Machine learning on rigid classes of Euclidean clouds of unordered points

Abstract

Most real objects allow infinitely many different representations. Robust machine learning aims to use only invariant features independent of object representations to guarantee that any output (class label or predicted property) is preserved if the same object is represented differently. For Euclidean clouds of unordered points under rigid motion, we introduce complete invariants (with no false negatives, no false positives) and a Lipschitz continuous distance that satisfies all metric axioms and is computable in polynomial time of the number of points. The new realizability property implies that the space of all rigid clouds is efficiently parametrized by vectorial invariants like geographic coordinates. The proposed invariants distinguished all rigid classes of atomic clouds in the world's largest collections of molecules with 3D coordinates and predicted chemical elements by pure geometry with over 98% accuracy.

1. Importance of complete and bi-continuous invariants for ML on data with real values

This paper formalizes practically important conditions for application-driven ML on real objects with ambiguous representations and develops new canonical representations satisfying these conditions for any *clouds* (finite sets) of unordered points in Euclidean space $\mathbb{R}^n$. Such a cloud is the most basic form of a real object from cars to molecules (Wang & Solomon, 2019), e.g. a set of corners or atoms.

Many objects are *rigid* in the sense that their shape and properties are preserved under *rigid motion* composed of translations and rotations in $\mathbb{R}^n$ (Atz et al., 2021), which form the group $\mathrm{SE}(n)$. The slightly weaker relation is by *isometries* (distance-preserving transformations), which form the group $\mathrm{E}(n)$. The practical cases are dimensions $n \leq 3$ and larger numbers m (hundreds) of unordered points without outliers (Shi et al., 2021) because atoms have stable nuclei.

. Correspondence to: Anonymous Author <anon.email@domain.com>.

Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.

Any rigid cloud has infinitely many representations, e.g. lists of point coordinates, but the shape and properties of an object should be independent of a coordinate system. Points are usually unordered and even simple molecules have many indistinguishable atoms. Hence predictions should not depend on point ordering. On another hand, different rigid classes of chemically identical molecules can have different functional properties such as solubility and hence therapeutic effectiveness. If not all rigid classes are distinguished, drugs can become useless, implying human suffering and financial losses for manufacturers (Morissette et al., 2003).

A repeated scan or measurement of the same object can produce a slightly different cloud that cannot be exactly matched with the original one by rigid motion, also due to atomic vibrations (Feynman, 1971). If noise is ignored up to any threshold $\varepsilon > 0$, sufficiently many tiny perturbations make all clouds equivalent by the transitivity axiom: if $A \sim B$ and $B \sim C$, then $A \sim C$ (Brink et al., 1997).

Since all small deviations between rigid classes of point clouds should be distinguished, all these classes live in a continuous space of rigid clouds, see Fig. 1 (left). This space was continuously parametrized only in dimension $n = 1$ or for $m = 3$ points or Fig. 1 (right) leaving other cases open.

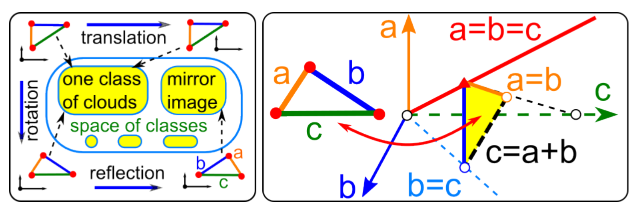

Figure 1. **Left**: rigid classes of m unordered points in $\mathbb{R}^n$ form a continuous space, which had no complete and bi-continuous invariants for $m > 3$, $n > 1$. **Right**: the space of 3 points under isometry is parametrized by distances $0 < a \leq b \leq c \leq a + b$.

Machine learning previously focused on discrete classifications or success measures for finite datasets, which can be considered discrete samples (of measure 0) in continuous spaces. For generalizability to all real data outside finite datasets, application-driven ML needs new conditions formalized in Problem 1.1 below. (Li et al., 2021; Dym & Gortler, 2024; Maennel et al., 2024; Nigam et al., 2024) studied complete invariants without realizability and Lipschitz bi-continuity (Morris et al., 2024; Cahill et al., 2024).

Problem 1.1. *Find a complete and bi-continuous invariant $I : \{$clouds of unordered points in $\mathbb{R}^n\} \to$ a space X with a distance d such that all the conditions below hold.*

***(a) Completeness**: any clouds A, B of unordered points are related by a rigid motion of $\mathbb{R}^n$ if and only if $I(A) = I(B)$.*

***(b) Metric axioms**: 1) $d(\alpha, \beta) = 0 \Leftrightarrow \alpha = \beta$; 2) $d(\alpha, \beta) = d(\beta, \alpha)$; 3) $d(\alpha, \beta) + d(\beta, \gamma) \geq d(\alpha, \gamma)$ for all $\alpha, \beta, \gamma \in X$.*

***(c) Lipschitz continuity**: there is a constant λ such that if each point of a cloud $A \subset \mathbb{R}^n$ is perturbed up to Euclidean distance ε, then $I(A)$ changes by at most $\lambda\varepsilon$ in the metric d.*

***(d) Realizability**: the image $\{I(A) \mid$ clouds $A \subset \mathbb{R}^n$ of unordered points$\}$ is parametrized so that one can reconstruct A up to rigid motion from any* realizable *value of I.*

***(e) Point matching**: there is a constant μ that guarantees for any clouds A, B a rigid motion matching all points of A, B up to Euclidean distance $\mu d(I(A), I(B))$.*

***(f) Computability**: for a fixed dimension n, the invariant I, the metric d, and all constructions in (d) and (e) are computable in polynomial time of the number of points.*

Clouds and rigid motion can be replaced with any data (graphs, meshes) and equivalences (also allowing reflections or uniform scaling), respectively, so Problem 1.1 makes sense for any real data with ambiguous representations.

