JMLR: Workshop and Conference Proceedings vol 35:1–28, 2014

Uniqueness of Ordinal Embedding

Matthäus Kleindessner KLEINDESSNER@INFORMATIK.UNI-HAMBURG.DE
Ulrike von Luxburg LUXBURG@INFORMATIK.UNI-HAMBURG.DE
Department of Computer Science, University of Hamburg, Germany

Abstract

Ordinal embedding refers to the following problem: all we know about an unknown set of points $x_1, \ldots, x_n \in \mathbb{R}^d$ are ordinal constraints of the form $\|x_i - x_j\| < \|x_k - x_l\|$; the task is to construct a realization $y_1, \ldots, y_n \in \mathbb{R}^d$ that preserves these ordinal constraints. It has been conjectured since the 1960ies that upon knowledge of all ordinal constraints a large but finite set of points can be approximately reconstructed up to a similarity transformation. The main result of our paper is a formal proof of this conjecture.

Keywords: non-metric multidimensional scaling, monotone mapping, isotonic mapping

1. Introduction

We consider the problem of ordinal embedding, also called ordinal scaling, non-metric multidimensional scaling, monotonic embedding, or isotonic embedding. Consider a set $x_1, \ldots, x_n$ in some metric space $(\mathcal{X}, \mathrm{dist})$, but assume that the distances between these points are unknown. All we get to see are ordinal relationships, namely whether $\mathrm{dist}(x_i, x_j) < \mathrm{dist}(x_k, x_l)$ or vice versa. The goal of ordinal embedding is to construct $y_1, \ldots, y_n \in \mathbb{R}^d$ such that all ordinal constraints are preserved (throughout the paper, $\|\cdot\|$ denotes the Euclidean norm):

$$\mathrm{dist}(x_i, x_j) < \mathrm{dist}(x_k, x_l) \Rightarrow \|y_i - y_j\| < \|y_k - y_l\|.$$

The problem of ordinal embedding has first been studied in the psychometric community by Shepard (1962a,b) and Kruskal (1964a,b), see also the monograph Borg and Groenen (2005). Lately it has drawn quite some attention in the machine learning community (Quist and Yona, 2004; Rosales and Fung, 2006; Agarwal et al., 2007; Shaw and Jebara, 2009; McFee and Lanckriet, 2009; Jamieson and Nowak, 2011a; McFee and Lanckriet, 2011; Tamuz et al., 2011; Ailon, 2012), also in its special case of ranking (Ouyang and Gray, 2008; McFee and Lanckriet, 2010; Jamieson and Nowak, 2011b; Lan et al., 2012; Wauthier et al., 2013). Even though ordinal embedding dates back to the 1960ies and is widely used in practice, surprisingly little is known about its theoretical properties. Particularly striking, one of the most elementary properties, namely the uniqueness of ordinal embeddings, has never been established in a finite sample setting. It is widely believed that, upon knowledge of all ordinal relationships, a point set $x_1, \ldots, x_n \in \mathbb{R}^d$ can be approximately reconstructed up to a similarity transformation if n is "large enough" (p. 294 of Shepard, 1966; Section 2.2 of Borg and Groenen, 2005; Section 4.13.2 of Dattorro, 2005). Numerous simulation experiments have been published as supporting evidence (Shepard, 1966; Young, 1970; Sherman, 1972). The main result of our paper is a formal proof that this uniqueness conjecture is indeed true: Consider a sequence of points $(x_n)_{n \in \mathbb{N}}$ that are dense in some "nice" set $K \subseteq \mathbb{R}^d$. Let $y_1^n, \ldots, y_n^n$

be any ordinal embedding of $x_1, \dots, x_n$. Then, as $n \to \infty$, the set of embedded points always converges to the set of original points, up to similarity transformations such as rotations, translations, rescalings, or reflections. This even holds if we only know about "local ordinal relationships", that is distance comparisons between points in small subregions.

Our proofs are elementary in the sense that we do not apply any heavy mathematical machinery. However, details are delicate and require a careful treatment.

2. Setup, definitions and notation

We start this section with the definition of the two central notions in our paper, ordinal embeddings and isotonic functions. We will see below that these two notions are closely related.

Definition 1 (Ordinal embedding) *Consider two sets* $\mathcal{X}_n = \{x_1, \dots, x_n\} \subseteq \mathbb{R}^d$ *and* $\mathcal{Y}_n = \{y_1, \dots, y_n\} \subseteq \mathbb{R}^d$. $\mathcal{Y}_n$ *is an* ordinal embedding *of* $\mathcal{X}_n$ *if for all* $1 \leq i, j, k, l \leq n$,

$$\|x_i - x_j\| < \|x_k - x_l\| \Rightarrow \|y_i - y_j\| < \|y_k - y_l\|. \tag{1}$$

$\mathcal{Y}_n$ *is a* weak ordinal embedding *of* $\mathcal{X}_n$ *if* (1) *holds for all* $1 \leq i, j, k, l \leq n$ *with* $i = k$. $\mathcal{Y}_n$ *is a* strong ordinal embedding *of* $\mathcal{X}_n$ *if* (1) *holds for all* $1 \leq i, j, k, l \leq n$*, and additionally* $\|x_i - x_j\| = \|x_k - x_l\| \Rightarrow \|y_i - y_j\| = \|y_k - y_l\|$ *for all* $1 \leq i, j, k, l \leq n$.

Definition 2 (Isotonic functions) *Let* $\Omega \subseteq \mathbb{R}^d$ *and* $f : \Omega \to \mathbb{R}^d$ *be an arbitrary function.* f *is a* similarity *if there is* $\lambda > 0$ *such that for all* $x, y \in \Omega$ *we have* $\|f(x) - f(y)\| = \lambda \|x - y\|$. f *is* isotonic *or an* isotony *if for all* $x, y, z, w \in \Omega$,

$$\|x - y\| < \|z - w\| \Rightarrow \|f(x) - f(y)\| < \|f(z) - f(w)\|.$$

f *is* weakly isotonic *if this property only holds for* $x, y, z, w \in \Omega$ *with* $x = z$. f *is* strongly isotonic *if it is isotonic and additionally satisfies* $\|x - y\| = \|z - w\| \Rightarrow \|f(x) - f(y)\| = \|f(z) - f(w)\|$ *for all* $x, y, z, w \in \Omega$. *We say that* f *is* locally *a similarity / (weakly / strongly) isotonic if for each point* $x \in \Omega$ *there exists a neighborhood* $U(x)$ *in* Ω *such that* $f|_{U(x)}$ *has the corresponding property. If we want to emphasize that a function* $f : \Omega \to \mathbb{R}^d$ *has a property not only locally but on all of* Ω*, we sometimes say that* f *is* globally *a similarity / (weakly / strongly) isotonic.*

Let us mention some obvious but important observations. Similarities $f : \mathbb{R}^d \to \mathbb{R}^d$ are nothing else than the well-known similarity transformations, given by $f(x) = \lambda O x + b$ for some orthogonal matrix O and an offset $b \in \mathbb{R}^d$. For general Ω, they are simply given by the restrictions of similarity transformations to Ω (see Lemma A in Appendix A). Obviously, we have

$$\text{similarity} \Rightarrow \text{strongly isotonic} \Rightarrow \text{isotonic} \Rightarrow \text{weakly isotonic},$$

but for general Ω none of the converses are true. Any weakly isotonic function is injective. If f is a similarity or a strong isotony, so is f^{-1}, but this does not necessarily hold for isotonies. A composition of similarities / (weak / strong) isotonies is again a similarity / (weak / strong) isotony.

Obviously, $y_1, \dots, y_n$ is a (weak / strong) ordinal embedding of $x_1, \dots, x_n$ if and only if the mapping $f : \{x_1, \dots, x_n\} \to \{y_1, \dots, y_n\}$ given by $f(x_i) = y_i$ is (weakly / strongly) isotonic. The uniqueness question for ordinal embedding can thus be formalized as follows: if f is a (weakly / strongly) isotonic mapping between two finite point sets, can it be approximated by a similarity? It is well-known that any strongly isotonic function $f : \mathbb{R}^d \to \mathbb{R}^d$ defined on the full domain $\mathbb{R}^d$ is a similarity transformation. One can see this by exploiting properties of sphere-preserving mappings in Euclidean geometry (see McKemie and Väisälä (1999) and also the argumentation in Shepard, 1966), by an elegant argument related to positive definite functions (Schoenberg, 1938), and also by the Beckman-Quarles theorem (Beckman and Quarles, 1953). The key question of this paper is in what sense such a property already holds for functions defined on a finite set.

Let us conclude this section with some standard notation for the rest of the paper. For any subset $A \subseteq \mathbb{R}^d$ we denote its linear hull by $[A] = \{\sum_{i=1}^n \lambda_i a_i : n \in \mathbb{N}, a_i \in A, \lambda_i \in \mathbb{R}\}$ and its affine hull by $\mathcal{H}(A) = \{\sum_{i=1}^n \lambda_i a_i : n \in \mathbb{N}, a_i \in A, \lambda_i \in \mathbb{R}, \sum_{i=1}^n \lambda_i = 1\}$. For $z \in \mathbb{R}^d$ and $r > 0$ the open ball with center z and radius r is $U_r(z) = \{x \in \mathbb{R}^d : \|x - z\| < r\}$ and the closed ball is $\overline{U_r(z)} = \{x \in \mathbb{R}^d : \|x - z\| \leq r\}$. For a vector-valued function $f : X \to \mathbb{R}^d$ and $j = 1, \ldots, d$ we write f^j for the jth component of f. For two functions $f : X_1 \to \mathbb{R}^d$ and $g : X_2 \to \mathbb{R}^d$ and an arbitrary subset $X \subseteq X_1 \cap X_2$ we denote the supremum norm between f and g on X by $\|f - g\|_{\infty(X)} = \sup_{x \in X} \|f(x) - g(x)\|$. At some points we will speak of a cross-polytope. By this we mean the image $T(C)$ of the d-dimensional standard cross-polytope C, which is given by the convex hull of all permutations of $(\pm 1/0/0/\ldots/0) \in \mathbb{R}^d$, under some similarity transformation T.

3. Main results

In this section we present our main results. The proofs of the theorems are deferred to Sections 4 and 5. Our key question is to what extent an isotonic function f is uniquely determined, up to a similarity transformation. Our first result concerns the infinite case. We show that if f is defined on a dense subset of some "nice" set G and f is locally isotonic, then it is actually a similarity.

Theorem 3 (Isotonic on a dense set implies similarity) *Let $G \subseteq \mathbb{R}^d$ be an open and connected domain and $\Omega \subseteq G$ a dense subset. Let $f : \Omega \to \mathbb{R}^d$ be a locally isotonic function. Then there exists a unique extension of f to a similarity transformation $F : \mathbb{R}^d \to \mathbb{R}^d$.*

The next theorem deals with the finite case and is the main result of this paper. We consider $\mathcal{X}_n = \{x_1, \ldots, x_n\} \subseteq \mathbb{R}^d$ and an isotonic mapping $\varphi_n : \mathcal{X}_n \to \varphi_n(\mathcal{X}_n)$ — hence $\varphi_n(\mathcal{X}_n) = \{\varphi_n(x_1), \ldots, \varphi_n(x_n)\}$ is an ordinal embedding of $\mathcal{X}_n$. We prove that φ_n can be approximated by a similarity transformation, up to arbitrary precision as $n \to \infty$.

Theorem 4 (Isotonic on a finite set implies approximate similarity)

1. *Global Version: Let $K = \overline{U_r(z)} \subseteq \mathbb{R}^d$ be a closed and bounded ball (for some arbitrary $r > 0$, $z \in \mathbb{R}^d$). Let $(x_n)_{n \in \mathbb{N}}$ be a sequence of points $x_n \in K$ such that $\{x_n : n \in \mathbb{N}\}$ is dense in K. Let $0 < R < \infty$ and $(\varphi_n)_{n \in \mathbb{N}}$ be a sequence of isotonic functions $\varphi_n : \{x_1, \ldots, x_n\} \to U_R(0) \subseteq \mathbb{R}^d$. Then there exists a sequence $(S_n)_{n \in \mathbb{N}}$ of similarity transformations $S_n : \mathbb{R}^d \to \mathbb{R}^d$ such that*

$$\|S_n - \varphi_n\|_{\infty(\{x_1, \ldots, x_n\})} \to 0 \quad \text{as } n \to \infty. \tag{2}$$

2. *Local Version: More generally, let $K = \bigcup_{i=1}^k K_i \subseteq \mathbb{R}^d$ be a finite union of closed and bounded balls such that $\bigcup_{i=1}^k K_i^\circ$ is connected. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence of points $x_n \in K$ such that*

$\{x_n : n \in \mathbb{N}\}$ is dense in K. Let $0 < R < \infty$ and $(\varphi_n)_{n\in\mathbb{N}}$ be a sequence of functions $\varphi_n : \{x_1, \ldots, x_n\} \to U_R(0) \subseteq \mathbb{R}^d$ such that

$$\forall i \in \{1, \ldots, k\} : \varphi_n|_{\{x_1,\ldots,x_n\}\cap K_i} \text{ is isotonic.}$$

Then there exists a sequence $(S_n)_{n\in\mathbb{N}}$ of similarity transformations $S_n : \mathbb{R}^d \to \mathbb{R}^d$ with (2).

Our proofs show that we can replace the set K in Part 1 of Theorem 4 by a cross-polytope or any convex set "between a cross-polytope and a ball". Consequently, we can replace K in Part 2 by any finite union of such sets if we additionally assume that all these sets satisfy $K_i \subseteq \overline{K_i^\circ}$. Note that the assumption that all functions φ_n map to the same bounded ball $U_R(0)$ is necessary. Otherwise we could blow up the configuration of the image points by a larger and larger constant and prevent the approximation error $\|S_n - \varphi_n\|_\infty$ from converging.

4. Proof of Theorem 3 (the infinite case)

The proof of Theorem 3 consists of a number of steps, which we formulate as separate lemmas.

