf{e}}_1 < 0\}.\end{aligned}$$

Due to the orthogonal invariance and symmetry of the dot product, it holds that  $\mathbf{O} \in \mathcal{O}^+$  if and only if  $\mathbf{O}^T \in \mathcal{O}^+$ . Similarly, it holds that  $\mathbf{O} \in \mathcal{O}^-$  if and only if  $\mathbf{O}^T \in \mathcal{O}^-$ .

Now consider any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$ , any  $i \in \{1, \dots, n\}$  and any step size  $t \geq 0$ . It holds that

$$\begin{aligned}& F(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)} + t\vec{\mathbf{e}}_1, \dots, \mathbf{x}^{(n)}) - F(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)}, \dots, \mathbf{x}^{(n)}) \\&= \int_{\mathcal{O}} \mathbf{O}^T A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)} + t\mathbf{O}\vec{\mathbf{e}}_1, \dots, \mathbf{O}\mathbf{x}^{(n)}) d\mathbf{O} \\&\quad - \int_{\mathcal{O}} \mathbf{O}^T A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)}, \dots, \mathbf{O}\mathbf{x}^{(n)}) d\mathbf{O} \\&= \int_{\mathcal{O}} \mathbf{O}^T (A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)} + t\mathbf{O}\vec{\mathbf{e}}_1, \dots, \mathbf{O}\mathbf{x}^{(n)}) - A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)}, \dots, \mathbf{O}\mathbf{x}^{(n)})) d\mathbf{O} \\&= \int_{\mathcal{O}^+} \mathbf{O}^T (A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)} + t\mathbf{O}\vec{\mathbf{e}}_1, \dots, \mathbf{O}\mathbf{x}^{(n)}) - A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)}, \dots, \mathbf{O}\mathbf{x}^{(n)})) d\mathbf{O} \\&\quad + \int_{\mathcal{O}^-} \mathbf{O}^T (A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)} + t\mathbf{O}\vec{\mathbf{e}}_1, \dots, \mathbf{O}\mathbf{x}^{(n)}) - A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)}, \dots, \mathbf{O}\mathbf{x}^{(n)})) d\mathbf{O}.\end{aligned}$$

On the one hand, from Theorem 23, we know that, for any  $\mathbf{O} \in \mathcal{O}^+$ ,

$$(A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)} + t\mathbf{O}\vec{\mathbf{e}}_1, \dots, \mathbf{O}\mathbf{x}^{(n)}) - A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)}, \dots, \mathbf{O}\mathbf{x}^{(n)})) \cdot \vec{\mathbf{e}}_1 \geq 0,$$

and, therefore, since  $\vec{\mathbf{e}}_1 \cdot \mathbf{O}\vec{\mathbf{e}}_1 \geq 0$ ,

$$\mathbf{O}^T (A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)} + t\mathbf{O}\vec{\mathbf{e}}_1, \dots, \mathbf{O}\mathbf{x}^{(n)}) - A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)}, \dots, \mathbf{O}\mathbf{x}^{(n)})) \cdot \vec{\mathbf{e}}_1 \geq 0.$$

Thus, it holds that

$$\int_{\mathcal{O}^+} \mathbf{O}^T (A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)} + t\mathbf{O}\vec{\mathbf{e}}_1, \dots, \mathbf{O}\mathbf{x}^{(n)}) - A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)}, \dots, \mathbf{O}\mathbf{x}^{(n)})) d\mathbf{O} \cdot \vec{\mathbf{e}}_1 \geq 0.$$

On the other hand, also from Theorem 23, we know that, for any  $\mathbf{O} \in \mathcal{O}^-$ ,

$$\left( A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)} + t\mathbf{O}\vec{\mathbf{e}}_1, \dots, \mathbf{O}\mathbf{x}^{(n)}) - A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)}, \dots, \mathbf{O}\mathbf{x}^{(n)}) \right) \cdot \vec{\mathbf{e}}_1 \leq 0,$$

and, therefore, since  $\vec{\mathbf{e}}_1 \cdot \mathbf{O}\vec{\mathbf{e}}_1 < 0$ ,

$$\mathbf{O}^T \left( A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)} + t\mathbf{O}\vec{\mathbf{e}}_1, \dots, \mathbf{O}\mathbf{x}^{(n)}) - A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)}, \dots, \mathbf{O}\mathbf{x}^{(n)}) \right) \cdot \vec{\mathbf{e}}_1 \geq 0.$$

Thus, it holds that

$$\int_{\mathcal{O}^-} \mathbf{O}^T \left( A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)} + t\mathbf{O}\vec{\mathbf{e}}_1, \dots, \mathbf{O}\mathbf{x}^{(n)}) - A(\mathbf{O}\mathbf{x}^{(1)}, \dots, \mathbf{O}\mathbf{x}^{(i)}, \dots, \mathbf{O}\mathbf{x}^{(n)}) \right) d\mathbf{O} \cdot \vec{\mathbf{e}}_1 \geq 0.$$

The result finally follows from the linearity of the dot product.  $\square$

The following result states that an idempotent function that is SP-monotone with respect to the directions given by the canonical basis and one additional direction must necessarily be SP-monotone (with respect to all directions). The only functions of this kind are weighted centroids. Actually, the result is also easily proved to hold if SP-monotonicity holds with respect to an orthogonal basis and one additional direction.

**Theorem 28.** *Let  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  be an idempotent function and  $d \geq 2$ . Then the following conditions are equivalent:*

- (i) *A is SP-monotone w.r.t. the directions  $\{\vec{\mathbf{e}}_j\}_{j=1}^d$  of the canonical basis as well as some other direction  $\vec{\mathbf{v}} = \sum_{i=1}^d v_i \vec{\mathbf{e}}_i$  with  $v_i \neq 0$  for at least two  $i \in \{1, \dots, d\}$ ;*
- (ii) *There exist  $k_1, \dots, k_n \geq 0$  with  $\sum_{i=1}^n k_i = 1$  such that, for any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$ , it holds that  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) = \sum_{i=1}^n k_i \mathbf{x}^{(i)}$ ;*
- (iii) *A is SP-monotone (w.r.t. all directions).*

*Proof.* (i)  $\Rightarrow$  (ii). We adapt the proof of Theorem 5 in [29] to prove the result. From Theorem 23, it follows that  $A$  is a componentwise extension of some  $d$  monotone unidimensional functions  $A_j : \mathbb{R}^n \rightarrow \mathbb{R}$  (with  $j \in \{1, \dots, d\}$ ).

Consider  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$ . We define  $(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(n)}) \in \mathbb{R}^{d \times n}$  such that  $\mathbf{y}^{(i_0)} = \mathbf{x}^{(i_0)} + t\vec{\mathbf{v}}$  for some step size  $t \geq 0$  and some fixed  $i_0 \in \{1, \dots, n\}$ , whereas  $\mathbf{y}^{(i)} = \mathbf{x}^{(i)}$  for all other  $i \neq i_0$ . From the  $\vec{\mathbf{v}}$ -SP-monotonicity of  $A$ , we know that  $A(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(n)}) = A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) + k\vec{\mathbf{v}}$  for some  $k \geq 0$  dependent on  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})$ ,  $t$  and  $i_0$ . Since  $A$  is a componentwise extension

of  $d$  monotone unidimensional functions, it follows that  $A(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(n)})_j - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})_j = A_j(y_j^{(1)}, \dots, y_j^{(n)}) - A_j(x_j^{(1)}, \dots, x_j^{(n)}) = k$  for any fixed  $j \in \{1, \dots, d\}$ , thus  $k$  being dependent on at most  $(x_j^{(1)}, \dots, x_j^{(n)})$ ,  $t$  and  $i_0$ . Since this holds for any  $j$  and  $\vec{v}$  is a combination of at least two directions in the canonical basis, we conclude that  $k$  must be independent of  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})$ . We thus use the notation  $k_{i_0, t}$  to indicate that it depends on the choice of  $i_0$  and  $t$ .