The completeness (or injectivity) in 1.1(a) fully answers the question "same or different?" A complete invariant I has the ultimate expressive power and always distinguishes all clouds $A \not\cong B$ (not only from a finite dataset) that cannot be matched by rigid motion, so I is a descriptor with *no false negatives* and *no false positives*. The universal approximation aims for the completeness of infinite-size invariants (Maron et al., 2019; Keriven & Peyré, 2019; Yarotsky, 2022), so polynomial time in 1.1(f) makes all conditions harder.

A complete invariant can give a discontinuous metric, say $d(A, B) = 1$ for all non-equivalent clouds without quantifying the similarity of near-duplicates. The continuity in 1.1(c) is necessary for smoothness and hence for any gradient-based optimisation Due to the first axiom in 1.1(b), any metric d detects rigidly equivalent clouds by checking if $d(A, B) = 0$. Without the first axiom, many more distances including the zero $d \equiv 0$ satisfy the other axioms and are called *pseudo-metrics* (Brécheteau, 2019). If the third axiom in 1.1(b) fails with any additive error $\varepsilon > 0$, results of clustering may not be trustworthy (Rass et al., 2024).

The realizability in 1.1(d) implies that the invariant I is an invertible 1-1 map from the complicated *Cloud Rigid Space* $\mathrm{CRS}(\mathbb{R}^n; m)$ of classes of clouds under rigid motion to the explicitly parametrized space $I(\mathrm{CRS}(\mathbb{R}^n; m))$ of realizable values. Then with 100% certainty, we can sample any value in $I(\mathrm{CRS}(\mathbb{R}^n; m))$ and reconstruct its cloud $A \subset \mathbb{R}^n$.

The 1-1 point matching in 1.1(e) can be interpreted as the Lipschitz continuity of the inverse map I^{-1} so that any close values $I(A), I(B)$ guarantee the closeness of A, B under rigid motion. Conditions 1.1(c,e) mean that I is bi-Lipschitz: $\varepsilon/\mu \leq d(I(A), I(B)) \leq \lambda\varepsilon$, where ε is the minimum perturbation needed to match all points of A, B.

A partial matching, e.g. ignoring outliers, is harder to formalize. Indeed, if any clouds sharing all points except one are called equivalent, the transitivity axiom allows us to build a chain of equivalences $A_1 \sim \cdots \sim A_k$ changing one point at a time, which can make all clouds equivalent.

One can define metrics satisfying 1.1(a,b,c) by minimizing or deviations of unordered points over infinitely many transformations but polynomial time in 1.1(f) makes Problem 1.1 notoriously hard, previously solved only for $m = 3$ points.

Conditions 1.1(a,b,c,f) and 1.1(d,e,f) formalize the *discriminative* and *generative* goals, respectively. A full solution to Problem 1.1 will imply that the rigid classes of clouds can be efficiently visualized in the *moduli* space $I(\mathrm{CRS}(\mathbb{R}^n; m))$ replacing any latent space of non-invariants or incomplete (or discontinuous or non-realizable) invariants. Geographically, $I(\mathrm{CRS}(\mathbb{R}^n; m))$ can be compared with Earth's map, where any location can be reconstructed with all properties (altitude, precipitation, images, ...) from the latitude and longitude coordinates in known (realizable) ranges.

Contributions. Problem 1.1 formalizes the necessary conditions for any application-driven ML on real objects. The new invariant Nested Distributed Projection solves Problem 1.1 for all clouds of m unordered points in dimension $n = 2$. Any cloud $A \subset \mathbb{R}^n$ can be reconstructed from a small part of the invariant (a vector in $\mathbb{R}^{n(m-(n+1)/2)}$) whose realizability in 1.1(d) is guaranteed by explicitly written inequalities. Hence coordinates of this vector can be chosen in known ranges like latitude and longitude on Earth maps. The appendices cover all dimensions $n > 2$. The Python/C++ code is in the supplementary materials.

2. Past work on continuous metrics for clouds

Ordered points. Kendall's shape theory (Kendall et al., 2009) studies m ordered points $p_1, \ldots, p_m \in \mathbb{R}^n$ under isometries from $\mathrm{E}(n)$. In this case, a complete invariant is the distance matrix (Schoenberg, 1935; Kruskal & Wish, 1978) or the Gram matrix of scalar products $p_i \cdot p_j$, see chapter 2.9 in (Weyl, 1946), (Villar et al., 2021). A brute-force extension to m unordered points requires $m!$ matrices due to $m!$ permutations, which is ruled out by 1.1(f).

Point cloud registration for unordered points samples rotations (Lin et al., 1986; Yang et al., 2020) and uses scale-invariant features (Lowe, 1999; 2004; Huang et al., 2006) to approximately match clouds. If approximately matched clouds are called equivalent, sufficiently many gradual perturbations make all clouds equivalent due to the transitivity axiom. Hence all rigid classes should be distinguished by a distance d that becomes zero only on rigidly equivalent clouds. Trying to sort points along a fixed direction or in a clockwise order around their center of mass leads to discontinuities because distant points can have equal projections to a line or a circle. A basis (say, of principal directions) of a cloud (Spezialetti et al., 2019; Zhu et al., 2022; Kurlin, 2024) is similarly unstable under perturbations of points in cases of high symmetry, e.g. when eigenvalues become equal, which often happens for real molecules in our main application. Converting a cloud by using extra parameters into a more complex object such as a continuous field $\mathbb{R}^3 \to \mathbb{R}$ (Chauvin et al., 2022) or the persistent homology transform leads to the harder analog of Problem 1.1 for continuous surfaces instead of discrete clouds (Turner et al., 2014).