Lemma 5 (Isotonic implies continuous) *Let $\Omega \subseteq \mathbb{R}^d$ and $f : \Omega \to \mathbb{R}^d$ be a locally isotonic function. Then f is continuous. If we additionally assume Ω to be a set with at least one limit point which is contained in it and f to be globally isotonic, then f is even uniformly continuous.*

Proof (sketch) Since continuity is a local property, it suffices to show that for any point $x \in \Omega$ there is a neigborhood $U(x)$ in Ω such that $f|_{U(x)}$ is continuous. Hence, w.l.o.g. we may assume f to be globally isotonic. The key observation is that if f was discontinuous at one point, the distance between different points in $f(\Omega)$ would be bounded from below by a positive constant. In case that Ω is uncountable, this immediately contradicts the separability of $\mathbb{R}^d$. In general, a compactness argument leads to the desired contradiction. In case Ω has a limit point which is contained in it, denote one such point by x_0 and let $\varepsilon > 0$ be arbitrary. We already know that f is continuous, and hence there exists $\delta > 0$ such that $\|f(x) - f(x_0)\| < \varepsilon$ for all $x \in \Omega$ with $\|x - x_0\| < \delta$. Let $x' \in \Omega$ with $0 < \|x' - x_0\| = \delta' < \delta$ (since x_0 is a limit point, there is such a point x'). For all $x, y \in \Omega$ with $\|x - y\| < \delta'$ we have $\|x - y\| < \|x' - x_0\|$ and hence $\|f(x) - f(y)\| < \|f(x') - f(x_0)\| < \varepsilon$. ■

The next lemma shows that if $\Omega \subseteq \mathbb{R}^d$ is a ball and $f : \Omega \to \mathbb{R}^d$ is weakly isotonic, then it is even strongly isotonic, at least on a slightly smaller ball.

Lemma 6 (Weakly isotonic implies strongly isotonic on balls) *Let $\Omega = U_\varepsilon(z) \subseteq \mathbb{R}^d$ and $f : \Omega \to \mathbb{R}^d$ be weakly isotonic. Then $f|_{U_{\varepsilon/4}(z)}$ is strongly isotonic.*

Proof (sketch) In order to prove that $f|_{U_{\varepsilon/4}(z)}$ is isotonic, we have to show that $\|f(x) - f(y)\| < \|f(v) - f(w)\|$ for all $x, y, v, w \in U_{\varepsilon/4}(z)$ with $\|x - y\| < \|v - w\|$. The idea is to use intermediate points $u_1, \ldots, u_n \in \Omega$ such that $\|x - y\| < \|y - u_1\| < \|u_1 - u_2\| < \ldots < \|u_{n-1} - u_n\| < \|u_n - v\| < \|v - w\|$. Since f is assumed to be weakly isotonic, it follows that $\|f(x) - f(y)\| < \|f(y) - f(u_1)\| < \|f(u_1) - f(u_2)\| < \ldots < \|f(u_{n-1}) - f(u_n)\| < \|f(u_n) - f(v)\| < \|f(v) - f(w)\|$. Using continuity we can show that $f|_{U_{\varepsilon/4}(z)}$ is even strongly isotonic. ■

The following proposition already shows that for functions defined on all points of an open and connected domain, all the properties we defined in Definition 2 are equivalent. The key ingredient in the proof is that the midpoint of a line segment between two points in Ω is mapped by an isotony to the midpoint of the line segment between the corresponding image points.

Proposition 7 (Weakly isotonic implies similarity) *Let $\Omega \subseteq \mathbb{R}^d$ be an open and connected domain and $f : \Omega \to \mathbb{R}^d$ be a locally weakly isotonic function. Then f is globally a similarity.*

Proof (details can be found in Appendix B) First we consider a globally strongly isotonic function $f : \Omega = \overline{U_r(z)} \to \mathbb{R}^d$. This allows us to define a function $\mu : [0, \operatorname{diam} \Omega] \to [0, \operatorname{diam} f(\Omega)]$ by $\mu(\|x-y\|) = \|f(x) - f(y)\|$ for all $x, y \in \Omega$. In order to show that f is a similarity, we have to show that μ is linear. By showing that the midpoint of a line segment between two points in Ω is mapped by f to the midpoint of the line segment between the corresponding image points, we iteratively obtain $\mu(\frac{j}{2^i} \operatorname{diam} \Omega) = \frac{j}{2^i} \operatorname{diam} f(\Omega)$, $i \in \mathbb{N}$, $j \in \{0, \ldots, 2^i\}$ (see Appendix B for details). By Lemma 5, f is continuous and so is μ, implying that $\mu(t) = t \cdot (\operatorname{diam} f(\Omega) / \operatorname{diam} \Omega)$.

Now assume that Ω is open and connected and $f : \Omega \to \mathbb{R}^d$ is a locally weakly isotonic function. By Lemma 6, f is locally strongly isotonic. Hence, given $x \in \Omega$ we can choose $\varepsilon(x) > 0$ such that $\overline{U_{\varepsilon(x)}(x)} \subseteq \Omega$ and that $f|_{\overline{U_{\varepsilon(x)}(x)}} : \overline{U_{\varepsilon(x)}(x)} \to \mathbb{R}^d$ is globally strongly isotonic. It follows from the above that $f|_{\overline{U_{\varepsilon(x)}(x)}}$ is a similarity and hence $f : \Omega \to \mathbb{R}^d$ is locally a similarity. By Lemma B (see Appendix A), f is even globally a similarity. ■

Finally, a continuous extension of an isotonic mapping is isotonic too. The proof is elementary.

Lemma 8 (Continuous extension inherits isotony) *Let $\Omega \subseteq \mathbb{R}^d$ such that $K = \overline{\Omega}$ is convex. Let $f : \Omega \to \mathbb{R}^d$ be isotonic and $F : K \to \mathbb{R}^d$ be a continuous extension of f. Then F is isotonic.*

Now, we have collected all ingredients to prove Theorem 3.

Proof of Theorem 3 (sketch) In case that f is globally isotonic and $\overline{\Omega} = \overline{G}$ is convex, we consider the unique continuous extension $\widetilde{F}$ of f to $\overline{\Omega}$. This is possible since f is uniformly continuous by Lemma 5. By Lemma 8, $\widetilde{F}$ is isotonic. According to Proposition 7, $\widetilde{F}|_G$ is even a similarity. By Lemma A (see Appendix A), $\widetilde{F}|_G$ can be uniquely extended to a similarity $F : \mathbb{R}^d \to \mathbb{R}^d$. For the general case we restrict f to several intersections of Ω and small balls. Considering one such a restriction, we are in the situation of the previous case and obtain a unique extension. We can show that all these extensions have to coincide similarly to the proof of Lemma B from Appendix A. ■

5. Proof of Theorem 4 (the finite case)

Case $d = 1$. The case $d = 1$ is particularly simple: it is easy to see that any weakly isotonic function $f : \Omega \to \mathbb{R}$ (with $\Omega \subseteq \mathbb{R}$) is either strictly increasing or decreasing. The following lemma is the main step of the proof in the one-dimensional case. It considers points that approximate a grid, and proves that this property remains intact after an isotonic mapping. See Figure 1 for an illustration.

Lemma 9 (Isotonic maps approximately preserve a grid) *Let $N \in \mathbb{N}$. For some $\varepsilon_1 < 1/2^{2N+1}$ set $\varepsilon_k = \varepsilon_1 2^{k-1}$, $2 \leq k \leq N$, and $\delta = \varepsilon_1/2$. For $k \in \{1, \ldots, N\}$ and $i \in \{1, 3, \ldots, 2^k - 1\}$ set*

(a) $N = 1$ (b) $N = 2$

Figure 1: The idea of Lemma 9 is to place points in small intervals close to the grid points $i/2^N$ ($y^l_{k,i}$ on the left side, $y^r_{k,i}$ on the right side) in such a way that the ordinal constraints between all these points are sufficient to determine the grid cells they belong to, independent of their exact location within the intervals.

$x_{k,i} = i/2^k$ and let $y^l_{k,i}$, $y^r_{k,i}$ be arbitrary elements of $(x_{k,i} - \varepsilon_k - \delta, x_{k,i} - \varepsilon_k)$ and $(x_{k,i} + \varepsilon_k, x_{k,i} + \varepsilon_k + \delta)$, respectively. Let $\varphi : \{0, 1\} \cup \{y^m_{k,i} : m \in \{l, r\}, k \leq N, i \in \{1, 3, \ldots, 2^k - 1\}\} \to [0, 1]$ be a weakly isotonic function with $\varphi(0) = 0$ and $\varphi(1) = 1$. Then it holds that

$$\left|y^m_{k,i} - \varphi(y^m_{k,i})\right| < \frac{1}{2^N}, \quad m \in \{l, r\}, k \leq N, i \in \{1, 3, \ldots, 2^k - 1\}. \tag{3}$$

Proof (details can be found in Appendix B) By induction over N we prove

$$\varphi(y^l_{k,i}) \in \left(\frac{2^{N-k}i - 1}{2^N}, \frac{2^{N-k}i}{2^N}\right), \quad \varphi(y^r_{k,i}) \in \left(\frac{2^{N-k}i}{2^N}, \frac{2^{N-k}i + 1}{2^N}\right),$$

for all $k \leq N$, $i \in \{1, 3, \ldots, 2^k - 1\}$, which immediately implies (3). The basis is clear (see Figure 1(a)): Due to $\varphi(0) = 0$ and $\varphi(1) = 1$, φ is strictly increasing and hence $0 = \varphi(0) < \varphi(y^l_{1,1}) < \varphi(y^r_{1,1}) < \varphi(1) = 1$. Since $|y^l_{1,1} - 0| < |y^l_{1,1} - 1|$ and φ is weakly isotonic, we have $|\varphi(y^l_{1,1}) - \varphi(0)| < |\varphi(y^l_{1,1}) - \varphi(1)|$ and thus can conclude that $\varphi(y^l_{1,1}) \in (0, 1/2)$. In the same way we obtain $\varphi(y^r_{1,1}) \in (1/2, 1)$. We demonstrate the inductive step by proving that the statement also holds for $N = 2$ (see Figure 1(b)): We already know that $\varphi(y^l_{1,1}) \in (0, 1/2)$ and $\varphi(y^r_{1,1}) \in (1/2, 1)$. Furthermore, due to φ being strictly increasing, we have $0 < \varphi(y^l_{2,1}) < \varphi(y^r_{2,1}) < \varphi(y^l_{1,1}) < \varphi(y^r_{1,1}) < \varphi(y^l_{2,3}) < \varphi(y^r_{2,3}) < 1$. The choice of $(\varepsilon_k)_{1 \leq k \leq N}$ and δ guarantees that $|y^l_{2,1} - 0| < |y^l_{2,1} - y^l_{1,1}|$ and $|y^r_{2,1} - y^r_{1,1}| < |y^r_{2,1} - 0|$ leading to $|\varphi(y^l_{2,1}) - 0| < |\varphi(y^l_{2,1}) - \varphi(y^l_{1,1})|$ and $|\varphi(y^r_{2,1}) - \varphi(y^r_{1,1})| < |\varphi(y^r_{2,1}) - 0|$. This yields $2\varphi(y^l_{2,1}) < \varphi(y^l_{1,1}) < 1/2$ and $1/2 < \varphi(y^r_{1,1}) < 2\varphi(y^r_{2,1})$ and thus $\varphi(y^l_{2,1}) \in (0, 1/4)$ and $\varphi(y^r_{2,1}), \varphi(y^l_{1,1}) \in (1/4, 1/2)$. In the same way we can show that $\varphi(y^r_{2,3}) \in (3/4, 1)$ and $\varphi(y^l_{2,3}), \varphi(y^r_{1,1}) \in (1/2, 3/4)$. ■

Now it is straightforward to prove Theorem 4 for the case $d = 1$ (Proposition 10 implies Part 1 of Theorem 4; the proof of Part 2 is the same as for the case $d \geq 2$, which follows later on).

Proposition 10 (Statement for $d = 1$) *Let $I = [a, b]$ (for some $-\infty < a < b < \infty$) and let $(x_n)_{n \in \mathbb{N}}$ be a sequence of points $x_n \in I$ such that $\{x_n : n \in \mathbb{N}\}$ is dense in I. Let $0 < R < \infty$ and $(\varphi_n)_{n \in \mathbb{N}}$ be a sequence of weakly isotonic functions $\varphi_n : \{x_1, \ldots, x_n\} \to [-R, R]$. Then there exists a sequence $(S_n)_{n \in \mathbb{N}}$ of similarity transformations $S_n : \mathbb{R} \to \mathbb{R}$ with (2).*

Proof (sketch) By appropriately rescaling the domain and the image of φ_n we may assume that $I = [0, 1]$ and that φ_n maps to $[0, 1]$ with $\varphi_n(0) = 0$, $\varphi_n(1) = 1$. We use Lemma 9 in order to show that φ_n for large values of n can be approximated by the identity: Choose $N \in \mathbb{N}$ such that $1/2^N$ is sufficiently small. Since $\{x_n : n \in \mathbb{N}\}$ is dense in I, there exists $N_0 \in \mathbb{N}$ such that in each of the intervals $(x_{k,i} - \varepsilon_k - \delta, x_{k,i} - \varepsilon_k)$ and $(x_{k,i} + \varepsilon_k, x_{k,i} + \varepsilon_k + \delta)$ as defined in

Lemma 9 (for the chosen N) there lies an element of $\{x_1, \ldots, x_{N_0}\}$. If $n \geq N_0$, $y \in \{x_1, \ldots, x_n\}$, and y is one of these elements, we immediately obtain $|y - \varphi_n(y)| < 1/2^N$ according to (3). If y is not one of these elements, we can use the monotonicity of φ_n to infer that $|y - \varphi_n(y)|$ is small. ■

Case $d \geq 2$. The case $d \geq 2$ is harder to deal with. Our basic idea is to show that an isotonic mapping $\varphi_n : \{x_1, \ldots, x_n\} \to \mathbb{R}^d$, up to some rescaling, is an $\varepsilon(n)$-*nearisometry*, that is φ_n satisfies

$$\|x - y\| - \varepsilon(n) \leq \|\varphi_n(x) - \varphi_n(y)\| \leq \|x - y\| + \varepsilon(n), \quad x, y \in \{x_1, \ldots, x_n\}. \tag{4}$$

Then, by a theorem of Alestalo et al. (2001), φ_n can be approximated by an isometry up to an error depending (essentially) only on $\varepsilon(n)$ and going to zero as $\varepsilon(n) \to 0$.

For proving that φ_n is an $\varepsilon(n)$-nearisometry we observe the following: since φ_n is isotonic, it is sufficient to prove (4) only for some pairs x, y such that $\|x - y\|$ is roughly uniformly distributed in $[0, \operatorname{diam}\{x_1, \ldots, x_n\}]$. Hence, we would like to consider points close to a straight line and argue in a way similar to Lemma 9 that their relative positions along the line are almost preserved by an isotonic mapping. Yet the problem is that, in general, there is no guarantee that the points are still close to a straight line after applying an isotony. However, assuming that there are points located close to the vertices of a cross-polytope and that these are "fixed" points (this is Assumption ($\star$) in the following lemma), we can show that this is indeed the case and Lemma 9 can be generalized in the following sense. Here we just provide a sketch of the lemma (see also Figure 2 for an explanation). A detailed version can be found in Appendix C.