We now prove that  $k_{i_0, t} = t k_{i_0, 1}$ . We distinguish two cases depending on whether  $t$  is a rational number or an irrational number. If  $t$  is a rational number, then it can be expressed as  $t = \frac{p}{q}$ . Since  $k_{i_0, t}$  does not depend on  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})$ , we may apply the above result  $q$  times with a step size of  $\frac{1}{q}$  (instead of a single time with a step size of 1), and obtain  $k_{i_0, 1} = q k_{i_0, \frac{1}{q}}$ . Similarly, we may apply the above result  $p$  times with a step size of  $\frac{1}{q}$  (instead of a single time with a step size of  $\frac{p}{q}$ ), and obtain  $k_{i_0, t} = p k_{i_0, \frac{1}{q}}$ . We conclude that  $k_{i_0, t} = \frac{p}{q} k_{i_0, 1} = t k_{i_0, 1}$ . If  $t$  is an irrational number, consider a decreasing sequence  $\{a_\ell\}_{\ell \in \mathbb{N}}$  of rational numbers with limit  $t$  and an increasing sequence  $\{b_\ell\}_{\ell \in \mathbb{N}}$  of rational numbers with limit  $t$ . The result follows from the facts that  $k_{i_0, t}$  is obviously increasing with respect to  $t$  and that  $k_{i_0, a_\ell} = a_\ell k_{i_0, 1}$  and  $k_{i_0, b_\ell} = b_\ell k_{i_0, 1}$  for any  $\ell \in \mathbb{N}$ .

Finally, we conclude that, for any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$ , it holds that  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) = A(\mathbf{0}, \dots, \mathbf{0}) + \left( \sum_{i=1}^n k_i x_1^{(i)}, \dots, \sum_{i=1}^n k_i x_d^{(i)} \right) = \sum_{i=1}^n k_i \mathbf{x}^{(i)}$  for some  $k_1, \dots, k_n \geq 0$ , due to the fact that

$$\begin{aligned} (\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) &= (\mathbf{0}, \dots, \mathbf{0}) + \left( (x_1^{(1)}, 0, \dots, 0)^T, \mathbf{0}, \dots, \mathbf{0} \right) \\ &\quad + \left( (0, x_2^{(1)}, 0, \dots, 0)^T, \mathbf{0}, \dots, \mathbf{0} \right) \\ &\quad + \dots + \left( \mathbf{0}, \dots, \mathbf{0}, (0, \dots, 0, x_d^{(n)})^T \right) \end{aligned}$$

and  $A(\mathbf{0}, \dots, \mathbf{0}) = \mathbf{0}$  if  $A$  is idempotent. Also from the idempotence of  $A$ , we conclude that  $\sum_{i=1}^n k_i = 1$ .

As it is easily seen that (ii)  $\Rightarrow$  (iii) and trivially (iii)  $\Rightarrow$  (i), the proof is complete.  $\square$

It trivially follows that the only symmetric<sup>7</sup>, idempotent and SP-monotone function is the centroid.

---

<sup>7</sup>A function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is symmetric if  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) = A(\mathbf{x}^{(\sigma(1))}, \dots, \mathbf{x}^{(\sigma(n))})$  for any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$  and any permutation  $\sigma$  of  $\{1, \dots, n\}$ .

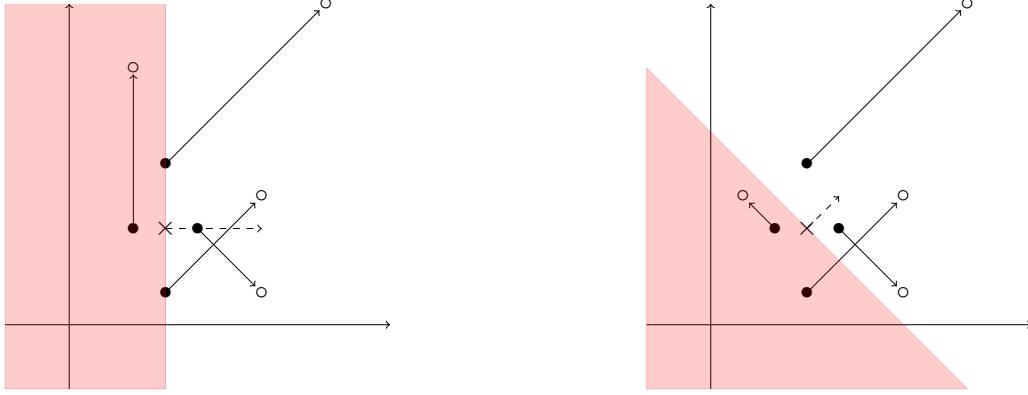


Figure 3: Graphical comparison of componentwise monotonicity w.r.t. the first component (left side) and ultramonotonicity (right side). The four black circles represent four initial data points, whereas the white circles represent the new positions of the initial data points. The point marked with the symbol  $\times$  represents the initial aggregate and the dashed arrow represents the direction in which we are considering the monotonicity. The red area indicates the area to which the aggregate must *not* go.

## 4. Componentwise monotonicity and ultramonotonicity

### 4.1. Componentwise monotonicity

Another natural type of monotonicity is what is referred to as componentwise monotonicity in [26].

**Definition 29.** A function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is called componentwisely monotone if, for any fixed  $j \in \{1, \dots, d\}$  and any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}), (\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(n)}) \in \mathbb{R}^{d \times n}$  satisfying that  $x_j^{(i)} \leq y_j^{(i)}$  for any  $i \in \{1, \dots, n\}$ , it holds that  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})_j \leq A(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(n)})_j$ .

Note that there is no condition on how  $x_{j'}^{(i)}$  and  $y_{j'}^{(i)}$  should relate if  $j' \neq j$ . Explicitly, this means that, as long as the  $j$ -th component of the data points increases, the  $j$ -th component of the aggregate should also increase (whatever happens in the other components). For simplicity, we just focus on componentwise monotonicity w.r.t. the first component in the illustration of this property in the left side of Figure 3.

An important result concerning componentwise monotonicity is that componentwise monotonicity of any function  $A$  is equivalent to  $A$  being a componentwise extension of some  $d$  monotone unidimensional functions [26].

**Theorem 30.** A function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is componentwisely monotone if and only if  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})_j = A_j(\mathbf{x}_j^{(1)}, \dots, \mathbf{x}_j^{(n)})$  for some  $d$  monotone unidimensional functions  $A_j : \mathbb{R}^n \rightarrow \mathbb{R}$  (with  $j \in \{1, \dots, d\}$ ).

*Remark 31.* The above result combined with Theorem 23 implies that componentwise monotonicity is equivalent to  $\vec{\mathbf{e}}_j$ -SP-monotonicity for any direction  $\vec{\mathbf{e}}_j$  in the canonical basis.

#### 4.2. Ultramonotonicity

The notion of componentwise monotonicity can be generalized to any possible direction. Hereinafter, we shall refer to such a property as ultramonotonicity. For a graphical interpretation of this property, see the right side of Figure 3.

**Definition 32.** For a given direction  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ , a function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is called  $\vec{\mathbf{u}}$ -ultramotone if, for any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$ , any non-negative step sizes  $t_1, \dots, t_n \in \mathbb{R}$  and any family of directions  $\vec{\mathbf{u}}_1, \dots, \vec{\mathbf{u}}_n \in \mathbb{R}^d$  such that  $\vec{\mathbf{u}} \cdot \vec{\mathbf{u}}_i \geq 0$  for any  $i \in \{1, \dots, n\}$ , it holds that

$$\left( A(\mathbf{x}^{(1)} + t_1 \vec{\mathbf{u}}_1, \dots, \mathbf{x}^{(n)} + t_n \vec{\mathbf{u}}_n) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \right) \cdot \vec{\mathbf{u}} \geq 0.$$

**Definition 33.** A function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is called ultramonotone if it is  $\vec{\mathbf{u}}$ -ultramotone with respect to all directions  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ .

Note that the magnitude and the sign of  $\vec{\mathbf{u}}$  do not matter for defining  $\vec{\mathbf{u}}$ -ultramotonicity.

**Proposition 34.** For any given direction  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ , a function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is  $\vec{\mathbf{u}}$ -ultramotone if and only if it is  $(c\vec{\mathbf{u}})$ -ultramotone for any  $c \neq 0$ .