Neural networks (Bronstein et al., 2021) can guarantee invariance or equivariance (Thomas et al., 2018; Kondor & Trivedi, 2018; Cohen et al., 2019; Fuchs et al., 2020; Deng et al., 2021). An *equivariant* descriptor E satisfies the weaker condition $E(f(A)) = T_f(E(A))$ for any rigid motion f of a cloud A, where T_f may not be the identity as required for invariants (Satorras et al., 2021; Chen et al., 2021; Aronsson, 2022; Assaad et al., 2023; Xu et al., 2022; Su et al., 2022). Any linear combination of points such as the center of mass is equivariant but cannot distinguish clouds under translation. Equivariants were used for predicting forces acting on atoms to move them to a more optimal configuration. These time-dependent clouds A_t can be studied directly by their invariant values $I(A_t)$ without intermediate forces. So neural networks optimize millions of parameters, see Table 4 in (Goyal et al., 2021), to improve accuracies (Dong et al., 2018; Akhtar & Mian, 2018; Laidlaw & Feizi, 2019; Guo et al., 2019; Colbrook et al., 2022) but need re-training any for new data and will have better generalizability if their inputs are invariants satisfying the conditions of Problem 1.1 for all possible clouds in $\mathbb{R}^n$.

General metrics between fixed clouds extend to their rigid classes by minimization over infinitely many rigid motions (Huttenlocher et al., 1993; Chew & Kedem, 1992; Chew et al., 1999). In $\mathbb{R}^2$, the time $O(m^5 \log m)$ (Chew et al., 1997) for the Hausdorff distance (Hausdorff, 1919) will be improved in Theorem 5.3 to $O(m^{3.5} \log m)$ for a new metric, see approximations in (Goodrich et al., 1999). The Gromov-Hausdorff and Gromov-Wasserstein metrics (Mémoli, 2011) are defined for metric-measure spaces also by minimizing over infinitely many correspondences between points, but cannot be approximated with a factor less than 3 in polynomial time unless P=NP, see Corollary 3.8 in (Schmiedl, 2017) and polynomial algorithms for partial cases in (Majhi et al., 2024). Also, computing a metric between rigid classes of clouds is only a small part of Problem 1.1. Indeed, to efficiently navigate on a real planet, in addition to distances between cities, we need a satellite-type view of the whole planet and hence a realizable bi-continuous invariant I, which can be considered an analog of the latitude and longitude coordinates on Earth.

Can we 'sense' a shape? Problem 1.1 asks the questions 'same or different clouds, and how much different?' The related problem 'Can we hear the shape of a drum?' (Kac, 1966) has the negative answer in terms of 2D polygons indistinguishable by spectral invariants (Gordon et al., 1992a;b; Reuter et al., 2006; Cosmo et al., 2019; Marin et al., 2021). Problem 1.1 looks for stronger invariants that can completely 'sense' (not only 'hear') all rigid clouds in any $\mathbb{R}^n$.

The partial cases when Problem 1.1 was solved are only $n = 1$ or $m \le 3$. In dimension $n = 1$, any rigid motion of $\mathbb{R}$ is a translation, so the Cloud Rigid Space $\mathrm{CRS}(\mathbb{R}; m)$ of m points $p_1, \dots, p_m \in \mathbb{R}$ is the space $\mathbb{R}_+^{m-1}$ of sequential inter-point distances $d_i = p_{i+1} - p_i > 0$ for $i = 1, \dots, m-1$. Including reflections, the *Cloud Isometry Space* $\mathrm{CIS}(\mathbb{R}; m)$ is the quotient of $\mathbb{R}_+^{m-1}$ under the cyclic equivalence $(d_1, \dots, d_{m-1}) \sim (d_{m-1}, \dots, d_1)$. For clouds of $m = 2$ points in any dimension $n \ge 1$, $\mathrm{CRS}(\mathbb{R}^n; 2)$ is parametrized by a single inter-point distance $d > 0$. The final known case is $m = 3$ due to the SSS theorem saying that any triangles are congruent (isometric) if and only if they have the same side lengths. The space $\mathrm{CIS}(\mathbb{R}^n; 3)$ of 3-point clouds has the geographic-style parametrization $\{0 < a \le b \le c \le a + b\}$ by inter-point distances a, b, c so that any $(a, b, c) \in \mathrm{CIS}(\mathbb{R}^n; 3)$ generates a uniquely triangle under isometry. Problem 1.1 asks for a similarly explicit parametrization of $\mathrm{CRS}(\mathbb{R}^n; m)$ for all $m \ge 4$ and $n \ge 2$.

Recent advances are the extensions (Delle Rose et al., 2024; Hordan et al., 2024) of the WL test (Leman & Weisfeiler, 1968), giving a binary answer (Brass & Knauer, 2000; 2004) by distinguishing all non-isometric clouds but without Lipschitz continuous metrics for all clouds including degenerate ones. Attempting to extend the SSS theorem, the Sorted Distance Vector (SDV) of all $\frac{m(m-1)}{2}$ distances between $m \ge 4$ unordered points distinguishes all non-isometric clouds in general position in $\mathbb{R}^n$ (Boutin & Kemper, 2004) but not infinitely many 4-point clouds in $\mathbb{R}^2$, see Fig. 2.