Lemma 11 (Under Assumption ($\star$) isotonic mappings preserve an approximately straight line) *Let $d \geq 2$. Let $N \in \mathbb{N}$ such that*

$$\omega = 24 \left(\frac{\Gamma(\frac{d}{2} + 1)}{\pi^{\frac{d}{2}}} \right)^{\frac{1}{d}} \left(\frac{1}{2^N - 1} \right)^{\frac{1}{d}} < \frac{1}{2(d-1)}$$

be fixed. Let U_s^+, U_s^-, $\widetilde{U}_s^+$, $\widetilde{U}_s^-$, $s = 1, \ldots, d$, and $U_{k,i}^j$, $U_{k,i}^l$, $U_{k,i}^r$, $k \leq N, i \in \{1, 3, \ldots, 2^k - 1\}, j \in \{2, \ldots, d\}$, be open balls with some certain properties (see Appendix C for details). Let $X_s^+, X_s^- \in \mathbb{R}^d$, $s = 1, \ldots, d$, be arbitrary elements of U_s^+ and U_s^-, respectively, $z_{k,i}^j \in \mathbb{R}^d$ be an arbitrary element of $U_{k,i}^j$, and $y_{k,i}^l$, $y_{k,i}^r \in \mathbb{R}^d$ be arbitrary elements of $U_{k,i}^l$ and $U_{k,i}^r$, respectively. Let $\varphi : \{X_1^+, X_1^-, \ldots, X_d^+, X_d^-\} \cup \{z_{k,i}^j : k \leq N, i \in \{1, 3, \ldots, 2^k - 1\}, j \in \{2, \ldots, d\}\} \cup \{y_{k,i}^m : m \in \{l, r\}, k \leq N, i \in \{1, 3, \ldots, 2^k - 1\}\} \to \mathbb{R}^d$ be an isotonic function and assume that

$$\varphi(X_s^+) \in \widetilde{U}_s^+, \quad \varphi(X_s^-) \in \widetilde{U}_s^-, \quad s = 1, \ldots, d. \tag{$\star$}$$

Set $\gamma(-1) = \gamma(1) = \tilde{\alpha}_1$ and $\gamma(0) = \tilde{\alpha}_1 + \frac{d-1}{2}(\omega + \rho)$ (where $\tilde{\alpha}_1$ is the radius of the balls $\widetilde{U}_1^+, \widetilde{U}_1^-$ and ρ a small number depending on size and location of the balls $\widetilde{U}_s^+, \widetilde{U}_s^-$, $s = 2, \ldots, d$), and define for $2 \leq k \leq N$ and $i \in \{1, 3, \ldots, 2^k - 1\}$ the positive expression $\gamma(-1 + i/2^{k-1})$ recursively by

$$\gamma\left(-1 + \frac{i}{2^{k-1}}\right) = \frac{1}{2}\left(\gamma\left(-1 + \frac{i-1}{2^{k-1}}\right) + \gamma\left(-1 + \frac{i+1}{2^{k-1}}\right) + (d-1)(\omega + 2\rho)\right).$$

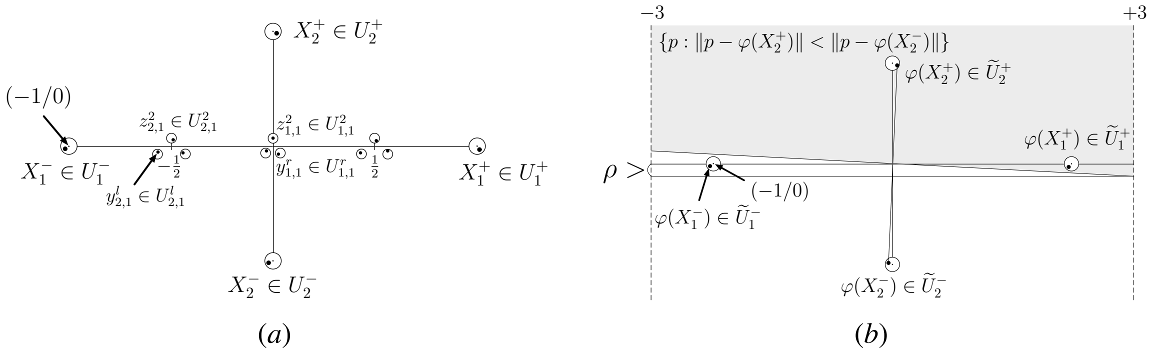

Figure 2: Explanation of Lemma 11 (for $d = 2$) We consider an isotonic map φ defined on the following point set (see 2(a)): (i) $X_1^+, X_1^-, X_2^+, X_2^-$ are located in small balls around the vertices of a cross-polytope and assumed to be "fixed" under φ (this is Assumption ($\star$)). (ii) The points $y_{k,i}^l, y_{k,i}^r$ approximate a grid as in Lemma 9 on the line segment between X_1^- and X_1^+ and are closer to X_2^- than to X_2^+. (iii) The points $z_{k,i}^2$ are close to the points $y_{k,i}^l$ and $y_{k,i}^r$ but are closer to X_2^+ than to X_2^-. Since φ is isotonic, the points $\varphi(y_{k,i}^l), \varphi(y_{k,i}^r)$ are closer to $\varphi(X_2^-)$ than to $\varphi(X_2^+)$ and hence $\varphi^2(y_{k,i}^l), \varphi^2(y_{k,i}^r) < \rho$ whereas for the points $\varphi(z_{k,i}^2)$ it is the other way round such that $\varphi^2(z_{k,i}^2) > -\rho$ (see 2(b)). However, $y_{k,i}^m$ ($m \in \{l, r\}$) and $z_{k,i}^2$ are close to each other and so are $\varphi(y_{k,i}^m)$ and $\varphi(z_{k,i}^2)$. We can conclude that all points $\varphi(y_{k,i}^m)$ are close to the first coordinate axis. This allows us to estimate the location of $\varphi(y_{k,i}^m)$ along similar lines as in Lemma 9.

Let $N^* < N$ such that $N^* \cdot 2^{N^*} < \frac{1}{5(d+1)(\omega+\rho+\tilde{\alpha}_1)}$. Then we have

$$\left\|y_{k,i}^m - \varphi\left(y_{k,i}^m\right)\right\| < \gamma(x_{k,i}) + \omega + (d-1)(\omega+\rho) < 3d\sqrt{\omega}, \quad m \in \{l, r\}, \tag{5}$$

where $x_{k,i} = -1 + \frac{i}{2^{k-1}}$, for all $1 \leq k \leq N^*$ and $i \in \{1, 3, \ldots, 2^k - 1\}$.

Proof (sketch) We prove that for all $1 \leq k \leq N^*$ and $i \in \{1, 3, \ldots, 2^k - 1\}$,

$$\begin{aligned} \varphi(y_{k,i}^l) &\in (x_{k,i} - \gamma(x_{k,i}) - \omega, x_{k,i} + \gamma(x_{k,i})) \times (-\rho - \omega, \rho)^{d-1}, \\ \varphi(y_{k,i}^r) &\in (x_{k,i} - \gamma(x_{k,i}), x_{k,i} + \gamma(x_{k,i}) + \omega) \times (-\rho - \omega, \rho)^{d-1}. \end{aligned} \tag{6}$$

It is elementary to show that $\gamma(x_{k,i}) < \frac{1}{2}(d-1)\sqrt{3\omega}$, $k \leq N^*$, and because of $y_{k,i}^l \in (x_{k,i} - \omega, x_{k,i}) \times (-\omega, 0)^{d-1}$, $y_{k,i}^r \in (x_{k,i}, x_{k,i} + \omega) \times (-\omega, 0)^{d-1}$ this immediately yields (5).

All points $y_{k,i}^l, y_{k,i}^r, z_{k,i}^j$ lie in the convex hull of the points $X_1^+, X_1^-, \ldots, X_d^+, X_d^-$. Since φ is isotonic and satisfies Assumption ($\star$), one can roughly estimate that

$$\varphi(y_{k,i}^l), \varphi(y_{k,i}^r), \varphi(z_{k,i}^j) \in [-3, 3]^d. \tag{7}$$

The idea for proving $\varphi^j(y_{k,i}^l), \varphi^j(y_{k,i}^r) \in (-\rho - \omega, \rho)$, $j \in \{2, \ldots, d\}$, is the following: Let j be fixed. For $m \in \{l, r\}$, $k \in \{1, \ldots, N\}$, $i \in \{1, 3, \ldots, 2^k - 1\}$ we have $\|y_{k,i}^m - X_j^-\| < \|y_{k,i}^m - X_j^+\|$ and $\|z_{k,i}^j - X_j^+\| < \|z_{k,i}^j - X_j^-\|$. Since φ is isotonic, it follows that $\|\varphi(y_{k,i}^m) - \varphi(X_j^-)\| < \|\varphi(y_{k,i}^m) - \varphi(X_j^+)\|$ and $\|\varphi(z_{k,i}^j) - \varphi(X_j^+)\| < \|\varphi(z_{k,i}^j) - \varphi(X_j^-)\|$. Because of ($\star$) and (7), we can conclude that $\varphi^j(y_{k,i}^m) < \rho$ and $\varphi^j(z_{k,i}^j) > -\rho$ (see Figure 2(b)). The distance between any two points $z^j_{k_1,i_1}, z^j_{k_2,i_2}$ is larger than the distance between any two points $z^j_{k,i}, y^l_{k,i}$ (or $z^j_{k,i}, y^r_{k,i}$, respectively), that is for $m \in \{l,r\}$, $k \in \{1,\dots,N\}$, $i \in \{1,3,\dots,2^k-1\}$ it holds that

$$\|z^j_{k,i} - y^m_{k,i}\| < \min\left\{\|u-v\| : u \neq v \in \left\{z^j_{\tilde{k},\tilde{i}} : \tilde{k} \leq N, \tilde{i} \in \left\{1,3,\dots,2^{\tilde{k}}-1\right\}\right\}\right\}. \tag{8}$$

Let $m \in \{l,r\}$, $k_0 \leq N$, $i_0 \in \{1,3,\dots,2^{k_0}-1\}$ be arbitrary and write $r = \|\varphi(z^j_{k_0,i_0}) - \varphi(y^m_{k_0,i_0})\|$. Due to (8) and φ being isotonic, all points $\varphi(z^j_{k,i})$ are located at distance larger than r to each other which implies that the intersection of two balls (whether open or closed) with radius $r/2$ and centers $\varphi(z^j_{k_1,i_1})$ and $\varphi(z^j_{k_2,i_2})$, respectively, is empty. Recall (7). Due to ($\star$) and, again, φ being isotonic, we clearly have $r \leq 3$. Hence, with each point $\varphi(z^j_{k,i})$ at least a fraction of $1/2^d$ of the volume of the ball $U_{r/2}(\varphi(z^j_{k,i}))$ is contained in $[-3,3]^d$ too. We can infer that

$$(2^N-1)\frac{1}{2^d}\frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}\left(\frac{r}{2}\right)^d \leq 6^d,$$

or equivalently $r \leq \omega$. Hence, we have $|\varphi^j(z^j_{k_0,i_0}) - \varphi^j(y^m_{k_0,i_0})| \leq \|\varphi(z^j_{k_0,i_0}) - \varphi(y^m_{k_0,i_0})\| \leq \omega$ and finally obtain $\varphi^j(y^l_{k_0,i_0}), \varphi^j(y^r_{k_0,i_0}) \in (-\rho-\omega, \rho)$. Similar to (8), we also have

$$\|y^l_{k,i} - y^r_{k,i}\| < \min\left\{\|u-v\| : u \neq v \in \left\{z^j_{\tilde{k},\tilde{i}} : \tilde{k} \leq N, \tilde{i} \in \left\{1,3,\dots,2^{\tilde{k}}-1\right\}\right\}\right\},$$

and with the same argument as above obtain $|\varphi^1(y^l_{k,i}) - \varphi^1(y^r_{k,i})| \leq \omega$, $k \leq N$, $i \in \{1,3,\dots,2^k-1\}$. Now, (6) can be shown by induction over k. ■

The following lemma shows that the Assumption ($\star$), which says that points close to the vertices of a cross-polytope are mapped approximately to themselves, can be taken as satisfied if the isotonic function acts on sufficiently many points. See Figure 3 for an explanation. Again, here we just provide a sketch of the lemma and the detailed version is in Appendix C.

Lemma 12 (Assumption ($\star$) can be taken as satisfied) *Let $d \geq 2$. Let $N' \in \mathbb{N}$ such that*

$$\omega' = 32\left(\frac{\Gamma(\frac{d}{2}+1)}{\pi^{\frac{d}{2}}}\right)^{\frac{1}{d}}\frac{1}{\sqrt[d]{N'}}$$

is sufficiently small and $r < 1$ and $\mu, \delta, \varepsilon > 0$ be appropriately chosen real numbers (see Appendix C for details). Define points $A, B \in \mathbb{R}^d$ and $Z^-_s, Z^+_s \in \mathbb{R}^d$, $s \in \{2,\dots,d\}$, by

$$A = (-1/0/\dots/0), \quad B = (1/0/\dots/0), \quad Z^-_2 = (0/-r/0/0/\dots),$$
$$Z^+_2 = (0/r/0/0/\dots), \quad Z^-_3 = (0/0/-r/0/\dots), \quad Z^+_3 = (0/0/r/0/\dots), \quad \textit{and so forth.}$$

For $s \in \{2,\dots,d\}$ and $v \in \{-1,1\}^d$ set $E^-_{s,v} = Z^-_s + \mu v$, $E^+_{s,v} = Z^+_s + \mu v$ and let $e^-_{s,v}, e^+_{s,v} \in \mathbb{R}^d$ be arbitrary elements of $U_\varepsilon(E^-_{s,v})$ and $U_\varepsilon(E^+_{s,v})$, respectively. For $i \in \{1,\dots,2N'-1\}$ let $x_i \in \mathbb{R}^d$ be an arbitrary element of $U_\delta((-1+\frac{i}{N'}/0/\dots/0))$. Let $\varphi : \{A,B\} \cup \{e^-_{s,v}, e^+_{s,v} : s \in \{2,\dots,d\}, v \in$

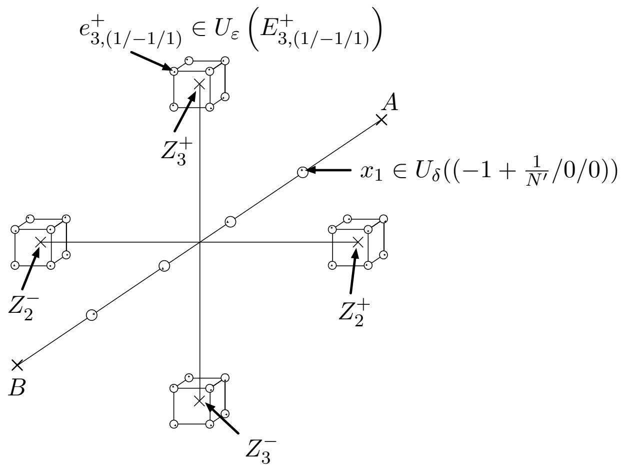

Figure 3: Explanation of Lemma 12 We consider an isotonic mapping φ defined on the following point set: (i) A and B are opposite vertices of a cross-polytope. (ii) The points $e^-_{s,v}, e^+_{s,v}$ are located in small balls around the vertices of hypercubes placed around the remaining vertices of the cross-polytope. (iii) Numerous points x_i are located in small balls which are placed equidistantly between A and B. This yields ordinal constraints sufficient to show that all points $e^-_{s,v}, e^+_{s,v}$ are "fixed" under φ up to some similarity transformation. The figure shows the setting of Lemma 12 for $d = 3$.