*Proof.* We only prove that if  $A$  is  $\vec{\mathbf{u}}$ -ultramotone, then  $A$  is  $(c\vec{\mathbf{u}})$ -ultramotone for any  $c \neq 0$  (the other implication is evident). Consider  $c > 0$ . Thus, for any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$ , any non-negative step sizes  $t_1, \dots, t_n \in \mathbb{R}$  and any family of directions  $\vec{\mathbf{u}}_1, \dots, \vec{\mathbf{u}}_n \in \mathbb{R}^d$  such that  $c\vec{\mathbf{u}} \cdot \vec{\mathbf{u}}_i \geq 0$  for any  $i \in \{1, \dots, n\}$ , it holds that

$$\begin{aligned} & \left( A(\mathbf{x}^{(1)} + t_1 c\vec{\mathbf{u}}_1, \dots, \mathbf{x}^{(n)} + t_n c\vec{\mathbf{u}}_n) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \right) \cdot c\vec{\mathbf{u}} \\ &= c \left( A(\mathbf{x}^{(1)} + t_1 \vec{\mathbf{u}}_1, \dots, \mathbf{x}^{(n)} + t_n \vec{\mathbf{u}}_n) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \right) \cdot \vec{\mathbf{u}} \\ &\geq 0. \end{aligned}$$

Similarly, consider now  $c < 0$ , any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$ , any non-negative step sizes  $t_1, \dots, t_n \in \mathbb{R}$  and any family of directions  $\vec{\mathbf{u}}_1, \dots, \vec{\mathbf{u}}_n \in \mathbb{R}^d$  such that  $c\vec{\mathbf{u}} \cdot \vec{\mathbf{u}}_i \geq 0$  for any  $i \in \{1, \dots, n\}$ . It follows that  $|c|\vec{\mathbf{u}} \cdot |c|(-\vec{\mathbf{u}}_i) \geq 0$  for any  $i \in \{1, \dots, n\}$ . We fix  $(\mathbf{x}^{(1)'}, \dots, \mathbf{x}^{(n)'}) = (\mathbf{x}^{(1)} - t_1|c|\vec{\mathbf{u}}_1, \dots, \mathbf{x}^{(n)} - t_n|c|\vec{\mathbf{u}}_n)$ . Due to the  $(|c|\vec{\mathbf{u}})$ -ultramonicity of  $A$ , it holds that

$$\left( A \left( \mathbf{x}^{(1)'} + t_1|c|(-\vec{\mathbf{u}}_1), \dots, \mathbf{x}^{(n)'} + t_n|c|(-\vec{\mathbf{u}}_n) \right) - A \left( \mathbf{x}^{(1)'}, \dots, \mathbf{x}^{(n)'} \right) \right) \cdot |c|\vec{\mathbf{u}} \geq 0,$$

or, equivalently, after substituting  $\mathbf{x}^{(i)'}$ ,

$$\left( A \left( \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)} \right) - A \left( \mathbf{x}^{(1)} - t_1|c|\vec{\mathbf{u}}_1, \dots, \mathbf{x}^{(n)} - t_n|c|\vec{\mathbf{u}}_n \right) \right) \cdot |c|\vec{\mathbf{u}} \geq 0.$$

We finally conclude that

$$\left( A \left( \mathbf{x}^{(1)} + t_1c\vec{\mathbf{u}}_1, \dots, \mathbf{x}^{(n)} + t_nc\vec{\mathbf{u}}_n \right) - A \left( \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)} \right) \right) \cdot c\vec{\mathbf{u}} \geq 0.$$

□

Ultramonotonicity trivially implies componentwise monotonicity.

**Proposition 35.** *Any ultramonotone function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is componentwisely monotone.*

*Proof.* Follows from the fact that componentwise monotonicity can be understood as  $\vec{\mathbf{u}}$ -ultramonicity with respect to all directions  $\vec{\mathbf{u}}$  in the canonical basis. □

Furthermore, any function that is rotation equivariant needs only to be  $\vec{\mathbf{u}}$ -ultramotone with respect to one direction  $\vec{\mathbf{u}} \in \mathbb{R}^d$  in order to be ultramonotone.

**Proposition 36.** *A rotation equivariant function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is ultramonotone if and only if it is  $\vec{\mathbf{u}}$ -ultramotone with respect to at least one direction  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ .*

*Proof.* Consider  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$  and assume that  $A$  is  $\vec{\mathbf{u}}$ -ultramotone. Consider any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$ , any non-negative step sizes  $t_1, \dots, t_n \in \mathbb{R}$  and any family of directions  $\vec{\mathbf{u}}_1, \dots, \vec{\mathbf{u}}_n \in \mathbb{R}^d$  such that  $\vec{\mathbf{u}} \cdot \vec{\mathbf{u}}_i \geq 0$  for any

$i \in \{1, \dots, n\}$ . Consider  $\vec{\mathbf{v}} = \mathbf{R}\vec{\mathbf{u}}$  for some rotation matrix  $\mathbf{R}$ , then it holds that

$$\begin{aligned}
& \left( A(\mathbf{x}^{(1)} + t_1 \vec{\mathbf{u}}_1, \dots, \mathbf{x}^{(n)} + t_n \vec{\mathbf{u}}_n) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \right) \cdot \vec{\mathbf{v}} \\
&= \mathbf{R} \left( A(\mathbf{R}^{-1} \mathbf{x}^{(1)} + t_1 \mathbf{R}^{-1} \vec{\mathbf{u}}_1, \dots, \mathbf{R}^{-1} \mathbf{x}^{(n)} + t_n \mathbf{R}^{-1} \vec{\mathbf{u}}_n) \right. \\
&\quad \left. - A(\mathbf{R}^{-1} \mathbf{x}^{(1)}, \dots, \mathbf{R}^{-1} \mathbf{x}^{(n)}) \right) \cdot \mathbf{R} \vec{\mathbf{u}} \\
&= \left( A(\mathbf{R}^{-1} \mathbf{x}^{(1)} + t_1 \mathbf{R}^{-1} \vec{\mathbf{u}}_1, \dots, \mathbf{R}^{-1} \mathbf{x}^{(n)} + t_n \mathbf{R}^{-1} \vec{\mathbf{u}}_n) \right. \\
&\quad \left. - A(\mathbf{R}^{-1} \mathbf{x}^{(1)}, \dots, \mathbf{R}^{-1} \mathbf{x}^{(n)}) \right) \cdot \vec{\mathbf{u}} \geq 0,
\end{aligned}$$

due to the fact that the dot product is rotation invariant (and, additionally  $\vec{\mathbf{u}} \cdot \vec{\mathbf{u}}_i \geq 0$  if and only if  $\vec{\mathbf{v}} \cdot \mathbf{R}^{-1} \vec{\mathbf{u}}_i \geq 0$ ). Thus,  $A$  is  $\vec{\mathbf{v}}$ -ultramotone.  $\square$

**Corollary 37.** *A rotation equivariant function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is ultramotone if and only if it is componentwisely monotone.*

#### 4.3. Ultramonotonicity and idempotence

It is widely known that, for a monotone unidimensional function, internality (being bounded from below by the minimum and from above by the maximum) is equivalent to idempotence. Note that there are two main extensions of the notion of internality to multiple dimensions [26]: BB (bounding box)-internality and CH (convex hull)-internality. The former requires the aggregate  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})$  to belong to the bounding box of  $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}$ , i.e., the set

$$\left[ \min_{i=1}^n x_1^{(i)}, \max_{i=1}^n x_1^{(i)} \right] \times \dots \times \left[ \min_{i=1}^n x_d^{(i)}, \max_{i=1}^n x_d^{(i)} \right].$$

The latter requires the aggregate  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})$  to belong to the convex hull of  $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}$ , i.e., the set

$$\left\{ \mathbf{x} = \sum_{i=1}^n \lambda_i \mathbf{x}^{(i)} \in \mathbb{R}^d \mid \left( \sum_{i=1}^n \lambda_i = 1 \right) \wedge \left( (\forall i \in \{1, \dots, n\}) (\lambda_i \geq 0) \right) \right\}.$$

For a componentwisely monotone function, idempotence is equivalent to BB-internality, whereas, for an ultramotone function, idempotence is equivalent to CH-internality.