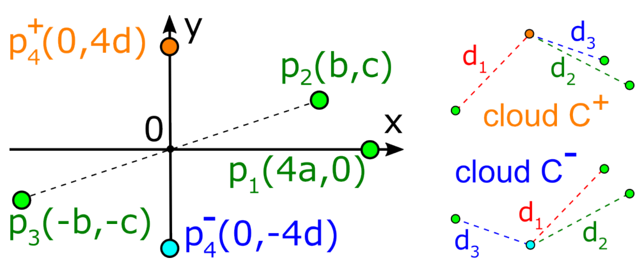

Figure 2. The infinite family of non-isometric clouds $C^+ \not\cong C^-$ sharing p_1, p_2, p_3 and depending on free parameters a, b, c, d.

The SDV was strengthened (Widdowson & Kurlin, 2022) to the Pointwise Distance Distribution (PDD), which still cannot distinguish infinitely many non-isometric clouds in $\mathbb{R}^3$, see Fig. S4 in (Pozdnyakov & Ceriotti, 2022). All these counter-examples were distinguished by the Simplexwise Centered Distributions from (Widdowson & Kurlin, 2023), which satisfy 1.1(a,b,c,f) but not 1.1(d,e). Distance-based invariants do not allow easy realizability already for $m = 4$ points in $\mathbb{R}^2$ whose 6 inter-point distances should satisfy a non-trivial polynomial equation saying that the tetrahedron on 4 points has volume 0 in $\mathbb{R}^2$. Hence random distances between $m > 3$ unordered points are realized by a point cloud in $\mathbb{R}^2$ with probability 0 (Duxbury et al., 2016).

3. Complete invariants of unordered clouds

Any point $p = (x_1, \dots, x_n) \in \mathbb{R}^n$ has *Euclidean* norm $|p| = \sqrt{\sum\limits_{i=1}^{n} x_i^2}$. Any points p and $q = (y_1, \dots, y_n) \in \mathbb{R}^n$ are also interpreted as vectors, have the *Euclidean* distance $|p - q|$ and the *scalar* (dot) product of $p \cdot q = \sum\limits_{i=1}^{n} x_i y_i$. Any vectors $p \perp q$ are *orthogonal* if and only if $p \cdot q = 0$.

While past representations used one basis (say, of principal directions of a given cloud $A \subset \mathbb{R}^n$), this section introduces a new representation based on variable projections that depend on $n-1$ ordered points in C consisting of m unordered points. For simplicity, we consider $n = 2$ when we have only m choices for a single point $p \in A$ in Fig. 3.

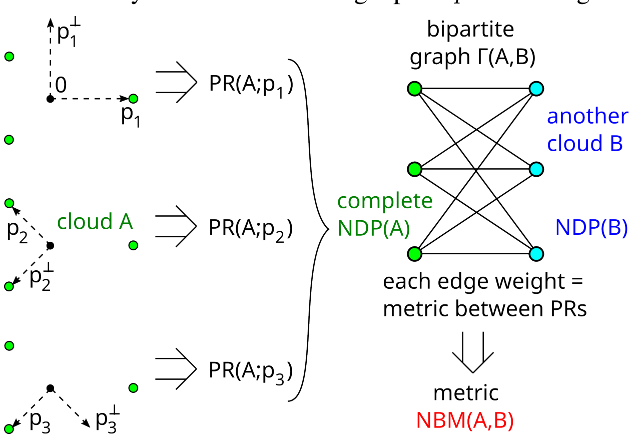

Figure 3. A Point-based Representation (PR) encodes a cloud A in the basis of a point $p \in A$. All PRs are combined into the complete invariant $\mathrm{NDP}(A)$. NDPs are compared by the Nested Bottleneck Metric (NBM) computed from a complete bipartite graph $\Gamma(A, B)$ with weights equal to distances between PRs.

For any cloud $A \subset \mathbb{R}^2$ of m unordered points, the *center of mass* is $O(A) = \dfrac{1}{m}\sum\limits_{p \in A} p$. Shift A so that $O(A)$ is the origin $0 \in \mathbb{R}^2$. For any $p = (x_1, x_2) \in A$, the vector $p^\perp = (-x_2, x_1)$ is orthogonal to p, so $p \cdot p^\perp = 0$, which holds even if $p = 0$. If p is not at the origin (center of mass of A), we use the orthogonal basis $p, p^\perp$ to represent all other points of A. Definition 3.1 makes sense for $p = 0$.

Definition 3.1 (point-based representation $\mathrm{PR}(A;p)$)**.** *Let $A \subset \mathbb{R}^2$ be a cloud with the center of mass at the origin 0. Fix a* base point *$p = (x, y) \in A$, set $p^\perp = (-y, x)$. For any $q \in A \setminus \{p\}$, the $2 \times (m-1)$ matrix $M(A;p)$ has a column of the scalar products $q \cdot p, q \cdot p^\perp$. The* point-based representation *of A is the pair $\mathrm{PR}(A;p) = \big[\, |p|^2,\ M(A;p)\, \big]$.*

We use $|p|^2$ and scalar products to make all components polynomial (smooth) in coordinates. The matrix $M(A;p)$ has two rows (ordered according to $p, p^\perp$) and $m-1$ unordered columns, and can be considered a *fixed cloud* of $m-1$ unordered points in $\mathbb{R}^2$, not under rigid.