$\{-1,1\}^d\} \cup \{x_i : i = 1, \ldots, 2N'-1\} \to \mathbb{R}^d$ be an isotonic function with $\|\varphi(A) - \varphi(B)\| = 2$. Then there exist a constant C (depending only on d) and an isometry $S : \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\begin{aligned}
&\|A - S(\varphi(A))\| \leq C\sqrt{A(\omega')}, \quad \|B - S(\varphi(B))\| \leq C\sqrt{A(\omega')},\\
&\|Z_s^m - S(\varphi(e^m_{s,\underline{v}}))\| \leq C\sqrt{A(\omega')}, \quad m \in \{-,+\}, s \in \{2, \ldots, d\},
\end{aligned}$$

where $\underline{v} = (1/1/1/\ldots/1)$ and $A(\omega')$ only depends on ω' and d and satisfies $A(\omega') \to 0$ as $\omega' \to 0$.

Proof (sketch) With an argument similar to the one subsequent to (8) in the proof of Lemma 11 we can show that all points p_1, p_2 in the domain of φ and with $\|p_1 - p_2\| < \|A - x_1\|$ must satisfy $\|\varphi(p_1) - \varphi(p_2)\| < \omega'/2$. The parameters μ, δ and ε are chosen in such a way that $\|e^m_{s,v_1} - e^m_{s,v_2}\| < \|A - x_1\|$ for any $m \in \{-,+\}$, $s \in \{2, \ldots, d\}$ and for all $v_1, v_2 \in \{-1,1\}^d$. Using this and the assumption of φ being isotonic, we can show that

$$\begin{aligned}
&2 - \omega' < \|p_s^+ - p_s^-\| \leq 2, \qquad s = 1, \ldots, d,\\
&\|p_s^+ - p_{s'}^-\| - \omega' < \|p_s^+ - p_{s'}^+\| < \|p_s^+ - p_{s'}^-\| + \omega', \quad s \neq s' \in \{1, \ldots, d\},\\
&\|p_s^- - p_{s'}^-\| - \omega' < \|p_s^- - p_{s'}^+\| < \|p_s^- - p_{s'}^-\| + \omega', \quad s \neq s' \in \{1, \ldots, d\},
\end{aligned} \tag{9}$$

where $p_1^+ = \varphi(B)$, $p_1^- = \varphi(A)$ and $p_s^+ = \varphi(e^+_{s,\underline{v}})$, $p_s^- = \varphi(e^-_{s,\underline{v}})$ for $s = 2, \ldots, d$. For example, let us prove $\|p_1^- - p_2^-\| - \omega' < \|p_1^- - p_2^+\| < \|p_1^- - p_2^-\| + \omega'$: Elementary calculations show that $\|A - e^-_{2,\underline{v}}\| < \|A - e^+_{2,\underline{v}}\|$ and $\|A - e^+_{2,v^c}\| < \|A - e^-_{2,v^c}\|$ with $v^c = (-1/-1/-1/\ldots/-1)$. We infer $\|p_1^- - p_2^-\| < \|p_1^- - p_2^+\|$ and $\|p_1^- - \varphi(e^+_{2,v^c})\| < \|p_1^- - \varphi(e^-_{2,v^c})\|$ and thus obtain

$$\begin{aligned}
\|p_1^- - p_2^+\| &\leq \|p_1^- - \varphi(e^+_{2,v^c})\| + \|\varphi(e^+_{2,v^c}) - p_2^+\| < \|p_1^- - \varphi(e^-_{2,v^c})\| + \|\varphi(e^+_{2,v^c}) - p_2^+\|\\
&\leq \|p_1^- - p_2^-\| + \|p_2^- - \varphi(e^-_{2,v^c})\| + \|\varphi(e^+_{2,v^c}) - p_2^+\| < \|p_1^- - p_2^-\| + \omega'.
\end{aligned}$$

From (9) we can infer that

$$|\langle p_s^+ - p_s^-, p_{s'}^+ - p_{s'}^- \rangle| < 10\omega', \quad s \neq s' \in \{1, \ldots, d\}. \tag{10}$$

Furthermore, we can show that $\|(p_s^+ + p_s^-) - (p_{s'}^+ + p_{s'}^-)\|$, $s \neq s' \in \{1, \ldots, d\}$, is small (provided ω' is small), that is

$$\|(p_s^+ + p_s^-) - (p_{s'}^+ + p_{s'}^-)\|^2 \leq d \left(\frac{20\omega' + 8\frac{10d\omega'}{4d-1}(4d)^{d-1}}{2 - \omega' - \frac{10d\omega'}{4d-1}(4d)^{d-1}} \right)^2, \quad s \neq s' \in \{1, \ldots, d\}. \tag{11}$$

This is done by first applying the Gram-Schmidt process to the vectors $(p_s^+ - p_s^-)$, $s = 1, \ldots, d$. By doing so we obtain an orthonormal basis of $\mathbb{R}^d$ whose elements (appropriately rescaled) differ from the vectors $(p_s^+ - p_s^-)$ only up to some small error (depending on ω'). Considering the Fourier coefficients of $(p_s^+ + p_s^-) - (p_{s'}^+ + p_{s'}^-)$ with respect to this orthonormal basis then leads to (11) .

Now, setting $Z_1^- = A$, $Z_1^+ = B$, we consider the mapping $f : \{Z_1^-, Z_1^+, \ldots, Z_d^-, Z_d^+\} \to \{p_1^-, p_1^+, \ldots, p_d^-, p_d^+\}$ given by $f(Z_s^m) = p_s^m$ for $m \in \{-, +\}$, $s \in \{1, \ldots, d\}$. Using (10) and (11) it is straightforward to show that f is a $2\sqrt{A(\omega')}$-nearisometry, that is it holds that

$$\|x - y\| - 2\sqrt{A(\omega')} \leq \|f(x) - f(y)\| \leq \|x - y\| + 2\sqrt{A(\omega')}, \quad x, y \in \{Z_1^-, Z_1^+, \ldots, Z_d^-, Z_d^+\}.$$

According to Alestalo et al. (2001), Theorem 3.3, there exists a constant C' (depending only on d — can be chosen independently of the parameters $r, \mu, \delta, \varepsilon$) and an isometry $T : \mathbb{R}^d \to \mathbb{R}^d$ such that $\|T(x) - f(x)\| \leq 2C'\sqrt{A(\omega')}$, $x \in \{Z_1^-, Z_1^+, \ldots, Z_d^-, Z_d^+\}$. Setting $S = T^{-1}$ and $C = 2C'$ the statement of Lemma 12 follows immediately. ■

Now we can prove Theorem 4 for $d \geq 2$.

Proof of Part 1 of Theorem 4 (sketch) By Lemma C (see Appendix A) it is sufficient to prove that for every $\varepsilon_0 > 0$ there exists $N(\varepsilon_0) \in \mathbb{N}$ such that for all $n \geq N(\varepsilon_0)$ there is a similarity transformation $S(n, \varepsilon_0) : \mathbb{R}^d \to \mathbb{R}^d$ with $\|\varphi_n - S(n, \varepsilon_0)\|_{\infty(\{x_1, \ldots, x_n\})} < \varepsilon_0$.

In a nutshell, the basic idea is the following: Assume K is a ball with diameter only slightly larger than two and containing all the balls of Lemma 11. If $n \in \mathbb{N}$ is sufficiently large, in each of these balls there is an element of $\{x_1, \ldots, x_n\}$. Assume for the moment that φ_n satisfies Assumption ($\star$) of Lemma 11. Then from (5) we obtain an estimate for the expression $\|\varphi_n(x) - \varphi_n(y)\|$ for roughly uniformly distributed values of $\|x - y\|$ in $[0, 2] \approx [0, \operatorname{diam}\{x_1, \ldots, x_n\}]$. Since φ_n is isotonic, this gives an estimate for $\|\varphi_n(x) - \varphi_n(y)\|$ for all $x, y \in \{x_1, \ldots, x_n\}$ which is sufficient to show that φ_n is an ε-nearisometry for some small ε. Hence, we can uniformly approximate φ_n by an isometry according to Alestalo et al. (2001). It remains to be argued why Assumption ($\star$) of Lemma 11 indeed can be taken as satisfied. However, this is the statement of Lemma 12.

A bit more precisely, the main steps of the proof can be summarized as follows:

1. Since $\{x_n : n \in \mathbb{N}\}$ is dense in K, we can choose $N_0 \in \mathbb{N}$ so large that there are points $x_A, x_B \in \{x_1, \ldots, x_{N_0}\}$ and a similarity transform $T : \mathbb{R}^d \to \mathbb{R}^d$ with the following properties:
    - $T(x_A) = (-1/0/0/\ldots/0)$, $T(x_B) = (1/0/0/\ldots/0)$
    - $U_1(0) \subseteq T(K)$, $\operatorname{diam} T(K)$ is "sufficiently small"
    - $\forall y \in T(K) : U_{r_0}(y) \cap \{T(x_1), \ldots, T(x_{N_0})\} \neq \emptyset$ where $r_0 > 0$ is smaller than the minimal radius of the finitely many open balls considered in Step 3 and smaller than δ_0 from Step 6.

    In the following, we consider φ_n for a fixed $n \geq N_0$.

2. Let $U : \mathbb{R}^d \to \mathbb{R}^d$ be a similarity transformation with $\|U(\varphi_n(x_A)) - U(\varphi_n(x_B))\| = 2$. For its scale factor $\lambda(U)$ we have

$$\lambda(U) = \frac{\|U(\varphi_n(x_A)) - U(\varphi_n(x_B))\|}{\|\varphi_n(x_A) - \varphi_n(x_B)\|} \geq \frac{2}{2R} = \frac{1}{R}. \tag{12}$$

3. (a) We choose $N \in \mathbb{N}$ such that $\omega = \omega(N)$ as in Lemma 11 is "sufficiently small" and N^* from Lemma 11 can be chosen "sufficiently large". We choose all parameters of Lemma 11 (see the detailed version in Appendix C) except the vectors $r, \tilde{r}$ appropriately — assuming that we will choose $r, \tilde{r}$ with $r = \tilde{r} \geq 1/2$ (in every component) afterwards.
   (b) We choose $N' \in \mathbb{N}$ such that $\omega' = \omega'(N')$ as in Lemma 12 satisfies $C\sqrt{A(\omega')} < \tilde{\alpha}_i$, $i = 1, \ldots, d$ (with $\tilde{\alpha}_i$ from (a)). We choose all parameters of Lemma 12 (see the detailed version in Appendix C) appropriately and such that
      - $U_\varepsilon(E^-_{s,v}) \subseteq U_{\alpha_s}(Z^-_s), U_\varepsilon(E^+_{s,v}) \subseteq U_{\alpha_s}(Z^+_s)$ (with α_s from (a)), $s \in \{2, \ldots, d\}$,
      - all corresponding balls are contained in $U_1(0)$.
   (c) Denoting the parameter r from (b) by r', we use r' to define r from Lemma 11 as $r = (1/r'/r'/\ldots/r')$ and set $\tilde{r} = r$. Hence, we have chosen all parameters of Lemma 11.

4. We consider the map $U \circ \varphi_n \circ T^{-1} : \{T(x_1), \ldots, T(x_n)\} \to \mathbb{R}^d$. According to Step 1, in every open ball of Lemma 12 there is an element of $\{T(x_1), \ldots, T(x_n)\}$. We denote these elements as in Lemma 12 ($A = T(x_A), B = T(x_B)$). According to Lemma 12 there exists an isometry $S : \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\|A - S(U \circ \varphi_n \circ T^{-1}(A))\| \leq C\sqrt{A(\omega')} < \tilde{\alpha}_1, \quad \|B - S(U \circ \varphi_n \circ T^{-1}(B))\| < \tilde{\alpha}_1,$$
$$\|Z^m_s - S(U \circ \varphi_n \circ T^{-1}(e^m_{s,\underline{v}}))\| \leq C\sqrt{A(\omega')} < \tilde{\alpha}_s, \quad m \in \{-, +\}, s \in \{2, \ldots, d\}.$$

5. We consider the map $S \circ U \circ \varphi_n \circ T^{-1} : \{T(x_1), \ldots, T(x_n)\} \to \mathbb{R}^d$. According to Step 1, in every open ball of Lemma 11 there is an element of $\{T(x_1), \ldots, T(x_n)\}$. We denote these elements as in Lemma 11 ($X^-_1 = A = T(x_A), X^+_1 = B = T(x_B), X^m_s = e^m_{s,\underline{v}}$). According to Step 4, Assumption ($\star$) of Lemma 11 is satisfied. Hence, we have

$$\|y^m_{k,i} - S \circ U \circ \varphi_n \circ T^{-1}(y^m_{k,i})\| < 3d\sqrt{\omega}, \quad m \in \{l, r\}, k \leq N^*, i \in \{1, 3, \ldots, 2^k - 1\},$$

   where $\omega = \omega(N)$ from Step 3(a).

6. We show that $f = S \circ U \circ \varphi_n \circ T^{-1}$ is an ε-nearisometry for some small ε (depending on ω and $1/2^{N^*}$), that is f satisfies

$$\|x - y\| - \varepsilon \leq \|f(x) - f(y)\| \leq \|x - y\| + \varepsilon, \quad x, y \in \{T(x_1), \ldots, T(x_n)\}. \tag{13}$$

   For elements x, y with $\|x - y\| \leq (2^{N^*} - 2)/2^{N^*-1}$ it is straightforward to show the inequality (13) by approximating $\|x - y\|$ by $\|\tilde{x} - \tilde{y}\|$ with elements $\tilde{x}, \tilde{y} \in \{y^m_{k,i} : m \in \{l, r\}, k \leq N^*, i \in \{1, 3, \ldots, 2^k - 1\}\}$ and using that f is isotonic.