**Proposition 38.** *A componentwisely monotone function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is idempotent if and only if it is BB-internal.*

*Proof.* By Theorem 30, it follows that  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})_j = A_j(\mathbf{x}_j^{(1)}, \dots, \mathbf{x}_j^{(n)})$  for some  $d$  monotone unidimensional functions  $A_j : \mathbb{R}^n \rightarrow \mathbb{R}$  (with  $j \in \{1, \dots, d\}$ ). Note that  $A$  is idempotent if and only if all  $A_j$  are idempotent, and that  $A$  is BB-internal if and only if all  $A_j$  are bounded from below by the minimum and from above by the maximum. Finally, the result follows from the fact that, for monotone unidimensional functions, it is known that idempotence is equivalent to being bounded from below by the min and from above by the max [1].  $\square$

**Proposition 39.** *An ultramonotone function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  is idempotent if and only if it is CH-internal.*

*Proof.* It is straightforward to see that, if  $A$  is CH-internal, then it is idempotent. We prove that, if  $A$  is idempotent, then it is CH-internal by reductio ad absurdum. Consider  $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}$  such that  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})$  does not belong to the convex hull of  $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}$ . Since the convex hull is a non-empty, closed and convex set, there exists a unique point  $\mathbf{x}^{(0)}$  in the convex hull that minimizes the distance to  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})$  and, additionally the halfspace formed by all points  $\mathbf{y}$  such that  $(\mathbf{y} - \mathbf{x}^{(0)}) \cdot (A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) - \mathbf{x}^{(0)}) \leq 0$  contains the convex hull of  $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}$  and does not contain  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})$ . Since  $A$  is idempotent, it follows that  $A(\mathbf{x}^{(0)}, \dots, \mathbf{x}^{(0)}) = \mathbf{x}^{(0)}$ . Consider  $\vec{\mathbf{u}} = A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) - \mathbf{x}^{(0)}$ . It follows that  $\vec{\mathbf{u}}_i = (\mathbf{x}^{(i)} - \mathbf{x}^{(0)})$  are such that  $\vec{\mathbf{u}}_i \cdot \vec{\mathbf{u}} \geq 0$ , thus, since  $A$  is ultramonotone, it follows that

$$(A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) - A(\mathbf{x}^{(0)}, \dots, \mathbf{x}^{(0)})) \cdot \vec{\mathbf{u}} \geq 0,$$

a contradiction with the halfspace  $(\mathbf{y} - \mathbf{x}^{(0)}) \cdot (A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) - \mathbf{x}^{(0)}) \leq 0$  not containing  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})$ .  $\square$

Note that this result complies with intuition since for rotation equivariant functions both componentwise monotonicity and ultramonotonicity coincide and both CH-internality and BB-internality coincide [26]. Later on in this section, it will follow that ultramonotone and idempotent functions also reduce to componentwise extensions of a single weighted arithmetic mean, which are always CH-internal.



#### 4.4. Relations to the other monotonicity properties

An important matter of study is the relation of  $\vec{\mathbf{u}}$ -ultramonicity w.r.t. all the monotonicity properties defined in Section 3.

**Proposition 40.** *For any function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$ , any direction  $\vec{\mathbf{u}} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$  and any orthogonal basis  $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_d \in \mathbb{R}^d$  such that  $\vec{\mathbf{v}}_1 = \vec{\mathbf{u}}$ , the following results hold:*

- (i) *If  $A$  is  $\vec{\mathbf{v}}_j$ -ultramonic for any  $j \in \{1, \dots, d\}$ , then it is  $\vec{\mathbf{u}}$ -SP-monotone.*
- (ii) *If  $A$  is  $\vec{\mathbf{v}}_j$ -SP-monotone for any  $j \in \{1, \dots, d\}$ , then it is  $\vec{\mathbf{u}}$ -ultra-monotone.*
- (iii) *If  $A$  is  $\vec{\mathbf{u}}$ -ultramonic, then it is  $\vec{\mathbf{u}}$ -SC-monotone.*
- (iv) *If  $A$  is  $\vec{\mathbf{v}}_j$ -ultramonic for any  $j \in \{1, \dots, d\}$ , then it is SC-monotone.*

*Proof.* (i) Consider any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$ , any  $i \in \{1, \dots, n\}$  and any step size  $t \geq 0$ . Since  $A$  is  $\vec{\mathbf{v}}_j$ -ultramonic for any  $j \in \{2, \dots, d\}$ , it follows that

$$\left( A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)} + t\vec{\mathbf{u}}, \dots, \mathbf{x}^{(n)}) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \right) \cdot \vec{\mathbf{v}}_j \geq 0.$$

Since  $A$  is also  $(-\vec{\mathbf{v}}_j)$ -ultramonic, it follows that:

$$\left( A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)} - t\vec{\mathbf{u}}, \dots, \mathbf{x}^{(n)}) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \right) \cdot (-\vec{\mathbf{v}}_j) \geq 0.$$

We conclude that:

$$\left( A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)} + t\vec{\mathbf{u}}, \dots, \mathbf{x}^{(n)}) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \right) \cdot \vec{\mathbf{v}}_j = 0.$$

Thus,  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)} + t\vec{\mathbf{u}}, \dots, \mathbf{x}^{(n)}) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)}, \dots, \mathbf{x}^{(n)})$  is orthogonal to any  $\vec{\mathbf{v}}_j$  with  $j \in \{2, \dots, d\}$ . Therefore, it holds that:

$$A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)} + t\vec{\mathbf{u}}, \dots, \mathbf{x}^{(n)}) - A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(i)}, \dots, \mathbf{x}^{(n)}) = k \vec{\mathbf{u}},$$

for some constant  $k$ . It follows from the  $\vec{\mathbf{u}}$ -ultramonicity of  $A$  that  $k \geq 0$ .

(ii) Consider any  $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) \in \mathbb{R}^{d \times n}$ , any non-negative step sizes  $t_1, \dots, t_n \in \mathbb{R}$  and any family of directions  $\vec{\mathbf{u}}_1, \dots, \vec{\mathbf{u}}_n \in \mathbb{R}^d \setminus \{\mathbf{0}\}$  such that

$\vec{\mathbf{u}} \cdot \vec{\mathbf{u}}_i \geq 0$  for any  $i \in \{1, \dots, n\}$ . Since  $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_d$  is an orthogonal basis, for any  $\vec{\mathbf{u}}_i$ , we can write  $t_i \vec{\mathbf{u}}_i = \sum_{j=1}^d \lambda_{i,j} \vec{\mathbf{v}}_j$  (note that  $\lambda_{i,1} \geq 0$  because  $\vec{\mathbf{v}}_1 \cdot \vec{\mathbf{u}}_i = \vec{\mathbf{u}} \cdot \vec{\mathbf{u}}_i \geq 0$ ). Decompose  $\left( A \left( \mathbf{x}^{(1)} + t_1 \vec{\mathbf{u}}_1, \dots, \mathbf{x}^{(n)} + t_n \vec{\mathbf{u}}_n \right) - A \left( \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)} \right) \right)$  in  $n \times d$  movements corresponding to the  $d$  movements  $\lambda_{i,j} \vec{\mathbf{v}}_j$  in each of the  $n$  directions  $\vec{\mathbf{u}}_i$ . Since  $\lambda_{i,1} \geq 0$  for any  $i \in \{1, \dots, n\}$ , we can apply  $(\pm \vec{\mathbf{v}}_j)$ -SP-monotonicity  $n \times d$  times and conclude that:

$$\begin{aligned} & \left( A \left( \mathbf{x}^{(1)} + t_1 \vec{\mathbf{u}}_1, \dots, \mathbf{x}^{(n)} + t_n \vec{\mathbf{u}}_n \right) - A \left( \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)} \right) \right) \cdot \vec{\mathbf{u}} \\ &= \left( A \left( \mathbf{x}^{(1)} + t_1 \vec{\mathbf{u}}_1, \dots, \mathbf{x}^{(n)} + t_n \vec{\mathbf{u}}_n \right) - A \left( \mathbf{x}^{(1)} + t_1 \vec{\mathbf{u}}_1, \dots, \mathbf{x}^{(n)} + t_n \vec{\mathbf{u}}_n - \lambda_{n,d} \vec{\mathbf{v}}_d \right) \right. \\ & \quad \left. + \dots + A \left( \mathbf{x}^{(1)} + \lambda_{1,1} \vec{\mathbf{v}}_1, \dots, \mathbf{x}^{(n)} \right) - A \left( \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)} \right) \right) \cdot \vec{\mathbf{u}} \geq 0, \end{aligned}$$

due to the linearity of the dot product.