Example 3.2 (regular polygons in $\mathbb{R}^2$)**.** ***(a)*** *For $m \geq 2$, let $A_m = \{R \exp \dfrac{2\pi i \sqrt{-1}}{m}\} \subset \mathbb{R}^2$, $i = 1, \dots, m$, be the vertex set of a regular m-sided polygon. Then A_m has the center of mass $O(A_m) = (0, 0)$ at the origin and is inscribed in the circle of the radius $R = R(A_m)$. In Definition 3.1, choose the point $p = (R, 0) \in A_m$, which doesn't affect $\mathrm{PR}(A_m; p)$ due to the rotational symmetry of A_m. Then the matrix $M(A_m;p)$ consists of $m-1$ columns $\begin{pmatrix} R^2 \cos(2\pi i/m) \\ R^2 \sin(2\pi i/m) \end{pmatrix}$, $i = 1, \dots, m-1$. The pair is $\mathrm{PR}(A_m;p) = \Big[R^2, \big(\begin{pmatrix} R^2 \cos \frac{2\pi i}{m} \\ R^2 \sin \frac{2\pi i}{m} \end{pmatrix}_{i=1}^{m-1}\big)\Big]$.*

(b) *Let the cloud $B_m \subset \mathbb{R}^2$ be A_m after adding the extra point at the origin $0 \in \mathbb{R}^2$. For any point $p \in A_m$, the new point-based representation $\mathrm{PR}(B_m;p)$ is obtained from $\mathrm{PR}(A_m;p)$ above by adding the zero column to the matrix $M(A_m;p)$. For the extra point at the origin 0, the representation is $\mathrm{PR}(B_m;0) = [0, M(B_m;0)]$, where $M(B_m;0)$ is the $2 \times m$ matrix consisting of zeros.*

Theorem 3.3 (realizability of abstract PR)**.** *Let $s > 0$ and M be any $2 \times (m-1)$ matrix for $m \geq 2$. The pair $[s, M]$ is realizable as a point-based representation $\mathrm{PR}(A;p)$ for a cloud $A \subset \mathbb{R}^n$ of m unordered points with $O(A) = 0$ and a point $p \in A$ if and only if $s + \sum\limits_{j=1}^{m-1} M_{1j} = 0 = \sum\limits_{j=1}^{m-1} M_{2j}$.*

In Theorem 3.3, $s = |p|^2$ is the squared distance from a point $p \in A$ to $0 \in \mathbb{R}^2$. The equations say that the sums of the scalar products $(q \cdot p)$ and $(q \cdot p^\perp)$ for all $q \in A$ equal to 0, which is equivalent to $\sum q \in A = 0$ meaning that the center of mass $O(A)$ is 0. Hence $s > 0$ and $m-2$ columns of M can be considered free parameters.

Definition 3.4 combines point-based representations $\mathrm{PR}(A;p)$ for all points $p \in A$ into one invariant NDP (Nested Distributed Projection) that will be proved to satisfy all conditions of Problem 1.1. The major advantage of NDP is its applicability to all real clouds $A \subset \mathbb{R}^2$ without any requirement of general position. Some points of a cloud A may coincide, so A can be a multiset of points.

Definition 3.4 (invariants NDP and NCP)**.** *Let $A \subset \mathbb{R}^2$ be any cloud of m unordered points. The* Nested Distributed Projection $\mathrm{NDP}(A)$ *is the unordered set of* $\mathrm{PR}(A;p)$ *for all $p \in A$. If $k > 1$ representations* $\mathrm{PR}(A;p)$ *are equal then we collapse them to one representation with the weight k/m. The resulting set of unordered* PR*s with weights is called the* Nested Compressed Projection $\mathrm{NCP}(A)$.

Table 1. Acronyms and references of all key concepts in the paper.

PR	POINT-BASED REPRESENTATION	DEF 3.1
NDP	NESTED DISTRIBUTED PROJECTION	DEF 3.4
PRM	POINT-BASED REPRESENT. METRIC	DEF 4.2
BMD	BOTTLENECK MATCHING DISTANCE	DEF 4.3
NBM	NESTED BOTTLENECK METRIC	DEF 4.4

For the cloud A_m from Example 3.2, the Nested Distributed Projection $\mathrm{NDP}(A_m)$ consists of m identical representations, so $\mathrm{NCP}(A_m)$ is the single representation $\mathrm{PR}(A_m;p)$ with weight 1. The invariant NDP is an expanded version of the NCP, where all PRs have equal weights $1/m$. The full invariant $\mathrm{NDP}(A)$ includes the faster (linear-time) vector of squared distances $|p|^2$ from the center of mass $O(A) = 0 \in \mathbb{R}^2$ to all points $p \in A$. If A has a distinguished point p, e.g. a special atom in a molecule, the point-based representation $\mathrm{PR}(A;p)$ is invariant.

Theorem 3.5 (completeness of NDP)**.** *The Nested Distributed Projection is complete in the sense that any clouds $A, B \subset \mathbb{R}^2$ of m unordered points are related by rigid motion in $\mathbb{R}^2$ if and only if* $\mathrm{NDP}(A) = \mathrm{NDP}(B)$ *so that there is a bijection* $\mathrm{NDP}(A) \to \mathrm{NDP}(B)$ *matching all* PR*s.*

Under a mirror reflection, for any $p \in A$, one can assume after applying rigid motion that the basis $p, p^{\perp}$ maps to its mirror image $p, -p^{\perp}$. The mirror image $\bar{A}$ has $\mathrm{NDP}(\bar{A})$ equal to $\overline{\mathrm{NDP}}(A)$ that is obtained from $\mathrm{NDP}(A)$ by reversing all signs in the last row of $M(A;p)$ for each $p \in A$.