   For elements x, y with $(2^{N^*} - 2)/2^{N^*-1} < \|x - y\| \leq \operatorname{diam} T(K)$ we approximate $\|x - y\| = \|x - x'\| + \|x' - y'\| + \|y' - y\|$ by $\|x - x^a\| + \|x^a - y^a\| + \|y^a - y\|$ where each summand is smaller than $(2^{N^*}-2)/2^{N^*-1}$ by approximating $x' = x+\frac{1}{3}(y-x)$, $y' = x+\frac{2}{3}(y-x)$ by elements $x^a \in U_{\delta_0}(x')$, $y^a \in U_{\delta_0}(y')$ (for some "sufficiently small" δ_0) with $x^a, y^a \in \{T(x_1), \ldots, T(x_n)\}$. Since K is assumed to be convex, so is $T(K)$ and hence $x', y' \in T(K)$. According to Step 1 there exist such elements x^a, y^a. Using that $\operatorname{diam} T(K)$ is small (see Step 1), then it is easy to show that $\|x-y\| - \varepsilon \le \|f(x) - f(y)\| \le \|x-y\| + \varepsilon$.

7. According to Alestalo et al. (2001) (Theorem 2.2 or Theorem 3.3) there exists an isometry $S' : \mathbb{R}^d \to \mathbb{R}^d$ such that S' uniformly approximates $S \circ U \circ \varphi_n \circ T^{-1}$ up to some small error (depending on ε). Since $\lambda(U)$ is bounded from below by (12), $U^{-1} \circ S^{-1} \circ S' \circ T$ is a good approximation of φ_n. ∎

Proof of Part 2 of Theorem 4 (details can be found in Appendix B) W.l.o.g. we may assume that $K = \cup_{i=1}^k K_i$ such that $K_i^\circ \cap K_{i+1}^\circ \neq \emptyset$, $i = 1, \ldots, k-1$. For every $i \in \{1, \ldots, k\}$, $\{x_n : n \in \mathbb{N}\} \cap K_i$ is dense in K_i because of $K_i \subseteq \overline{K_i^\circ}$. Hence, there exists a sequence $(S_n^i)_{n\in\mathbb{N}}$ of similarity transformations such that $\|S_n^i - \varphi_n|_{\{x_1,\ldots,x_n\}\cap K_i}\|_{\infty(\{x_1,\ldots,x_n\}\cap K_i)} \to 0$. We prove that

$$\|S_n^1 - \varphi_n|_{\{x_1,\ldots,x_n\}\cap(K_1\cup\ldots\cup K_j)}\|_{\infty(\{x_1,\ldots,x_n\}\cap(K_1\cup\ldots\cup K_j))} \to 0, \quad j = 1, \ldots, k,$$

by induction over j. The key observation for the inductive step is the following: If S_n^1 is a good approximation to φ_n for points in $K_1 \cup \ldots \cup K_{j-1}$ and S_n^j for points in K_j, the similarities S_n^1 and S_n^j differ only slightly for points in $(K_1 \cup \ldots \cup K_{j-1}) \cap K_j$. Since K is bounded, $S_n^1(x)$ cannot be too different from $S_n^j(x)$ for any point $x \in K$. ∎

6. Discussion

The main result of our paper is to establish the uniqueness of ordinal embedding, upon knowledge of all pairwise constraints in local regions. This result closes a long-standing gap in the literature on ordinal embedding. However, there are a number of interesting and important follow-up questions that are still open: (1) Our current Theorem 4 states the convergence of isotonic embeddings but does not give any error rates. It would be desirable to have a statement such as "if the points have been sampled from some nice probability density, then with high probability an embedding of n points has error at most ε". However, it seems difficult to get such a statement, our current proof techniques are not powerful enough to obtain strong bounds. (2) It also seems plausible that ordinal embedding is still possible in a noisy scenario where either the distance measurements are noisy or some of the constraints $\|x_i - x_j\| < \|x_k - x_l\|$ have been flipped. (3) In this paper we have shown that already mere local ordinal information guarantees a unique embedding. However, it seems quite reasonable that even fewer ordinal relationships might be sufficient for reconstructing a given set of points (in a similar spirit as turning partial orders to total orders). We also do not know whether it is sufficient in Theorem 4 to assume the functions φ_n to be only weakly isotonic (for $d \ge 2$ — compare with Proposition 10 for $d = 1$). So finally, what is the minimal amount of ordinal information that is necessary for reconstructing a given set of points up to a given precision?

Acknowledgments

Ulrike von Luxburg acknowledges funding of the German Research Foundation (grant LU1718/1-1 and Research Unit 1735 "Structural Inference in Statistics: Adaptation and Efficiency").

References

S. Agarwal, J. Wills, L. Cayton, G. Lanckriet, D. Kriegman, and S. Belongie. Generalized non-metric multidimensional scaling. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2007.

N. Ailon. An active learning algorithm for ranking from pairwise preferences with an almost optimal query complexity. *JMLR*, 13:137–164, 2012.

P. Alestalo, D. Trotsenko, and J. Väisälä. Isometric approximation. *Israel Journal of Mathematics*, 125:61–82, 2001.

F. Beckman and D. Quarles. On isometries of Euclidean spaces. *Proc. Amer. Math. Soc.*, 4:810–815, 1953.

I. Borg and P. Groenen. *Modern multidimensional scaling: Theory and applications*. Springer, 2005.

J. Dattorro. *Convex optimization and Euclidean distance geometry*. Meboo Publishing, 2005.

K. Jamieson and R. Nowak. Low-dimensional embedding using adaptively selected ordinal data. In *Conference on Communication, Control, and Computing*, 2011a.

K. Jamieson and R. Nowak. Active ranking using pairwise comparisons. In *Neural Information Processing Systems (NIPS)*, 2011b.

J. Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29(1):1–27, 1964a.

J. Kruskal. Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 29(2):115–129, 1964b.

Y. Lan, J. Guo, X. Cheng, and T. Liu. Statistical consistency of ranking methods in a rank-differentiable probability space. In *Neural Information Processing Systems (NIPS)*, 2012.

B. McFee and G. Lanckriet. Partial order embedding with multiple kernels. In *International Conference on Machine Learning (ICML)*, 2009.

B. McFee and G. Lanckriet. Metric learning to rank. In *International Conference on Machine Learning (ICML)*, 2010.

B. McFee and G. Lanckriet. Learning multi-modal similarity. *JMLR*, 12:491–523, 2011.

M. McKemie and J. Väisälä. Spherical maps of Euclidean spaces. *Results in Mathematics*, 35: 145–160, 1999.

H. Ouyang and A. Gray. Learning dissimilarities by ranking: from SDP to QP. In *International Conference on Machine Learning (ICML)*, 2008.

M. Quist and G. Yona. Distributional scaling: An algorithm for structure-preserving embedding of metric and nonmetric spaces. *JMLR*, 5:399–420, 2004.

R. Rosales and G. Fung. Learning sparse metrics via linear programming. In *International Conference on Knowledge Discovery and Data Mining (KDD)*, 2006.

I. Schoenberg. Metric spaces and completely monotone functions. *Ann. of Math.*, 39(4):811–841, 1938.

B. Shaw and T. Jebara. Structure preserving embedding. In *International Conference on Machine Learning (ICML)*, 2009.

R. Shepard. The analysis of proximities: Multidimensional scaling with an unknown distance function (I). *Psychometrika*, 27(2):125–140, 1962a.

R. Shepard. The analysis of proximities: Multidimensional scaling with an unknown distance function (II). *Psychometrika*, 27(3):219–246, 1962b.

R. Shepard. Metric structures in ordinal data. *Journal of Mathematical Psychology*, 3(2):287–315, 1966.

C. Sherman. Nonmetric multidimensional scaling: A Monte Carlo study of the basic parameters. *Psychometrika*, 37(3):323–355, 1972.

O. Tamuz, C. Liu, S. Belongie, O. Shamir, and A. Kalai. Adaptively learning the crowd kernel. In *International Conference on Machine Learning (ICML)*, 2011.

F. Wauthier, M. Jordan, and N. Jojic. Efficient ranking from pairwise comparisons. In *International Conference on Machine Learning (ICML)*, 2013.

F. Young. Nonmetric multidimensional scaling: Recovery of metric information. *Psychometrika*, 35(4):455–473, 1970.

Appendix A. Additional lemmas

Lemma A (Extending a similarity) *Let $\Omega \subseteq \mathbb{R}^d$ and $f : \Omega \to \mathbb{R}^d$ be a similarity. Then there exists an affine and surjective similarity $F : \mathbb{R}^d \to \mathbb{R}^d$ (that is F is a similarity transformation) such that $F(x) = f(x)$, $x \in \Omega$. The function F is uniquely determined by f if and only if $\mathcal{H}(\Omega) = \mathbb{R}^d$.*

Proof Let $\lambda > 0$ such that $\|f(x) - f(y)\| = \lambda\|x - y\|$, $x, y \in \Omega$. We may assume that $\lambda = 1$ since otherwise we can set $\tilde{f} = (1/\lambda)f$ and $F = \lambda\tilde{F}$ if $\tilde{F}$ is an extension of $\tilde{f}$. In the following we distinguish three cases:

- $0 \in \Omega$ and $f(0) = 0$

  This implies that $\|f(x)\| = \|x\|$, $x \in \Omega$, and because of

$$\begin{aligned}&\|f(x) - f(x')\|^2 = \|x - x'\|^2 \\ &\qquad \Rightarrow \|f(x)\|^2 - 2\langle f(x), f(x')\rangle + \|f(x')\|^2 = \|x\|^2 - 2\langle x, x'\rangle + \|x'\|^2,\end{aligned}$$

  we can conclude that $\langle f(x), f(x')\rangle = \langle x, x'\rangle$, $x, x' \in \Omega$.

  Let $x_1, \dots, x_n \in \Omega$ form a basis of $[\Omega]$. If $x \in \Omega$ and $x = \sum_{i=1}^n c_i x_i$, then

$$\begin{aligned}\left\|f(x) - \sum_{i=1}^n c_i f(x_i)\right\|^2 &= \|f(x)\|^2 - 2\left\langle f(x), \sum_{i=1}^n c_i f(x_i)\right\rangle + \left\|\sum_{i=1}^n c_i f(x_i)\right\|^2 \\ &= \|x\|^2 - 2\sum_{i=1}^n c_i \langle x, x_i\rangle + \sum_{i=1}^n\sum_{j=1}^n c_i c_j \langle x_i, x_j\rangle \\ &= \sum_{i=1}^n\sum_{j=1}^n c_i c_j \langle x_i, x_j\rangle - 2\sum_{i=1}^n\sum_{j=1}^n c_i c_j \langle x_i, x_j\rangle + \sum_{i=1}^n\sum_{j=1}^n c_i c_j \langle x_i, x_j\rangle \\ &= 0,\end{aligned}$$

  hence $f(x) = \sum_{i=1}^n c_i f(x_i)$. Thus, by setting

$$f'(\tilde{x}) = \sum_{i=1}^n \tilde{c}_i f(x_i) \quad \text{for } \tilde{x} = \sum_{i=1}^n \tilde{c}_i x_i \in [\Omega]$$

  we can define a linear map f' from $[\Omega]$ to $\mathbb{R}^d$ which coincides with f on Ω.

  Obviously, f' is a linear isometry from $[\Omega]$ onto $f'([\Omega])$. If $[\Omega] \neq \mathbb{R}^d$, we can choose an orthonormal basis of $[\Omega]^\perp$ and one of $f'([\Omega])^\perp$. These comprise the same number of basis vectors since $[\Omega]$ and $f'([\Omega])$ have the same dimension. Let f'' be a linear mapping from $[\Omega]^\perp$ to $f'([\Omega])^\perp$ which maps the orthonormal basis of $[\Omega]^\perp$ onto the one of $f'([\Omega])^\perp$. Then f'' is a linear isometry from $[\Omega]^\perp$ onto $f'([\Omega])^\perp$ and $F = f' \oplus f''$ a linear isometry from $\mathbb{R}^d$ onto $\mathbb{R}^d$.

  Concerning the uniqueness: Clearly, if $[\Omega] \neq \mathbb{R}^d$, we can choose different orthonormal bases of $[\Omega]^\perp$ and $f'([\Omega])^\perp$, respectively — or different mappings between them. On the other hand, if $[\Omega] = \mathbb{R}^d$, any linear extension of f to $\mathbb{R}^d$ is uniquely determined by $f(x_1), \dots, f(x_n)$. Since $0 \in \Omega$, we have $\mathcal{H}(\Omega) = [\Omega]$, and because of $f(0) = 0$, any affine extension of f is linear.

- $0 \in \Omega$ but $f(0) \neq 0$

  Define $f' : \Omega \to \mathbb{R}^d$ by $f'(x) = f(x) - f(0)$, $x \in \Omega$. We can apply the previous case to f' and obtain a linear and isometric extension F' of f'. Setting $F = F' + f(0)$ gives the desired extension of f. Obviously, F is uniquely determined if and only if F' is uniquely determined. As we have seen, this is the case if and only if $\mathcal{H}(\Omega) = \mathbb{R}^d$.

- $0 \notin \Omega$ (in fact, one could deal with the second case in the same way as with this case, and so one could merge them into one case "$0 \notin \Omega$, or $0 \in \Omega$ but $f(0) \neq 0$")

  Let $x' \in \Omega$ be fixed. Set $\Omega' = \Omega - x'$ and define $f' : \Omega' \to \mathbb{R}^d$ by $f'(x-x') = f(x) - f(x')$, $x \in \Omega$. Then it holds that $0 \in \Omega'$, $f'(0) = 0$ and f' is isometric on Ω'. Let F' be the linear and isometric extension of f' to $\mathbb{R}^d$ according to the first case. Define $F : \mathbb{R}^d \to \mathbb{R}^d$ by $F = F' - F'(x') + f(x')$. Since

  $$\begin{aligned} F(x) = F'(x) - F'(x') + f(x') = F'(x - x') + f(x') &= f'(x - x') + f(x') \\ &= f(x) - f(x') + f(x') = f(x) \end{aligned}$$

  for $x \in \Omega$, this gives an affine, surjective and isometric extension of f to $\mathbb{R}^d$. In order to prove the assertion concerning uniqueness of F it suffices to note that $\mathcal{H}(\Omega) = \mathbb{R}^d$ if and only if $[\Omega'] = \mathbb{R}^d$ and that F is unique if and only if F' is unique.