(iii) Follows straightforwardly by considering the step sizes  $t_1, \dots, t_n \in \mathbb{R}$  such that  $t_i \neq 0$  (where  $i$  is the index of the point to be moved) and  $t_{i'} = 0$  for any  $i' \in \{1, \dots, n\} \setminus \{i\}$  and  $\vec{\mathbf{u}}_i = \vec{\mathbf{u}}$ .

(iv) From (i), we obtain that  $A$  is  $\vec{\mathbf{v}}_j$ -SP-monotone for any  $j \in \{1, \dots, d\}$ . The result then follows from Proposition 19.  $\square$

The following theorem follows as a corollary from the proposition above. We illustrate the result in Figure 4.

**Theorem 41.** *For any function  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$ , the following results hold:*

- (i)  *$A$  is ultramonotone if and only if it is SP-monotone.*
- (ii) *If  $A$  is componentwisely monotone, then it is SC-monotone.*

Since SP-monotonicity is equivalent to ultramonotonicity, we conclude from Proposition 16 that both ultramonotonicity and componentwise monotonicity are equivalent in case  $d = 1$  and reduce to the classical monotonicity in the setting of real numbers.

**Proposition 42.** *For any unidimensional function  $A : \mathbb{R}^n \rightarrow \mathbb{R}$ , the following results are equivalent:*

- (i)  *$A$  is monotone in the sense of Definition 1.*
- (ii)  *$A$  is ultramonotone.*
- (iii)  *$A$  is componentwisely monotone.*

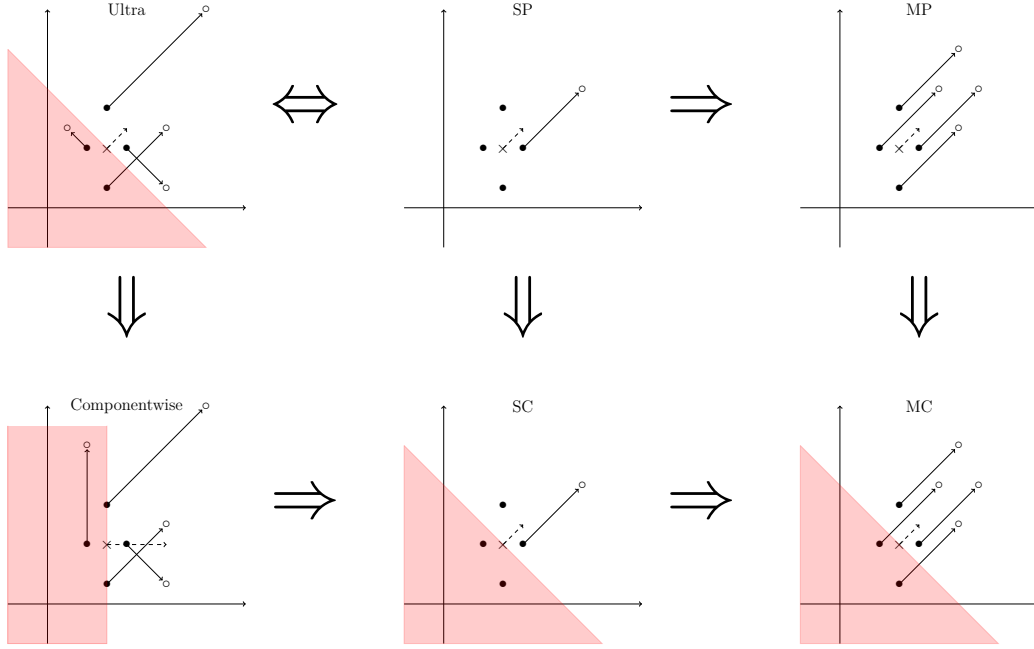


Figure 4: Graphical comparison of the different monotonicity properties defined in this section. The four black circles represent four initial data points, whereas the white circles represent the new positions of (some of) the initial data points. The point marked with the symbol  $\times$  represents the initial aggregate and the dashed arrow represents the direction in which we are considering the monotonicity. The red area indicates the area to which the aggregate must not go.

## 5. Examples

In the following, we provide some noteworthy examples of functions  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  and discuss which monotonicity properties they satisfy. A summary of these examples is anticipated in Table 1.

*Centroid.* The centroid is defined as the point that minimizes the sum of squared Euclidean distances to the points to be aggregated. Equivalently, the centroid is characterized as the componentwise extension of the arithmetic mean. Since it is a componentwise extension of  $d$  monotone unidimensional functions and it is rotation equivariant, we know from Proposition 26 that the centroid is SP-monotone. Furthermore, from Theorem 14, it follows that the centroid is MP-, SC-, and MC-monotone, and, from Theorem 41, it follows that the centroid is componentwisely monotone and ultramonotone. Note that any weighted centroid (the componentwise extension of a single weighted

Function	SP=Ultra	MP	SC	MC	Componentwise
Centroid	✓	✓	✓	✓	✓
Componentwise median	×	✓	✓	✓	✓
Spatial median	×	✓	×	✓	×
Orthomedian	×	✓	✓	✓	×
Euclidean center	×	✓	×	✓	×
Componentwise extension of some $d$ monotone unidimensional functions	?	?	✓	✓	✓
Orthogonalized version of componentwise extension of some $d$ monotone unidimensional functions	?	?	✓	✓	?
Translation equivariant function	?	✓	?	✓	?

Table 1: Noteworthy examples of functions  $A : \mathbb{R}^{d \times n} \rightarrow \mathbb{R}^d$  and summary of the monotonicity properties they satisfy. The symbols ✓, × and ? respectively represent that all, none or some (but not all) functions within the family satisfy the corresponding monotonicity property.

arithmetic mean) satisfies all these monotonicity properties. Actually, as stated in Theorem 28, weighted centroids are the only idempotent functions fulfilling SP-monotonicity (and, thus, ultramonotonicity).

*Componentwise median.* The componentwise median is the componentwise extension of the median. It satisfies the conditions of Propositions 18 (see [26]) and 25, thus, it follows that the componentwise median is MP- and SC-monotone. Furthermore, from Theorem 14, it follows that the componentwise median is MC-monotone. As it is componentwisely defined, it is evident that the componentwise median is componentwisely monotone (see Theorem 30). Finally, we can see that the componentwise median is not SP-monotone (and, thus, not ultramonotone) with the following example. Consider  $\mathbf{x}^{(1)} = (1, 1)^T$ ,  $\mathbf{x}^{(2)} = (3, 1)^T$  and  $\mathbf{x}^{(3)} = (2, 0)^T$ . The componentwise median of  $\mathbf{x}^{(1)}$ ,  $\mathbf{x}^{(2)}$  and  $\mathbf{x}^{(3)}$  is given by  $A(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}) = (2, 1)^T$ , whereas for  $\mathbf{x}^{(3)'} = \mathbf{x}^{(3)} + (1, 1)^T$ , we get  $A(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)'}) = (3, 1)^T$ . Thus, there does not exist  $k \geq 0$  such that  $(A(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)'}) - A(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)})) = (1, 0)^T$  equals  $k(1, 1)^T$ .

*Spatial median.* The spatial median is defined as the point that minimizes the sum of Euclidean distances to the points to be aggregated. It satisfies the conditions of Proposition 18 (see [26]), thus, it follows that the spatial

median is MP-monotone. Furthermore, from Theorem 14, it follows that the spatial median is MC-monotone. Finally, we can see that the spatial median is not SC-monotone (and, thus, neither SP-monotone, nor componentwisely monotone, nor ultramonotone) from the example in Figure 5.

*Orthomedian.* The orthomedian is defined as the average of the componentwise median over all orthogonal transformations. It satisfies the conditions of Proposition 27, thus, it follows that the orthomedian is SC-monotone. Additionally, it satisfies the conditions of Proposition 18 (see [26]), thus, it follows that the orthomedian is MP-monotone. Furthermore, from Theorem 14, it follows that the orthomedian is MC-monotone. Finally, we can see that the orthomedian is not componentwisely monotone (and, thus, neither SP-monotone, nor ultramonotone) from the following example. Consider  $\mathbf{x}^{(1)} = (-1, 0)^T$ ,  $\mathbf{x}^{(2)} = (1, 0)^T$  and  $\mathbf{x}^{(3)} = (0, 1)^T$ . The orthomedian of  $\mathbf{x}^{(1)}$ ,  $\mathbf{x}^{(2)}$  and  $\mathbf{x}^{(3)}$  is given by  $A(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}) = (0, 0.5)^T$ , whereas for  $\mathbf{x}^{(1)'} = \mathbf{x}^{(1)} + (1, 0)^T$ , we get  $A(\mathbf{x}^{(1)'}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}) = (0.25, 0.25)^T$ .