The completeness of $\mathrm{NDP}(A)$ Theorem 3.5 implies the completeness of the pair $\mathrm{NDP}(A), \overline{\mathrm{NDP}}(A)$ under isometry including reflections. Further work can simplify this pair to a smaller invariant while keeping the completeness. Since a bijection $\mathrm{NDP}(A) \to \mathrm{NDP}(B)$ between all (uncollapsed) PRs induces a bijection $\mathrm{NCP}(A) \to \mathrm{NCP}(B)$ respecting all weights of collapsed PRs, Theorem 3.5 implies the completeness of NCP under rigid motion in $\mathbb{R}^2$.

4. A metric on complete invariants of clouds

This section will define the metric NBM on invariants NDP by using the bottleneck distance BD in Definition 4.1, a metric on point-based representations (PRs) in Definition 4.2, and a bottleneck matching distance in Definition 4.3.

Definition 4.1 (bottleneck distance BD)**.** *For any $v = (v_1, \dots, v_n) \in \mathbb{R}^n$, the* Minkowski *norm is $||v||_\infty = \max\limits_{i=1,\dots,n} |v_i|$. For clouds $A, B \subset \mathbb{R}^n$ of m unordered points, the* bottleneck distance $\mathrm{BD}(A,B) = \inf\limits_{g:A\to B} \sup\limits_{p\in A} ||p - g(p)||_\infty$ *is minimized over all bijections $g : A \to B$.*

Though the bottleneck distance is defined as a minimum for $m!$ bijections $A \to B$ between m-point clouds, Theorem 6.5 in (Efrat et al., 2001) computes $\mathrm{BD}(A,B)$ in time $O(m^{1.5}\log^2 m)$ by filtering out distant points. The brute-force extension of $\mathrm{BD}(A,B)$ under rigid motion need a minimization for infinitely many rotations. $\mathrm{NDP}(A)$ consists of only m point-based representations $\mathrm{PR}(A;p) = [|p|^2, M(A;p)]$, one for each $p \in A$. The BD algorithm can compare any $2 \times (m-1)$ matrices $M(A;p)$ and $M(B;q)$ as fixed clouds of unordered columns (points in $\mathbb{R}^2$).

In Definition 4.2, the notation M/R means that all elements of the matrix $M(A;p)$ are divided by the *radius* $R(A) = \max\limits_{p\in A} |p|$ of a cloud A. Then PRM and further metrics have units of original points, e.g. in meters. One more division by $R(A)$ makes metrics invariant under uniform scaling.

Definition 4.2 (Point-Based Representation Metric PRM)**.** *Let* $\mathrm{PR}(A;p), \mathrm{PR}(B;q)$ *be point-based representations of clouds $A, B \subset \mathbb{R}^2$ of m unordered points for base points $p \in A$ and $q \in B$, respectively, see Definition 3.1. The* Point-based Representation Metric *between the* PR*s above is* $\mathrm{PRM} = \max\{\,|\,|p| - |q|\,|,\ |R(A) - R(B)|,\ w_M\,\}$, *where* $w_M = \mathrm{BD}\left(\dfrac{M(A;p)}{R(A)}, \dfrac{M(B;q)}{R(B)}\right)$, *see Definition 4.1.*

We defined PRM as the maximum of 3 metrics to guarantee the metric axiom (if $\mathrm{PRM} = 0$ then $A \cong B$) and the simplest Lipschitz constant $\lambda = 2$ in 1.1(d), see all proofs in appendix D. Replacing the maximum with (say) a sum gives a metric with a higher constant λ depending on m.

Definition 4.3 (bottleneck matching distance $\mathrm{BMD}(\Gamma)$)**.** *Let Γ be a complete bipartite graph with m white vertices and m black vertices so that every white vertex is connected to every black vertex by an edge e of a weight $w(e) \geq 0$. A* vertex matching *in Γ is a set E of m disjoint edges of Γ. The* weight $W(E) = \max\limits_{e\in E} w(e)$ *is the largest weight in E. The* bottleneck matching distance *of the graph Γ is* $\mathrm{BMD}(\Gamma) = \min\limits_{E} W(E)$ *is minimized over all vertex matchings.*

Because Γ is bipartite, any edge from a vertex matching E joins a white vertex with a black vertex. Then $\mathrm{BMD}(\Gamma)$ is minimized for all bijections E between all white vertices and all black vertices of Γ similar to Definition 4.1. Definition 4.4 builds a graph $\Gamma(A,B)$ on all point-based representations of $A, B \subset \mathbb{R}^n$ and introduces the Nested Bottleneck Metric $\mathrm{NBM}(A,B)$ as BMD of $\Gamma(A,B)$.

Definition 4.4 (NBM : Nested Bottleneck Metric). *Let clouds $A, B \subset \mathbb{R}^2$ consist of m unordered points. The complete bipartite graph $\Gamma(A, B)$ has m white vertices (one for each $p \in A$) and m black vertices (one for each $q \in B$). Any edge e of $\Gamma(A, B)$ has endpoints associated with point-based representations* $\mathrm{PR}(A; p)$, $\mathrm{PR}(B; q)$, *and the* weight $w(e) = \mathrm{PRM}\big(\,\mathrm{PR}(A; p),\ \mathrm{PR}(B; q)\,\big)$. *The* Nested Bottleneck Metric *is defined as* $\mathrm{NBM}(A, B) = \mathrm{BMD}(\Gamma(A, B))$.