■

Lemma B (Local similarity implies global similarity) *Let $\Omega \subseteq \mathbb{R}^d$ be an open and connected domain and $f : \Omega \to \mathbb{R}^d$ be locally a similarity. Then f is globally a similarity.*

Proof For $z \in \Omega$ we can choose $\varepsilon_z > 0$ and $\lambda_z > 0$ such that $U_{\varepsilon_z}(z) \subseteq \Omega$ and

$$\|f(u) - f(v)\| = \lambda_z \|u - v\| \quad \forall u, v \in U_{\varepsilon_z}(z)$$

since Ω is open and f is locally a similarity. Fix an arbitrary element $x_0 \in \Omega$ and consider the mapping $f|_{U_{\varepsilon_{x_0}}(x_0)} : U_{\varepsilon_{x_0}}(x_0) \to \mathbb{R}^d$ which is a similarity. By Lemma A there exists a unique extension $F_{x_0} : \mathbb{R}^d \to \mathbb{R}^d$ which is a similarity. We will show that $f = F_{x_0}|_\Omega$.

Let $y \neq x_0$ be an arbitrary element of Ω. It is well known that an open and connected subset of $\mathbb{R}^d$ is path-connected, hence there exists a continuous path $\varphi : [0,1] \to \Omega$ with $\varphi(0) = x_0$ and $\varphi(1) = y$. Its image $\varphi([0,1])$ is compact. Hence, we can choose $x_1, \ldots, x_n \in \varphi([0,1])$ with $x_n = y$ such that it is covered by the open balls $U_{\varepsilon_{x_i}}(x_i)$, $i = 0, \ldots, n$. W.l.o.g. we may assume that

$$\forall i = 1, \ldots, n \; \exists w_i \in \varphi([0,1]) \subseteq \Omega : w_i \in U_{\varepsilon_{x_{i-1}}}(x_{i-1}) \cap U_{\varepsilon_{x_i}}(x_i).$$

We will prove by induction that $f|_{U_{\varepsilon_{x_i}}(x_i)} = F_{x_0}|_{U_{\varepsilon_{x_i}}(x_i)}$ for $i = 0, \ldots, n$. This implies $f(y) = F_{x_0}(y)$, and since $y \in \Omega$ was chosen arbitrarily, we can conclude that $f = F_{x_0}|_\Omega$.

The basis ($i = 0$) is clear by construction of F_{x_0}. For the inductive step from $i-1$ to i let $\varepsilon > 0$ such that

$$U_\varepsilon(w_i) \subseteq U_{\varepsilon_{x_{i-1}}}(x_{i-1}) \cap U_{\varepsilon_{x_i}}(x_i).$$

Note that it immediately follows that $\lambda_{x_i} = \lambda_{x_0}$. By Lemma A there exists a unique extension of $f|_{U_\varepsilon(w_i)}$ to a similarity defined on $\mathbb{R}^d$ (which is obviously given by F_{x_0}). There also exists a unique extension of $f|_{U_{\varepsilon_{x_i}}(x_i)}$. However, these extensions have to coincide and thus we have $f|_{U_{\varepsilon_{x_i}}(x_i)} = F_{x_0}|_{U_{\varepsilon_{x_i}}(x_i)}$. ■

Lemma C (A diagonal argument) *Let X be an arbitrary set and $(A_n)_{n\in\mathbb{N}}$, $A_n \subseteq X$, be a sequence of subsets of X. Let $(\varphi_n)_{n\in\mathbb{N}}$ be a sequence of functions $\varphi_n : A_n \to \mathbb{R}^d$. Assume that for every $\varepsilon > 0$ there exists $N(\varepsilon) \in \mathbb{N}$ such that for all $l \geq N(\varepsilon)$ there is a function $S(l, \varepsilon) : X \to \mathbb{R}^d$ with*

$$\|\varphi_l - S(l, \varepsilon)\|_{\infty(A_l)} < \varepsilon.$$

Then there exists a sequence of functions $(S_n)_{n\in\mathbb{N}}$, $S_n : X \to \mathbb{R}^d$, with

$$\|\varphi_n - S_n\|_{\infty(A_n)} \to 0 \quad \text{as } n \to \infty,$$

where every S_n equals a function $S(l_n, \varepsilon_n)$.

Proof We can choose a strictly decreasing sequence of positive reals $(\varepsilon_n)_{n\in\mathbb{N}}$ converging to zero and a strictly increasing sequence of natural numbers $(N_n)_{n\in\mathbb{N}}$ such that for every $l \geq N_n$ there is a function $S(l, \varepsilon_n) : X \to \mathbb{R}^d$ with $\|\varphi_l - S(l, \varepsilon_n)\|_{\infty(A_l)} < \varepsilon_n$. Let $\varepsilon_0 > 0$ and $l_0 \geq N(\varepsilon_0)$ be arbitrary. Set $S_k = S(l_0, \varepsilon_0)$ for $k < N_1$ and $S_k = S(k, \varepsilon_n)$ for $N_n \leq k < N_{n+1}$. In order to show that $\|\varphi_n - S_n\|_{\infty(A_n)} \to 0$, let $\delta > 0$ be arbitrary. Let $n_0 \in \mathbb{N}$ such that $\varepsilon_{n_0} < \delta$. If $m \geq N_{n_0}$, then we have $N_{\tilde{n}} \leq m < N_{\tilde{n}+1}$ for some $\tilde{n} \geq n_0$, and it holds that

$$\|\varphi_m - S_m\|_{\infty(A_m)} = \|\varphi_m - S(m, \varepsilon_{\tilde{n}})\|_{\infty(A_m)} < \varepsilon_{\tilde{n}} < \varepsilon_{n_0} < \delta.$$

■

Appendix B. Detailed proofs of Proposition 7, Lemma 9 and Part 2 of Theorem 4

Proof Proposition 7 Here we want to prove the statement for the case that $\Omega = \overline{U_r(z)}$ is a closed and bounded ball (for some arbitrary $r > 0$, $z \in \mathbb{R}^d$) and $f : \Omega \to \mathbb{R}^d$ is globally strongly isotonic. How to derive the general result from this special case is shown in Section 4.

Consider the set $f(\Omega)$. Since f is continuous by Lemma 5 and Ω is compact, so is $f(\Omega)$. In particular, $f(\Omega)$ is bounded, that is $\operatorname{diam} f(\Omega) < \infty$. We can define a function $\mu : [0, \operatorname{diam}\Omega] \to [0, \operatorname{diam} f(\Omega)]$ as follows:

$$\forall x, y \in \Omega : \|f(x) - f(y)\| = \mu(\|x - y\|).$$

Since f is strongly isotonic, μ is indeed well-defined. Note that μ is definitely defined on the whole interval $[0, \operatorname{diam}\Omega]$ since Ω naturally contains a line segment of length $\operatorname{diam}\Omega$. In order to show that f is a similarity, we have to show that μ is linear (that is given by $\mu(t) = \lambda t$, $t \in [0, \operatorname{diam}\Omega]$, for some $\lambda > 0$).

It follows from f being strongly isotonic that μ is strictly increasing. Obviously, we have $\mu(0) = 0$. Due to the compactness of Ω and $f(\Omega)$ and f being strongly isotonic, we can conclude that $\mu(\operatorname{diam}\Omega) = \operatorname{diam} f(\Omega)$.

Choose points x_0 and y_0 on the boundary of Ω with $\|x_0 - y_0\| = \operatorname{diam}\Omega$ (thus x_0 and y_0 are elements of a straight line going through z). We can write $\mu(t)$ as

$$\mu(t) = \left\| f(x_0) - f\left(x_0 + t\frac{y_0 - x_0}{\|y_0 - x_0\|}\right)\right\|, \quad t \in [0, \operatorname{diam}\Omega].$$

This shows that μ is continuous.

Let $m = (x_0 + y_0)/2$ be the midpoint of the line segment between x_0 and y_0 (in fact, $m = z$). We want to show that $f(m) = (f(x_0) + f(y_0))/2$. If $d = 1$, this immediately follows from f being strongly isotonic. In general, set $r_0 = x_0 - y_0$ and let $R = [r_0]$ be the linear hull of r_0. Let $\{e_1, \ldots, e_{d-1}\}$ be an orthonormal basis of $R^\perp$. We can choose $\varepsilon > 0$ such that all points $p_i^+ = m + \varepsilon e_i$ and $p_i^- = m - \varepsilon e_i$, $i = 1, \ldots, d-1$, are elements of Ω (in fact, we can choose any $\varepsilon \le r$). Set $p_0^+ = x_0$ and $p_0^- = y_0$.
Now we have

$$\|m - p_i^+\| = \|m - p_i^-\|, \quad i = 0, \ldots, d-1,$$

and

$$\begin{aligned}\|p_j^+ - p_i^+\| &= \|p_j^+ - p_i^-\|, \quad i \neq j \in \{0, \ldots, d-1\},\\ \|p_j^- - p_i^+\| &= \|p_j^- - p_i^-\|, \quad i \neq j \in \{0, \ldots, d-1\}.\end{aligned}$$

Since f is strongly isotonic, it follows that

$$\|f(m) - f(p_i^+)\| = \|f(m) - f(p_i^-)\|, \quad i = 0, \ldots, d-1,$$

and

$$\begin{aligned}\|f(p_j^+) - f(p_i^+)\| &= \|f(p_j^+) - f(p_i^-)\|, \quad i \neq j \in \{0, \ldots, d-1\},\\ \|f(p_j^-) - f(p_i^+)\| &= \|f(p_j^-) - f(p_i^-)\|, \quad i \neq j \in \{0, \ldots, d-1\}.\end{aligned}$$

This implies

$$\left\langle f(p_i^+) - f(p_i^-), f(m)\right\rangle = \left\langle f(p_i^+) - f(p_i^-), \frac{f(p_i^+) + f(p_i^-)}{2}\right\rangle, \quad i = 0, \ldots, d-1, \tag{14}$$

and

$$\begin{aligned}\left\langle f(p_i^+) - f(p_i^-), f(p_j^+)\right\rangle &= \left\langle f(p_i^+) - f(p_i^-), \frac{f(p_i^+) + f(p_i^-)}{2}\right\rangle, \quad i \neq j \in \{0, \ldots, d-1\},\\ \left\langle f(p_i^+) - f(p_i^-), f(p_j^-)\right\rangle &= \left\langle f(p_i^+) - f(p_i^-), \frac{f(p_i^+) + f(p_i^-)}{2}\right\rangle, \quad i \neq j \in \{0, \ldots, d-1\}.\end{aligned} \tag{15}$$

We show that under the conditions (15) the point $f(m) = (f(p_0^+) + f(p_0^-))/2 = (f(x_0) + f(y_0))/2$ is the unique solution to (14):

1. $f(m) = (f(p_0^+) + f(p_0^-))/2$ is a solution to (14): Choose $j = 0$ ($i \in \{1, \ldots, d-1\}$ arbitrary) in (15). Add the first line in (15) to the second and divide by two. Hence, $f(m) = (f(p_0^+) + f(p_0^-))/2$ is a solution to (14) for $i = 1, \ldots, d-1$ and obviously also for $i = 0$.

2. there is a unique solution to (14): (14) is a linear system involving d equations for the d unknown coordinates of $f(m)$. It suffices to show that the vectors $f(p_i^+) - f(p_i^-)$, $i = 0, \ldots, d-1$, are linearly independent. Subtracting the two lines of (15) yields

$$\left\langle f(p_i^+) - f(p_i^-), f(p_j^+) - f(p_j^-) \right\rangle = 0, \quad i \neq j \in \{0, \ldots, d-1\}.$$

   We see that the vectors $(f(p_i^+) - f(p_i^-))$, $i = 0, \ldots, d-1$, even form an orthogonal system.

Hence, we have $f(m) = (f(x_0)+f(y_0))/2$ and can conclude that $\mu\left(\operatorname{diam}\Omega/2\right) = \operatorname{diam} f(\Omega)/2$.

By repeating this procedure (once starting with $x_0 = x_0$, $y_0 = m$, once starting with $x_0 = m$, $y_0 = y_0$) we see that

$$\mu\left(\frac{1}{4}\operatorname{diam}\Omega\right) = \frac{1}{4}\operatorname{diam} f(\Omega) \quad \text{and} \quad \mu\left(\frac{3}{4}\operatorname{diam}\Omega\right) = \frac{3}{4}\operatorname{diam} f(\Omega)$$

and in general

$$\mu\left(\frac{j}{2^i}\operatorname{diam}\Omega\right) = \frac{j}{2^i}\operatorname{diam} f(\Omega), \quad i \in \mathbb{N}, j \in \{0, \ldots, 2^i\}.$$

Note that Ω being a ball allows us to find a proper ε in each iteration step. By continuity, this shows

$$\mu(t) = t\,\frac{\operatorname{diam} f(\Omega)}{\operatorname{diam}\Omega}.$$

■

Proof of Lemma 9 We want to prove Lemma 9 in the following slightly more general form:

Let $N \in \mathbb{N}$ and $(\varepsilon_k)_{1\leq k\leq N}$, $(\delta_k)_{1\leq k\leq N}$ be finite sequences of positive real numbers satisfying

$$\varepsilon_k < \varepsilon_{k+1}, \quad \delta_k \geq \delta_{k+1} \quad \varepsilon_k > \varepsilon_j + \delta_j,\ j < k, \quad \varepsilon_N + \delta_N + \max_{j=1,\ldots,N-1}(\varepsilon_j + \delta_j) < \frac{1}{2^N}. \quad (16)$$

For $k \in \{1, \ldots, N\}$ and $i \in \{1, 3, \ldots, 2^k - 1\}$ set $x_{k,i} = i/2^k$ and let $y_{k,i}^l$, $y_{k,i}^r$ be arbitrary elements of $(x_{k,i} - \varepsilon_k - \delta_k, x_{k,i} - \varepsilon_k)$ and $(x_{k,i} + \varepsilon_k, x_{k,i} + \varepsilon_k + \delta_k)$, respectively.

Let $\varphi : \{0,1\} \cup \{y_{k,i}^m : m \in \{l,r\}, k \leq N, i \in \{1, 3, \ldots, 2^k - 1\}\} \to [0,1]$ be a weakly isotonic function with $\varphi(0) = 0$ and $\varphi(1) = 1$. Then it holds for $k \leq N$ and $i \in \{1, 3, \ldots, 2^k - 1\}$ that

$$\varphi(y_{k,i}^l) \in \left(\frac{2^{N-k}i - 1}{2^N}, \frac{2^{N-k}i}{2^N}\right), \quad \varphi(y_{k,i}^r) \in \left(\frac{2^{N-k}i}{2^N}, \frac{2^{N-k}i + 1}{2^N}\right), \quad (17)$$

and hence

$$\left|y^m_{k,i} - \varphi(y^m_{k,i})\right| < \frac{1}{2^N}, \quad m \in \{l,r\}, k \le N, i \in \{1,3,\dots,2^k-1\}. \tag{18}$$

Due to (16) we have

$$y^l_{k,i} \in \left(\frac{2^{N-k}i-1}{2^N}, \frac{2^{N-k}i}{2^N}\right), \quad y^r_{k,i} \in \left(\frac{2^{N-k}i}{2^N}, \frac{2^{N-k}i+1}{2^N}\right),$$

and thus (18) follows from (17). We prove (17) by induction over N. Note that φ is strictly increasing due to $\varphi(0)=0$ and $\varphi(1)=1$.