*Euclidean center.* The Euclidean center is defined as the point that minimizes the maximum Euclidean distance to the points to be aggregated. It satisfies the conditions of Proposition 18 (see [26]), thus, it follows that the Euclidean center is MP-monotone. Furthermore, from Theorem 14, it follows that the Euclidean center is MC-monotone. Finally, we can see that the Euclidean center is not SC-monotone (and, thus, neither SP-monotone, nor componentwisely monotone, nor ultramonotone) from the example in Figure 6.

*Componentwise cubic mean.* The cubic mean  $A : \mathbb{R}^n \rightarrow \mathbb{R}$  is the function defined as  $A(\mathbf{x}) = \sqrt[3]{\frac{1}{n} \sum_{i=1}^n x_i^3}$ . The componentwise extension of the cubic mean satisfies the conditions of Proposition 25, thus, it follows that the componentwise cubic mean is SC-monotone. Furthermore, from Theorem 14, it follows that the componentwise cubic mean is MC-monotone. As it is componentwisely defined, it is evident that the componentwise cubic mean is componentwisely monotone (see Theorem 30). Finally, we can see that the componentwise cubic mean is not MP-monotone (and, thus, neither SP-monotone, nor ultramonotone) with the following example. Consider  $\mathbf{x}^{(1)} = (-2, 0)^T$ ,  $\mathbf{x}^{(2)} = (1, 0)^T$  and  $\mathbf{x}^{(3)} = (-0.5, 1)^T$ . The componentwise cubic mean of  $\mathbf{x}^{(1)}$ ,  $\mathbf{x}^{(2)}$  and  $\mathbf{x}^{(3)}$  is given by  $A(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}) \approx (-1.3342008, 0.6933613)^T$ , whereas after moving all points in the direction  $\vec{\mathbf{u}} = (1, 1)^T$  and step size  $t = 1$ , we get  $A(\mathbf{x}^{(1)} + \vec{\mathbf{u}}, \mathbf{x}^{(2)} + \vec{\mathbf{u}}, \mathbf{x}^{(3)} + \vec{\mathbf{u}}) \approx (1.334201, 1.493802)^T$ . Thus, there does

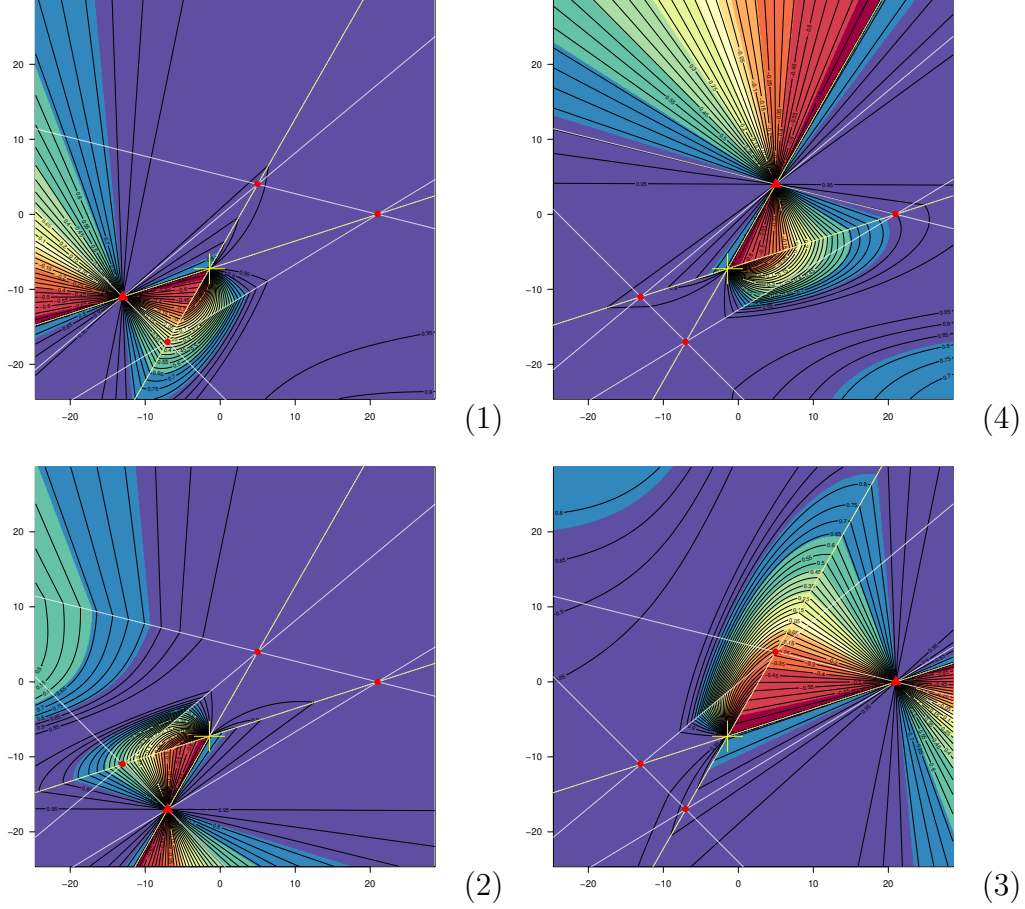


Figure 5: Consider  $\mathbf{x}^{(1)} = (-13, -11)^T$ ,  $\mathbf{x}^{(2)} = (-7, -17)^T$ ,  $\mathbf{x}^{(3)} = (21, 0)^T$ ,  $\mathbf{x}^{(4)} = (5, 4)^T$ . The corresponding spatial median  $\mathbf{y}$  (marked with a yellow +) is located near  $(-1.43, -7.26)^T$ . In each subplot ( $i$ ) ( $i \in \{1, 2, 3, 4\}$ ), we depict the cosine of the angle between  $\mathbf{y}' - \mathbf{y}$  and  $(x, y)^T - \mathbf{x}^{(i)}$  for all possible  $(x, y)$  on a rectangular grid spanning  $[-24, 29] \times [-24, 29]$ , where  $\mathbf{y}'$  is the spatial median of  $(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}, \mathbf{x}^{(4)})$  but with the corresponding  $\mathbf{x}^{(i)}$  substituted with  $(x, y)^T$ . Red-colored regions mark destinations where the spatial median actually moves in a conflicting direction (cosine  $< 0$ ) to the point being relocated.