Example 4.5 (4-point clouds $C^{\pm}$). *In $\mathbb{R}^2$, consider the 4-point clouds $C^{\pm} = \{p_1, p_2, p_3, p_4^{\pm}\}$, where $p_1 = (4a, 0)$, $p_2 = (b, c)$, $p_3 = -p_2 = (-b, -c)$, $p_4^+ = (0, 4d)$, and $p_4^- = (0, -4d)$ for parameters $a, b, c, d \geq 0$, see Fig. 2. Appendix C will explicitly compute* $\mathrm{NDP}(C^{\pm})$ *to distinguish all clouds $C^+ \not\cong C^-$. Fig. 4 shows the new metric* NBM *for variable parameters a, b and fixed c, d.* $\mathrm{NBM} > 0$ *implies that $C^+ \not\cong C^-$, except in the singular cases below. If $a = 0$ or $d = 0$ or $b = c = 0$, the clouds are related by a 2-fold rotation around the origin 0. If $a = \frac{\sqrt{3}}{2} \approx 0.87$, $b = 0$, $c = 2$, $d = 0.5$, then C^+ consists of the vertices $(0, \pm 2), (2\sqrt{3}, 0)$ of an equilateral triangle, where $(0, 2)$ is the double point $p_2 = p_4^+$. Then C^- is the same equilateral triangle but its vertex $(0, -2)$ is the double point $p_3 = p_4^-$. Because these clouds are related by rotation,* $\mathrm{NBM} = 0$ *in the black pixel at $a = \frac{\sqrt{3}}{2} \approx 0.87$, $b = 0$ in Fig. 4.*

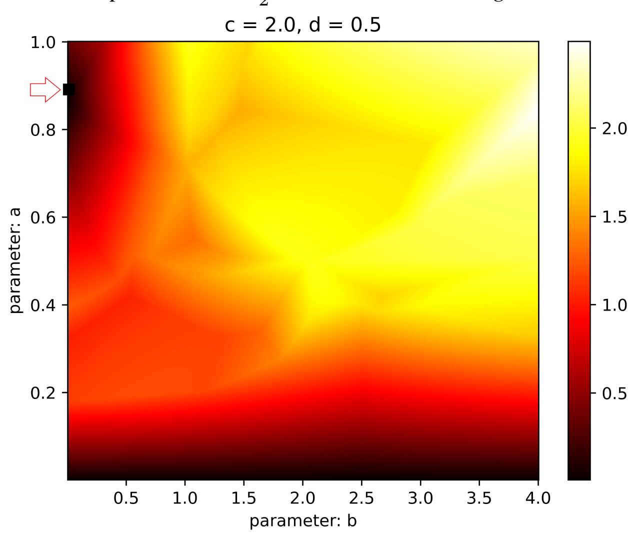

Figure 4. The Nested Bottleneck Metric NBM in Definition 4.4 for the clouds $C^{\pm} \subset \mathbb{R}^2$ that depend on parameters a, b and are not distinguished by 6 pairwise distances in Fig. 2, see Example C.1.

5. Bi-continuity and polynomial algorithms

For a *fixed dimension* n, all algorithms for m unordered points will have polynomial times in m in the RAM model.

Theorem 5.1 (Lipschitz continuity of NBM). *Let $B \subset \mathbb{R}^2$ be obtained from a cloud $A \subset \mathbb{R}^2$ by perturbing every point of A up to Euclidean distance ε. Then* $\mathrm{NBM}(A, B) \leq 6\varepsilon$.

To illustrate Theorem 5.1, we generated uniformly random clouds A in the unit square and cube. To get a perturbation B of A, we shifted every point of A by adding a uniformly random value in $[-\varepsilon, \varepsilon]$ to each coordinate, where $\varepsilon \in [0.01, 0.1]$ is a noise bound. Fig. 5 shows how the Nested Bottleneck Metric (NBM, averaged over several clouds) linearly increases with respect to the noise bound.

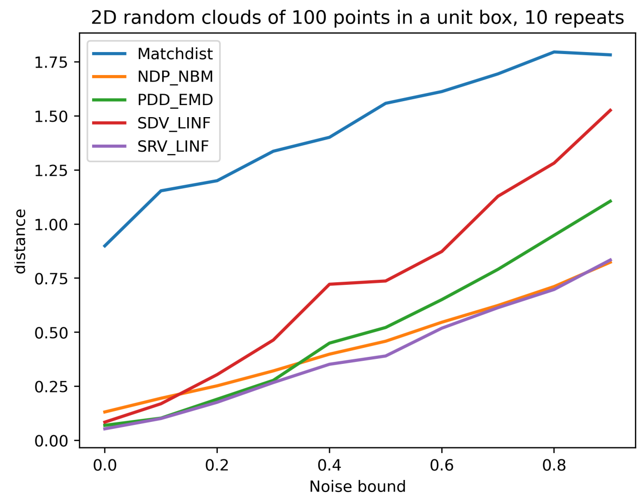

Figure 5. The metric $\mathrm{NBM}(\mathrm{NDP}(A), \mathrm{NDP}(B))$ for a random cloud A and its ε-perturbation B increases at most linearly in the noise bound ε with a Lipschitz constant $\lambda_2 < 6$ as in Theorem 5.1.