For the basis let $N=1$. Then we have $y^l_{1,1} \in (0,1/2)$ and $y^r_{1,1} \in (1/2,1)$ implying that $|0-y^l_{1,1}| < |1-y^l_{1,1}|$ and $|0-y^r_{1,1}| > |1-y^r_{1,1}|$. Since φ is weakly isotonic, $\varphi(0)=0$ and $\varphi(1)=1$, it follows that $|0-\varphi(y^l_{1,1})| < |1-\varphi(y^l_{1,1})|$ and $|0-\varphi(y^r_{1,1})| > |1-\varphi(y^r_{1,1})|$ and hence

$$\varphi(y^l_{1,1}) \in \left(0, \frac{1}{2}\right), \quad \varphi(y^r_{1,1}) \in \left(\frac{1}{2}, 1\right).$$

Assume that the statement holds for N and we want to infer that it also holds for $N+1$. If the assumptions of the lemma are satisfied for $N+1$, $(\varepsilon_k)_{1\le k\le N}$, $(\delta_k)_{1\le k\le N}$ and $\varphi|_{\{0,1\}\cup\{y^m_{k,i}: m\in\{l,r\}, k\le N, i\in\{1,3,\dots,2^k-1\}\}}$ satisfy the assumptions with N, and hence the induction hypothesis yields for $k \le N$ and $i \in \{1,3,\dots,2^k-1\}$

$$\varphi(y^l_{k,i}) \in \left(\frac{2^{N-k}i-1}{2^N}, \frac{2^{N-k}i}{2^N}\right), \quad \varphi(y^r_{k,i}) \in \left(\frac{2^{N-k}i}{2^N}, \frac{2^{N-k}i+1}{2^N}\right).$$

First, consider $y^l_{N+1,1}$ and $y^r_{N+1,1}$. Due to (16) we have

$$\frac{1}{2^{N+1}} + \varepsilon_{N+1} + \delta_{N+1} < \frac{1}{2^N} - \varepsilon_N - \delta_N$$

and hence

$$0 < y^l_{N+1,1} < y^r_{N+1,1} < y^l_{N,1} < y^r_{N,1}.$$

We have

$$|y^l_{N+1,1} - 0| < \frac{1}{2^{N+1}} - \varepsilon_{N+1}$$

and

$$|y^l_{N,1} - y^l_{N+1,1}| > \frac{1}{2^N} - \varepsilon_N - \delta_N - \left(\frac{1}{2^{N+1}} - \varepsilon_{N+1}\right) = \frac{1}{2^{N+1}} + \varepsilon_{N+1} - \varepsilon_N - \delta_N.$$

Because of $\varepsilon_N + \delta_N < 2\varepsilon_{N+1}$ according to (16), this yields

$$|y^l_{N+1,1} - 0| < |y^l_{N,1} - y^l_{N+1,1}|$$

implying that

$$|\varphi(y^l_{N+1,1}) - 0| < |\varphi(y^l_{N,1}) - \varphi(y^l_{N+1,1})|$$

and

$$2\varphi(y^l_{N+1,1}) < \varphi(y^l_{N,1}),$$

respectively. Due to the induction hypothesis we finally obtain

$$\varphi(y^l_{N+1,1}) < \frac{1}{2}\varphi(y^l_{N,1}) < \frac{1}{2}\frac{1}{2^N} = \frac{1}{2^{N+1}}$$

and hence

$$\varphi(y^l_{N+1,1}) \in \left(0, \frac{1}{2^{N+1}}\right).$$

We have

$$|y^r_{N+1,1} - 0| > \frac{1}{2^{N+1}} + \varepsilon_{N+1}$$

and

$$|y^r_{N,1} - y^r_{N+1,1}| < \frac{1}{2^N} + \varepsilon_N + \delta_N - \left(\frac{1}{2^{N+1}} + \varepsilon_{N+1}\right) = \frac{1}{2^{N+1}} + \varepsilon_N - \varepsilon_{N+1} + \delta_N$$

implying that (due to (16))

$$|y^r_{N,1} - y^r_{N+1,1}| < |y^r_{N+1,1} - 0|.$$

It follows that

$$\frac{1}{2^{N+1}} < \frac{1}{2}\varphi(y^r_{N,1}) < \varphi(y^r_{N+1,1}).$$

Because of

$$y^r_{N+1,1} < y^l_{N,1},$$

we have

$$\varphi(y^r_{N+1,1}) < \varphi(y^l_{N,1}) < \frac{1}{2^N} = \frac{2}{2^{N+1}}$$

and hence

$$\varphi(y^r_{N+1,1}) \in \left(\frac{1}{2^{N+1}}, \frac{2}{2^{N+1}}\right).$$

We also obtain

$$\varphi(y^l_{N,1}) \in \left(\frac{1}{2^{N+1}}, \frac{2}{2^{N+1}}\right).$$

In the same manner one can show (17) for $y^l_{N+1,2^{N+1}-1}$, $y^r_{N+1,2^{N+1}-1}$ and $y^r_{N,2^N-1}$.

Now, let $i \in \{3, 5, \ldots, 2^{N+1}-3\}$ be arbitrary. Consider the reduced fractions $\frac{i-1}{2^{N+1}} = \frac{j_1}{2^{k_1}}$ and $\frac{i+1}{2^{N+1}} = \frac{j_2}{2^{k_2}}$ with $1 \leq k_1, k_2 \leq N$ and $j_1 \in \{1, 3, \ldots, 2^{k_1}-1\}$, $j_2 \in \{1, 3, \ldots, 2^{k_2}-1\}$. Due to (16) we have

$$y^l_{k_1,j_1} < y^r_{k_1,j_1} < y^l_{N+1,i} < y^r_{N+1,i} < y^l_{k_2,j_2} < y^r_{k_2,j_2}.$$

We have to show (17) for $y^r_{k_1,j_1}$, $y^l_{N+1,i}$, $y^r_{N+1,i}$ and $y^l_{k_2,j_2}$.

We have

$$|y^l_{k_1,j_1} - y^l_{N+1,i}| < \frac{i}{2^{N+1}} - \varepsilon_{N+1} - \left(\frac{i-1}{2^{N+1}} - \varepsilon_{k_1} - \delta_{k_1}\right) = \frac{1}{2^{N+1}} - \varepsilon_{N+1} + \varepsilon_{k_1} + \delta_{k_1}$$

and

$$|y^l_{k_2,j_2} - y^l_{N+1,i}| > \frac{i+1}{2^{N+1}} - \varepsilon_{k_2} - \delta_{k_2} - \left(\frac{i}{2^{N+1}} - \varepsilon_{N+1}\right) = \frac{1}{2^{N+1}} - \varepsilon_{k_2} + \varepsilon_{N+1} - \delta_{k_2}.$$

Since $\delta_{k_1} + \delta_{k_2} + \varepsilon_{k_1} + \varepsilon_{k_2} < 2\varepsilon_{N+1}$ according to (16), this yields

$$|y^l_{k_1,j_1} - y^l_{N+1,i}| < |y^l_{k_2,j_2} - y^l_{N+1,i}|$$

and hence

$$|\varphi(y^l_{k_1,j_1}) - \varphi(y^l_{N+1,i})| < |\varphi(y^l_{k_2,j_2}) - \varphi(y^l_{N+1,i})|.$$

Using the induction hypothesis we can conclude that

$$\varphi(y^l_{N+1,i}) < \frac{\varphi(y^l_{k_2,j_2}) + \varphi(y^l_{k_1,j_1})}{2} < \frac{1}{2}\left(\frac{i+1}{2^{N+1}} + \frac{i-1}{2^{N+1}}\right) = \frac{i}{2^{N+1}}.$$

The induction hypothesis also yields

$$\varphi(y^r_{k_1,j_1}) > \frac{i-1}{2^{N+1}}$$

and hence we have

$$\varphi(y^r_{k_1,j_1}) \in \left(\frac{i-1}{2^{N+1}}, \frac{i}{2^{N+1}}\right), \quad \varphi(y^l_{N+1,i}) \in \left(\frac{i-1}{2^{N+1}}, \frac{i}{2^{N+1}}\right).$$

We have

$$|y^r_{k_1,j_1} - y^r_{N+1,i}| > \frac{i}{2^{N+1}} + \varepsilon_{N+1} - \left(\frac{i-1}{2^{N+1}} + \varepsilon_{k_1} + \delta_{k_1}\right) = \frac{1}{2^{N+1}} + \varepsilon_{N+1} - \varepsilon_{k_1} - \delta_{k_1},$$

$$|y^r_{k_2,j_2} - y^r_{N+1,i}| < \frac{i+1}{2^{N+1}} + \varepsilon_{k_2} + \delta_{k_2} - \left(\frac{i}{2^{N+1}} + \varepsilon_{N+1}\right) = \frac{1}{2^{N+1}} - \varepsilon_{N+1} + \varepsilon_{k_2} + \delta_{k_2}$$

and hence (due to (16))

$$|y^r_{k_2,j_2} - y^r_{N+1,i}| < |y^r_{k_1,j_1} - y^r_{N+1,i}|.$$

In the same manner as above we can conclude that

$$\varphi(y^r_{N+1,i}) \in \left(\frac{i}{2^{N+1}}, \frac{i+1}{2^{N+1}}\right), \quad \varphi(y^l_{k_2,j_2}) \in \left(\frac{i}{2^{N+1}}, \frac{i+1}{2^{N+1}}\right).$$

■

Remark 13

- *The assumptions (16) on the sequences $(\varepsilon_k)_{1\leq k\leq N}$ and $(\delta_k)_{1\leq k\leq N}$ are equivalent to $\varepsilon_k > 0$, $\delta_k > 0$, $\delta_k \geq \delta_{k+1}$, $\varepsilon_k + \delta_k < \varepsilon_{k+1}$ and $\varepsilon_N + \delta_N + \varepsilon_{N-1} + \delta_{N-1} < 1/2^N$.*
- *Sequences $(\varepsilon_k)_{1\leq k\leq N}$ and $(\delta_k)_{1\leq k\leq N}$ satisfying these assumptions always exist. For example, we can choose $\varepsilon_k = \varepsilon_1 2^{k-1}$ and $\delta_k = \varepsilon_1/2$ with $\varepsilon_1 < 1/2^{2N+1}$ as in Section 5.*

Proof of Part 2 of Theorem 4 Since $\cup_{i=1}^k K_i^\circ$ is assumed to be connected, we can write K as $K = \cup_{i=1}^{k'} K_i'$ with $K_i' \in \{K_1, \ldots, K_k\}$ and such that ${K_i'}^\circ \cap {K_{i+1}'}^\circ \neq \emptyset$, $i = 1, \ldots, k' - 1$. Hence, w.l.o.g. we may assume that $K = \cup_{i=1}^k K_i$ such that $K_i^\circ \cap K_{i+1}^\circ \neq \emptyset$, $i = 1, \ldots, k-1$.

For every $i \in \{1, \ldots, k\}$ it holds that $\{x_n : n \in \mathbb{N}\} \cap K_i$ is dense in K_i because of $K_i \subseteq \overline{K_i^\circ}$, and hence there exists a sequence $(S_n^i)_{n\in\mathbb{N}}$ of similarity transformations such that

$$\|S_n^i - \varphi_n|_{\{x_1,\ldots,x_n\}\cap K_i}\|_{\infty(\{x_1,\ldots,x_n\}\cap K_i)} \to 0. \tag{19}$$

We prove that

$$\|S_n^1 - \varphi_n|_{\{x_1,\ldots,x_n\}\cap(K_1\cup\ldots\cup K_j)}\|_{\infty(\{x_1,\ldots,x_n\}\cap(K_1\cup\ldots\cup K_j))} \to 0, \quad j = 1, \ldots, k, \tag{20}$$

by induction over j. The basis is clear. For the inductive step assume that (20) holds for some $j < k$. We have to infer that it also holds for $j + 1$. So let $\varepsilon > 0$ be arbitrary.

Since $K_j^\circ \cap K_{j+1}^\circ \neq \emptyset$, we can choose $\widehat{N} \in \mathbb{N}$ such that $\{x_1, \ldots, x_{\widehat{N}}\}$ contains $d + 1$ affinely independent points of $K_j \cap K_{j+1}$. Denote these points by $u_1, u_2, \ldots, u_{d+1}$. Any point $u \in \mathbb{R}^d$ can be written as $u = \sum_{i=1}^{d+1} \lambda_i(u) u_i$ for some (unique) coefficients $\lambda_i(u) \in \mathbb{R}$ with $\sum_{i=1}^{d+1} \lambda_i(u) = 1$. However, since K is bounded, there exists $C > 0$ such that $|\lambda_i(u)| \leq C$, $u \in K$, $i \in \{1, \ldots, d+1\}$.