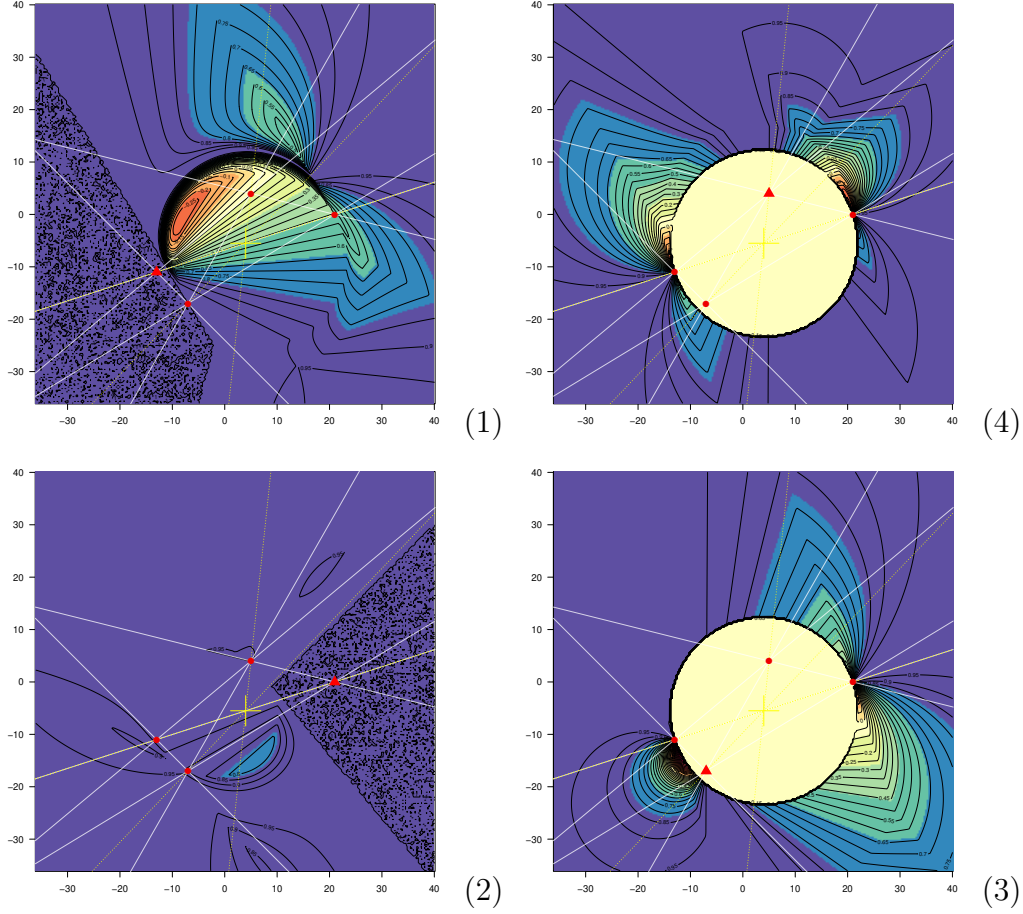


Figure 6: Consider  $\mathbf{x}^{(1)} = (-13, -11)^T$ ,  $\mathbf{x}^{(2)} = (-7, -17)^T$ ,  $\mathbf{x}^{(3)} = (21, 0)^T$ ,  $\mathbf{x}^{(4)} = (5, 4)^T$ . The corresponding Euclidean center  $\mathbf{y}$  (marked with a yellow +) is located at  $(4, -5.5)^T$ . In each subplot  $(i)$  ( $i \in \{1, 2, 3, 4\}$ ), we depict the cosine of the angle between  $\mathbf{y}' - \mathbf{y}$  and  $(x, y)^T - \mathbf{x}^{(i)}$  for all possible  $(x, y)$  on a rectangular grid spanning  $[-34, 40] \times [-34, 40]$ , where  $\mathbf{y}'$  is the Euclidean center of  $(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}, \mathbf{x}^{(4)})$  but with the corresponding  $\mathbf{x}^{(i)}$  substituted with  $(x, y)^T$ . Red-colored regions mark destinations where the Euclidean center actually moves in a conflicting direction (cosine  $< 0$ ) to the point being relocated.

not exist  $k \geq 0$  such that  $(A(\mathbf{x}^{(1)} + \vec{\mathbf{u}}, \mathbf{x}^{(2)} + \vec{\mathbf{u}}, \mathbf{x}^{(3)} + \vec{\mathbf{u}}) - A(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)})) \approx (2.6684016, 0.8004403)^T$  equals  $k(1, 1)^T$ .

*Convex combination of the componentwise cubic mean and the spatial median.* We have seen that the spatial median is not SC-monotone and that the componentwise cubic mean is not MP-monotone. One could think of a convex combination of these two functions for finding an example of an MC-monotone function that is neither SP-, nor MP-, nor SC-monotone. First, we see that it is MC-monotone by noting that both the spatial median and the componentwise cubic mean are MC-monotone, and, thus, also any convex combination of them. We prove that this function is not MP-monotone (and, thus, neither SP-monotone, nor ultramonotone) with the following example. Consider  $\mathbf{x}^{(1)} = (-0.2, 1.2)^T$ ,  $\mathbf{x}^{(2)} = (0.8, -1.3)^T$  and  $\mathbf{x}^{(3)} = (1.5, -0.3)^T$ . The chosen convex combination (0.8 for the spatial median and 0.2 for the componentwise cubic mean) of  $\mathbf{x}^{(1)}$ ,  $\mathbf{x}^{(2)}$  and  $\mathbf{x}^{(3)}$  is given by  $A(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}) \approx (1.123976, -0.4035746)^T$ , whereas after moving all points in the direction  $\vec{\mathbf{u}} = (0.35, 0.12)^T$  and step size  $t = 1$ , we get  $A(\mathbf{x}^{(1)} + \vec{\mathbf{u}}, \mathbf{x}^{(2)} + \vec{\mathbf{u}}, \mathbf{x}^{(3)} + \vec{\mathbf{u}}) \approx (1.461759, -0.07761333)^T$ . Thus, there does not exist  $k \geq 0$  such that  $(A(\mathbf{x}^{(1)} + \vec{\mathbf{u}}, \mathbf{x}^{(2)} + \vec{\mathbf{u}}, \mathbf{x}^{(3)} + \vec{\mathbf{u}}) - A(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)})) \approx (0.3377828, 0.3259613)^T$  equals  $k(0.35, 0.12)^T$ . Similarly, we prove that this function is not SC-monotone (and, thus, neither SP-monotone, nor componentwisely monotone, nor ultramonotone) with the following example. Consider  $\mathbf{x}^{(1)} = (-0.3, 0.7)^T$ ,  $\mathbf{x}^{(2)} = (1.1, 0.2)^T$  and  $\mathbf{x}^{(3)} = (0.5, 1.2)^T$ . The chosen convex combination (again, 0.8 for the spatial median and 0.2 for the componentwise cubic mean) of  $\mathbf{x}^{(1)}$ ,  $\mathbf{x}^{(2)}$  and  $\mathbf{x}^{(3)}$  is given by  $A(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}) \approx (0.4886453, 0.8882840)^T$ , whereas after moving  $\mathbf{x}^{(1)}$  in the direction  $\vec{\mathbf{u}} = (0.35, 0.075)^T$  and step size  $t = 1$ , we get  $A(\mathbf{x}^{(1)} + \vec{\mathbf{u}}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}) \approx (0.4894801, 0.8808175)^T$ . Thus, we conclude that  $(A(\mathbf{x}^{(1)} + \vec{\mathbf{u}}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}) - A(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)})) \cdot \vec{\mathbf{u}} \approx -0.000267812345 < 0$ .

*A non-monotone function.* As an example that does not fulfill any of the above monotonicity properties, consider  $A(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) = \max_{i=1}^n \max_{j=1}^d x_j^{(i)}$ . Assume  $\mathbf{x}^{(1)} = (3, 2)^T$  and  $\mathbf{x}^{(2)} = (5, 0)^T$ . Applying this function to  $\mathbf{x}^{(1)}$  and  $\mathbf{x}^{(2)}$  we obtain  $A(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) = (5, 5)^T$ , whereas after moving all points in the direction  $\vec{\mathbf{u}} = (-1, 2)^T$  and step size  $t = 1$ , we get  $A(\mathbf{x}^{(1)} + \vec{\mathbf{u}}, \mathbf{x}^{(2)} + \vec{\mathbf{u}}) = (4, 4)^T$ . Thus, we conclude that  $(A(\mathbf{x}^{(1)} + \vec{\mathbf{u}}, \mathbf{x}^{(2)} + \vec{\mathbf{u}}) - A(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})) \cdot \vec{\mathbf{u}} = -1 < 0$ . Obviously, this function is not idempotent, however, the function  $F$  defined as  $F(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) = \mathbf{x}^{(1)}$ , if  $\mathbf{x}^{(1)} = \dots = \mathbf{x}^{(n)}$ , and,  $F(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}) =$



$\max_{i=1}^n \max_{j=1}^d x_j^{(i)}$ , otherwise, serves as an example of an idempotent function that still fails to fulfill any of the monotonicity properties above.

## 6. Discussion

Different monotonicity properties for the framework of the aggregation of multidimensional data have been studied. It is worth noting that, prior to the current work, with the goal of allowing more functions to fulfill the classical definition of an aggregation function, a further concept related to the property of weak monotonicity (see Definition 2) has been studied under the name of directional monotonicity [34]. Functions increasing in one variable and decreasing in another one (e.g., ‘fuzzy implication’ functions) are examples of functions justifying the analysis of directional monotonicity. However, from the perspective of this contribution, this concept is not really semantically valid. The directional monotonicity of a function  $A : \mathbb{R}^n \rightarrow \mathbb{R}$  treats the input to  $A$  as a single point in an  $n$ -dimensional space that can be moved to an arbitrary position. Formally, however, from the geometrical perspective, there are only two directions (‘left’ and ‘right’) to which we can move any of the  $n$  points on the real line.

In this paper, we have pointed out that the classical monotonicity based on the preservation of the product order makes no sense in the setting of multidimensional data if not combined with some desirable behaviour with regard to rotations. Furthermore, under rotation equivariance, this monotonicity property coincides with both SP-monotonicity and ultramonotonicity. Regrettably, if we combine this with the fact that the only idempotent functions that are SP-monotone are componentwise extensions of a single weighted arithmetic mean, we restrict the aggregation of multidimensional data to this latter type of functions. However, thorough studies on many different aggregation methods for multidimensional data that are not extensions of a single weighted arithmetic mean have been addressed. For instance, we highlight the Euclidean center [35] and different extensions of the unidimensional median [27]. Aggregation functions of the latter type include the spatial median [36], Oja’s simplex median [37], Tukey’s halfspace median [38], the simplicial depth median [39] and the convex hull stripping median [40]. Note that all of the previously mentioned extensions of the median are rotation and translation equivariant (and all of them but the spatial median actually are affine equivariant), thus not being monotone in

the classical sense of aggregation theory. Fortunately, different monotonicity properties are proved to hold for these functions.

We conclude by noting that many of the presented results may be generalized to arbitrary vector spaces, e.g., polynomials, complex numbers, etc.

*Acknowledgements.* Raúl Pérez-Fernández acknowledges the support of the Research Foundation of Flanders (FWO17/PDO/160) and the Spanish MINECO (TIN2017-87600-P). Marek Gagolewski's research was supported by the Czech Science Foundation (project no. 18-06915S).

## References

- [1] G. Beliakov, A. Pradera, T. Calvo, Aggregation Functions: A Guide for Practitioners, Vol. 221 of Studies in Fuzziness and Soft Computing, Springer, Berlin, Heidelberg, 2007.
- [2] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap, Aggregation Functions, Cambridge University Press, Cambridge, 2009.
- [3] P. Bullen, Handbook of Means and Their Inequalities, Springer Science+Business Media, Dordrecht, 2003.
- [4] G. Beliakov, H. Bustince, T. Calvo, A Practical Guide to Averaging Functions, Springer, 2016.
- [5] M. Gagolewski, Data Fusion: Theory, Methods, and Applications, Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland, 2015.
- [6] D. Dubois, H. Prade, On the use of aggregation operations in information fusion processes, Fuzzy Sets and Systems 142 (2004) 143–161.
- [7] M. Komorníková, R. Mesiar, Aggregation functions on bounded partially ordered sets and their classification, Fuzzy Sets and Systems 175 (2011) 48–56.
- [8] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap, Aggregation functions: Means, Information Sciences 181 (2011) 1–22.
- [9] A. L. Cauchy, Cours d'analyse de l'Ecole Royale Polytechnique, Vol. 1 of Analyse algébrique, Debure, Paris, 1821.

- [10] T. Wilkin, G. Beliakov, Weakly monotonic averaging functions, *International Journal of Intelligent Systems* 30 (2) (2015) 144–169.
- [11] G. Beliakov, T. Calvo, T. Wilkin, On the weak monotonicity of Gini means and other mixture functions, *Information Sciences* 300 (2015) 70–84.
- [12] G. Beliakov, J. Špirková, Weak monotonicity of Lehmer and Gini means, *Fuzzy Sets and Systems* 299 (2016) 26–40.
- [13] J. C. Borda, *Mémoire sur les Élections au Scrutin*, Histoire de l’Académie Royale des Sciences, Paris, 1781.
- [14] M. Condorcet, *Essai sur l’Application de l’Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix*, De l’Imprimerie Royale, Paris, 1785.
- [15] H. P. Young, Condorcet’s theory of voting, *American Political Science Review* 82 (4) (1988) 1231–1244.
- [16] B. Monjardet, An axiomatic theory of tournament aggregation, *Mathematics of Operations Research* 3 (4) (1978) 334–351.
- [17] J. P. Barthélemy, B. Monjardet, The median procedure in cluster analysis and social choice theory, *Mathematical Social Sciences* 1 (1981) 235–267.
- [18] Y. Wakabayashi, The complexity of computing medians of relations, *Resenhas* 3 (3) (1998) 323–349.
- [19] T. Kohonen, Median strings, *Pattern Recognition Letters* 3 (1985) 309–313.
- [20] R. R. Yager, Toward a general theory of information aggregation, *Information Sciences* 68 (1993) 191–206.
- [21] R. Pérez-Fernández, M. Rademaker, B. De Baets, Monometrics and their role in the rationalisation of ranking rules, *Information Fusion* 34 (2017) 16–27.

- [22] R. Pérez-Fernández, B. De Baets, On the role of monometrics in penalty-based data aggregation, *IEEE Transactions on Fuzzy Systems*, in press, DOI: 10.1109/TFUZZ.2018.2880716.
- [23] R. R. Yager, A. Rybalov, Understanding the median as a fusion operator, *International Journal of General Systems* 26 (1997) 239–263.
- [24] T. Calvo, G. Beliakov, Aggregation functions based on penalties, *Fuzzy Sets and Systems* 161 (2010) 1420–1436.
- [25] H. Bustince, G. Beliakov, G. P. Dimuro, B. Bedregal, R. Mesiar, On the definition of penalty functions in data aggregation, *Fuzzy Sets and Systems* 323 (2017) 1–18.
- [26] M. Gagolewski, Penalty-based aggregation of multidimensional data, *Fuzzy Sets and Systems* 325 (2017) 4–20.
- [27] C. G. Small, A survey of multidimensional medians, *International Statistical Review* 58 (3) (1990) 263–277.
- [28] P. J. Rousseeuw, A. Struyf, Computation of robust statistics: depth, median, and related measures, in: J. E. Goodman, J. O’Rourke (Eds.), *The Handbook of Discrete and Computational Geometry*, Chapman & Hall/CRC, Boca Raton, 2004, pp. 1279–1292.
- [29] M. Gagolewski, R. Pérez-Fernández, B. De Baets, An inherent difficulty in the aggregation of multidimensional data, *IEEE Transactions on Fuzzy Systems*, in press, DOI: 10.1109/TFUZZ.2019.2908135.
- [30] J. P. Gram, Ueber die Entwicklung reeller Functionen in Reihen mittelst der Methode der kleinsten Quadrate, *Journal für die reine und angewandte Mathematik* 94 (1883) 41–73.
- [31] E. Schmidt, Zur Theorie der linearen und nichtlinearen Integralgleichungen, *Mathematische Annalen* 63 (4) (1907) 433–476.
- [32] S. J. Leon, Å. Björck, W. Gander, Gram-Schmidt orthogonalization: 100 years and more, *Numerical Linear Algebra with Applications* 20 (3) (2013) 492–532.
- [33] R. Grübel, Orthogonalization of multivariate location estimators: The orthomedian, *Annals of Statistics* 24 (4) (1996) 1457–1473.

- [34] H. Bustince, J. Fernandez, A. Kolesárová, R. Mesiar, Directional monotonicity of fusion functions, *European Journal of Operational Research* 244 (2015) 300–308.
- [35] J. J. Sylvester, A question in the geometry of situation, *The Quarterly Journal of Mathematics* 1 (1857) 79.
- [36] A. Weber, *Ueber den Standort der Industrien*, Mohr Siebeck Verlag, Tübingen, 1909.
- [37] H. Oja, Descriptive statistics for multivariate distributions, *Statistics and Probability Letters* 1 (1983) 327–332.
- [38] J. W. Tukey, Mathematics and the picturing of data, in: *Proceedings of the International Congress of Mathematicians*, Vancouver, 1975, pp. 523–531.
- [39] R. Y. Liu, On a notion of data depth based on random simplices, *The Annals of Statistics* 18 (1) (1990) 405–414.
- [40] W. F. Eddy, Convex hull peeling, in: *Proceedings of the COMPSTAT Symposium*, Toulouse, 1982, pp. 42–47.

Please cite this paper as:

R. Pérez-Fernández, B. De Baets, M. Gagolewski, A taxonomy of monotonicity properties for the aggregation of multidimensional data, *Information Fusion* 52, 322–334, 2019, doi:10.1016/j.inffus.2019.05.006