Choose $\tilde{\varepsilon} > 0$ such that $2(d + 1)C\tilde{\varepsilon} + \tilde{\varepsilon} < \varepsilon$. Choose $N_1 \in \mathbb{N}$ such that

$$\|S_n^1 - \varphi_n|_{\{x_1,\ldots,x_n\}\cap(K_1\cup\ldots\cup K_j)}\|_{\infty(\{x_1,\ldots,x_n\}\cap(K_1\cup\ldots\cup K_j))} < \tilde{\varepsilon}, \quad n \geq N_1,$$

which is possible by the induction hypothesis. Choose $N_2 \in \mathbb{N}$ such that

$$\|S_n^{j+1} - \varphi_n|_{\{x_1,\ldots,x_n\}\cap K_{j+1}}\|_{\infty(\{x_1,\ldots,x_n\}\cap K_{j+1})} < \tilde{\varepsilon}, \quad n \geq N_2,$$

which is possible because of (19). For $n \geq \max\{\widehat{N}, N_1, N_2\}$ and $x \in \{x_1, \ldots, x_n\} \cap (K_1 \cup \ldots \cup K_{j+1})$ we then have:

- if $x \in K_1 \cup \ldots \cup K_j$, then clearly $\|S_n^1(x) - \varphi_n(x)\| < \tilde{\varepsilon} < \varepsilon$
- if $x \in K_{j+1}$, then $\|S_n^1(x) - \varphi_n(x)\| \leq \|S_n^1(x) - S_n^{j+1}(x)\| + \|S_n^{j+1}(x) - \varphi_n(x)\| < \|S_n^1(x) - S_n^{j+1}(x)\| + \tilde{\varepsilon}$

  Since $x = \sum_{i=1}^{d+1} \lambda_i u_i$ with $\sum_{i=1}^{d+1} \lambda_i = 1$ and $|\lambda_i| \leq C$, $i = 1, \ldots, d+1$, we have

$$\begin{aligned}
\|S_n^1(x) - S_n^{j+1}(x)\| &= \left\| \sum_{i=1}^{d+1} \lambda_i S_n^1(u_i) - \sum_{i=1}^{d+1} \lambda_i S_n^{j+1}(u_i) \right\| \\
&\leq \sum_{i=1}^{d+1} |\lambda_i| \|S_n^1(u_i) - S_n^{j+1}(u_i)\| \\
&\leq (d+1)C \max_{i=1,\ldots,d+1} \|S_n^1(u_i) - S_n^{j+1}(u_i)\|.
\end{aligned}$$

  Since $\|S_n^1(u_i) - S_n^{j+1}(u_i)\| \leq \|S_n^1(u_i) - \varphi_n(u_i)\| + \|\varphi_n(u_i) - S_n^{j+1}(u_i)\| < 2\tilde{\varepsilon}$ for $i \in \{1, \ldots, d+1\}$, this yields

$$\|S_n^1(x) - \varphi_n(x)\| \leq 2(d+1)C\tilde{\varepsilon} + \tilde{\varepsilon} < \varepsilon.$$

■

Appendix C. Detailed versions of Lemma 11 and Lemma 12

Lemma 11 *Let $d \geq 2$. Let $N \in \mathbb{N}$ such that*

$$\omega = 24 \left(\frac{\Gamma(\frac{d}{2}+1)}{\pi^{\frac{d}{2}}} \right)^{\frac{1}{d}} \left(\frac{1}{2^N - 1} \right)^{\frac{1}{d}} < \frac{1}{2(d-1)}$$

be fixed. Let $r, \tilde{r}, \alpha, \tilde{\alpha} \in \mathbb{R}^d$, let $\mu > 0$ and let $(\varepsilon_k)_{1 \leq k \leq N}$, $(\delta_k)_{1 \leq k \leq N}$ be real sequences with

$$\begin{array}{ll}
r, \tilde{r} > 0, \quad \alpha, \tilde{\alpha} \geq 0, & \varepsilon_k > 0, \quad \delta_k > 0, \\
\alpha < r, \quad \tilde{\alpha} < \tilde{r}, & \varepsilon_{k+1} > \varepsilon_k, \quad \delta_{k+1} \leq \delta_k, \\
r_1 = \tilde{r}_1 = 1, & \delta_1 < \mu, \quad \delta_1 < \varepsilon_1, \\
r_j \leq 1, \quad \tilde{r}_j \leq 1, \quad j = 2, \ldots, d, & \alpha_1 + \delta_1 + d\mu < \varepsilon_1, \\
\tilde{\alpha}_1 < \omega, \quad \max_{s=1,\ldots,d} \tilde{\alpha}_s < \frac{1}{2}, & 4\varepsilon_N + 4\delta_1 + d\mu < \frac{1}{2^N}, \\
\rho = \max_{j=2,\ldots,d} \frac{\tilde{\alpha}_j(\tilde{r}_j + 3\sqrt{d-1})}{\tilde{r}_j - \tilde{\alpha}_j} < \omega, & \varepsilon_{k+1} > \varepsilon_k + 2\delta_1 + d\mu + \alpha_1, \\
\multicolumn{2}{l}{4\mu r_j - 4\alpha_j \sqrt{1 + (d-1)\mu^2} - 4r_j\alpha_j - 4r_j\delta_1 - 4\alpha_j\delta_1 - \alpha_j^2 > 0, \quad j = 2, \ldots, d,}
\end{array} \tag{21}$$

and such that all the balls $U_{k,i}^l$, $U_{k,i}^r$ and $U_{k,i}^j$, which we define below, lie in the convex hull of the points $X_1^+, X_1^-, \ldots, X_d^+, X_d^-$ defined in the next paragraph.

Define the points $m_s^+, m_s^-, \widetilde{m}_s^+, \widetilde{m}_s^- \in \mathbb{R}^d$, $s = 1, \ldots, d$, by

$$\begin{aligned} m_1^+ &= (r_1/0/0/0/\ldots), \quad & m_1^- &= (-r_1/0/0/0/\ldots), \\ \widetilde{m}_1^+ &= (\tilde{r}_1/0/0/0/\ldots), \quad & \widetilde{m}_1^- &= (-\tilde{r}_1/0/0/0/\ldots), \\ m_2^+ &= (0/r_2/0/0/\ldots), \quad & m_2^- &= (0/-r_2/0/0/\ldots), \\ \widetilde{m}_2^+ &= (0/\tilde{r}_2/0/0/\ldots), \quad & \widetilde{m}_2^- &= (0/-\tilde{r}_2/0/0/\ldots), \\ & & & \textit{and so forth.} \end{aligned}$$

Let $X_s^+, X_s^- \in \mathbb{R}^d$, $s = 1, \ldots, d$, be arbitrary elements of $U_{\alpha_s}(m_s^+)$ and $U_{\alpha_s}(m_s^-)$, respectively.

For $k \in \{1, \ldots, N\}$, $i \in \{1, 3, \ldots, 2^k - 1\}$ and $j \in \{2, \ldots, d\}$ set

$$x_{k,i} = -1 + \frac{i}{2^{k-1}}, \qquad o_{k,i}^j = (x_{k,i}/-\mu/\ldots/-/\mu/\underbrace{+\mu}_{jth\ \ entry}/-\mu/\ldots/-\mu) \in \mathbb{R}^d,$$

$$u_{k,i}^l = (x_{k,i} - \varepsilon_k/-\mu/-\mu/\ldots/-\mu) \in \mathbb{R}^d, \quad u_{k,i}^r = (x_{k,i} + \varepsilon_k/-\mu/-\mu/\ldots/-\mu) \in \mathbb{R}^d,$$

and define the open balls

$$U_{k,i}^j = U_{\delta_k}(o_{k,i}^j), \quad U_{k,i}^l = U_{\delta_k}(u_{k,i}^l), \quad U_{k,i}^r = U_{\delta_k}(u_{k,i}^r).$$

Let $z_{k,i}^j$ be an arbitrary element of $U_{k,i}^j$ and $y_{k,i}^l$, $y_{k,i}^r$ be arbitrary elements of $U_{k,i}^l$ and $U_{k,i}^r$, respectively.

Let $\varphi : \{X_1^+, X_1^-, \ldots, X_d^+, X_d^-\} \cup \{z_{k,i}^j : k \leq N, i \in \{1, 3, \ldots, 2^k - 1\}, j \in \{2, \ldots, d\}\} \cup \{y_{k,i}^m : m \in \{l, r\}, k \leq N, i \in \{1, 3, \ldots, 2^k - 1\}\} \to \mathbb{R}^d$ be an isotonic function and assume that

$$\varphi(X_s^+) \in U_{\tilde{\alpha}_s}(\widetilde{m}_s^+), \quad \varphi(X_s^-) \in U_{\tilde{\alpha}_s}(\widetilde{m}_s^-), \quad s = 1, \ldots, d. \tag{$\star$}$$

Set $\gamma(-1) = \gamma(1) = \tilde{\alpha}_1$ and $\gamma(0) = \tilde{\alpha}_1 + \frac{d-1}{2}(\omega + \rho)$, and define for $k \in \{2, \ldots, N\}$ and $i \in \{1, 3, \ldots, 2^k - 1\}$ the positive expression $\gamma(-1 + i/2^{k-1})$ recursively by

$$\gamma\left(-1 + \frac{i}{2^{k-1}}\right) = \frac{1}{2}\left(\gamma\left(-1 + \frac{i-1}{2^{k-1}}\right) + \gamma\left(-1 + \frac{i+1}{2^{k-1}}\right) + (d-1)(\omega + 2\rho)\right).$$

Let $N^* < N$ such that $N^* \cdot 2^{N^*} < \frac{1}{5(d+1)(\omega+\rho+\tilde{\alpha}_1)}$. Then we have

$$\begin{aligned} \varphi(y_{k,i}^l) &\in (x_{k,i} - \gamma(x_{k,i}) - \omega, x_{k,i} + \gamma(x_{k,i})) \times (-\rho - \omega, \rho)^{d-1}, \\ \varphi(y_{k,i}^r) &\in (x_{k,i} - \gamma(x_{k,i}), x_{k,i} + \gamma(x_{k,i}) + \omega) \times (-\rho - \omega, \rho)^{d-1} \end{aligned}$$

and hence

$$\left\|y_{k,i}^m - \varphi\left(y_{k,i}^m\right)\right\| < \gamma(x_{k,i}) + \omega + (d-1)(\omega + \rho) < 3d\sqrt{\omega}, \quad m \in \{l, r\},$$

for all $1 \leq k \leq N^*$ and $i \in \{1, 3, \ldots, 2^k - 1\}$.

Remark 14 *It is straightforward to see that for any $N \in \mathbb{N}$ there exist $r, \tilde{r}, \alpha, \tilde{\alpha} \in \mathbb{R}^d$, a positive constant μ and sequences $(\varepsilon_k)_{1\le k\le N}$, $(\delta_k)_{1\le k\le N}$ satisfying (21) and which have the property that the balls $U^l_{k,i}$, $U^r_{k,i}$ and $U^j_{k,i}$ lie in the convex hull of $X_1^+, X_1^-, \ldots, X_d^+, X_d^-$ for any choice of these points within the balls $U_{\alpha_s}(m_s^+)$ and $U_{\alpha_s}(m_s^-)$, respectively. At the same time, we can choose all components α_s, $\tilde{\alpha}_s$ to be strictly positive.*

Lemma 12 *Let $d \ge 2$. Let $N' \in \mathbb{N}$ such that*

$$\omega' = 32\left(\frac{\Gamma(\frac{d}{2}+1)}{\pi^{\frac{d}{2}}}\right)^{\frac{1}{d}} \frac{1}{\sqrt[d]{N'}} < 1$$

and all denominators of fractions in

$$A(\omega') = \frac{1}{4}d\left(\frac{20\omega' + 8\frac{10d\omega'}{4d-1}(4d)^{d-1}}{2-\omega'-\frac{10d\omega'}{4d-1}(4d)^{d-1}}\right)^2 + 2\sqrt{d}\,\frac{20\omega' + 8\frac{10d\omega'}{4d-1}(4d)^{d-1}}{2-\omega'-\frac{10d\omega'}{4d-1}(4d)^{d-1}} + 5\omega'$$

are larger than one be fixed.

Let $r < 1$ and $\mu, \delta, \varepsilon > 0$ be real numbers such that

$$r \ge \frac{1}{2}, \quad r > 1 - \sqrt{A(\omega')/2}, \quad \sqrt{1+r^2} > \sqrt{2} - \sqrt{A(\omega')}, \quad 2(r+\mu) - 2\varepsilon > \frac{2N'-1}{N'} + \delta,$$
$$2(r+\mu) + 2\varepsilon < 2, \quad \delta < \frac{1}{3N'}, \quad \sqrt{2} + \mu\sqrt{d} + \varepsilon < 2, \quad \delta + 2\mu\sqrt{d} + 2\varepsilon < \frac{1}{N'},$$

$$\sqrt{(1-\mu)^2 + (-\mu\tilde{v} - mr)^2 + (d-2)\mu^2} + \varepsilon < \sqrt{(1+\mu)^2 + (-\mu\tilde{v} - mr)^2 + (d-2)\mu^2} - \varepsilon \quad \text{for } \tilde{v}, m \in \{-1, +1\},$$

$$\sqrt{(-1-\mu\tilde{v})^2 + (r-\mu)^2 + (d-2)\mu^2} + \varepsilon < \sqrt{(-1-\mu\tilde{v})^2 + (r+\mu)^2 + (d-2)\mu^2} - \varepsilon \quad \text{for } \tilde{v} \in \{-1, +1\},$$

$$\sqrt{(r+\mu(\tilde{v}-\tilde{v}'))^2 + (r-2\mu)^2 + \sum_{k=1}^{d-2}\mu^2(v_k - v'_k)^2} + 2\varepsilon < \sqrt{(r+\mu(\tilde{v}-\tilde{v}'))^2 + (r+2\mu)^2 + \sum_{k=1}^{d-2}\mu^2(v_k - v'_k)^2} - 2\varepsilon$$
$$\text{for } \tilde{v}, \tilde{v}', v_k, v'_k \in \{-1, +1\} \ (k = 1, \ldots, d-2). \tag{22}$$

Define points $A, B \in \mathbb{R}^d$ *and* $Z_s^-, Z_s^+ \in \mathbb{R}^d$, $s \in \{2, \dots, d\}$, *by*

$$A = (-1/0/\dots/0), \quad B = (1/0/\dots/0), \quad Z_2^- = (0/-r/0/0/\dots),$$
$$Z_2^+ = (0/r/0/0/\dots), \quad Z_3^- = (0/0/-r/0/\dots), \quad Z_3^+ = (0/0/r/0/\dots), \quad \textit{and so forth.}$$

For $s \in \{2, \dots, d\}$ *and* $v \in \{-1, 1\}^d$ *set* $E_{s,v}^- = Z_s^- + \mu v$, $E_{s,v}^+ = Z_s^+ + \mu v$ *and let* $e_{s,v}^-, e_{s,v}^+ \in \mathbb{R}^d$ *be arbitrary elements of* $U_\varepsilon(E_{s,v}^-)$ *and* $U_\varepsilon(E_{s,v}^+)$*, respectively. For* $i \in \{1, \dots, 2N' - 1\}$ *let* $x_i \in \mathbb{R}^d$ *be an arbitrary element of* $U_\delta((-1 + \frac{i}{N'}/0/\dots/0))$.

Let $\varphi : \{A, B\} \cup \{e_{s,v}^-, e_{s,v}^+ : s \in \{2, \dots, d\}, v \in \{-1, 1\}^d\} \cup \{x_i : i = 1, \dots, 2N' - 1\} \to \mathbb{R}^d$ *be an isotonic function with* $\|\varphi(A) - \varphi(B)\| = 2$*. Then there exist a constant* C *(depending only on* d*) and an isometry* $S : \mathbb{R}^d \to \mathbb{R}^d$ *such that*

$$\|A - S(\varphi(A))\| \leq C\sqrt{A(\omega')}, \quad \|B - S(\varphi(B))\| \leq C\sqrt{A(\omega')},$$
$$\|Z_s^m - S(\varphi(e_{s,\underline{v}}^m))\| \leq C\sqrt{A(\omega')}, \quad m \in \{-, +\}, s \in \{2, \dots, d\},$$

where $\underline{v} = (1/1/1/\dots/1)$.

Remark 15 *For any* $N' \in \mathbb{N}$ *there exist real numbers* $r < 1$ *and* $\mu, \delta, \varepsilon > 0$ *satisfying* (22)*. We can even choose them in such a way that all the considered open balls are contained in* $U_1(0)$.