Theorem 5.2 (NDP time). *For any cloud $A \subset \mathbb{R}^2$ of m unordered points, the Nested Distributed Projection* $\mathrm{NDP}(A)$ *is computed in time $O(m^2)$ with space $O(m^2)$.*

Theorem 5.3 (NBM time). *For any clouds $A, B \subset \mathbb{R}^2$ of m unordered points, the Nested Bottleneck Metric* $\mathrm{NBM}(A, B)$ *is computable in time $O(m^{3.5} \log m)$ with space $O(m^3)$.*

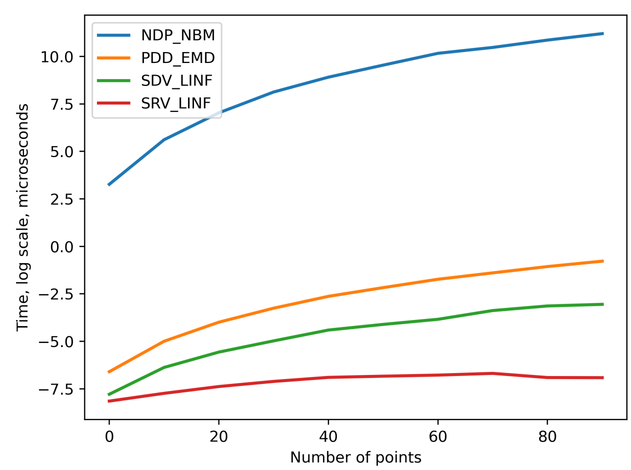

Figure 6. Times (microseconds, log scale) of metrics on invariants.

Fig. 6 illustrates a polynomial dependence of the NBM time in Theorem 5.3. Theorem 5.4 says that any m-point clouds $A, B \subset \mathbb{R}^2$ can be matched up to a perturbation proportional to the Nested Bottleneck Metric $d = \text{NBM}$. If d is small, all points of A, B can be matched up to a perturbation $3\sqrt{2}d$ by rigid motion. In section 6, the experimental maximum of this approximate factor is $2.2 < 3\sqrt{2}$.

Theorem 5.4 (point matching)**.** *For any m-point clouds $A, B \subset \mathbb{R}^2$, one can find in time $O(m^{3.5} \log m)$ a rigid motion f of $\mathbb{R}^2$ and a bijection $\beta : A \to B$ such that the match distance $\max\limits_{q \in A} |f(q) - \beta(q)| \leq 3\sqrt{2}\text{NBM}(A, B)$, see the comparison of this distance with others in Fig. 5.*

By Theorem 5.1, perturbing every atom up to ε (due to the ever-present thermal vibrations) changes NDP up to 6ε in the metric NBM. Conversely, by Theorem 5.4, if $\text{NBM}(A, B) = \delta > 0$ is small, the clouds A, B can be approximately matched by rigid motion up to $3\sqrt{2}\delta$ pointwise.

If clouds $A, B \subset \mathbb{R}^n$ have ordered points, one can *morph* (continuously transform) A to B by moving every i-th point of A along a straight-line to the i-th point of B for $i = 1, \dots, m$. If m points are unordered, there are $m!$ potential transformations, one for each permutation of m points.

Associating every point $p \in A$ to its nearest neighbor $q \in B$ is justified only for fixed clouds because a rigid motion of A can change a nearest neighbor of any point $p \in A$ in B.

6. Experiments on large molecular databases

The big databases of molecules with *3D conformers* (embeddings in $\mathbb{R}^3$) are QM9 (130K+ entries) (Ramakrishnan et al., 2014) and GD (GEOM_drugs, 31M+ entries) containing hundreds of 3D conformers of *unordered* atoms for each of 61607 chemical compositions (Axelrod & Gomez-Bombarelli, 2022). The Protein Data Bank has backbones of *ordered* atoms classified by simpler invariants (Anosova et al., 2025). All experiments took a few hours on Ryzen 9 3950X 3.5 GHz, 64 MB of L3 cache, RAM 82GB.

The ICML guide for application-driven ML says that "novel ideas that are simple to apply may be especially valuable", so we start with simpler and much faster invariants below.

Definition 6.1 (invariants $\text{SRV}, \text{SDV}, \text{PDD}$)**.** *Let $A \subset \mathbb{R}^n$ be a cloud of m unordered points with the center of mass at $0 \in \mathbb{R}^n$. The* Sorted Radial Vector $\text{SRV}(A)$ *has m radial distances $|p|$ in decreasing order for all $p \in A$. The* Sorted Distance Vector $\text{SDV}(A)$ *is the vector of $\frac{m(m-1)}{2}$ pairwise distances $|p - q|$ in decreasing order for distinct $p, q \in A$. For any point $p \in A$, let $d_1(p) \leq \dots \leq d_{m-1}(p)$ be Euclidean distances from p to all other points $q \in A \setminus \{p\}$ in increasing order. These distance lists become rows of the $m \times (m - 1)$ matrix $D(S; k)$. Any $l > 1$ identical rows are collapsed into a single row with the* weight l/m. *The final matrix with at most m unordered weighted rows and $m - 1$ ordered columns is the* Pointwise Distance Distribution.

For a PDD on m points, we sort m distance lists in time $O(m^2 \log m)$. Then PDDs are compared by the Earth Mover's Distance EMD (Rubner et al., 2000) in time $O(m^3)$. Table 2 emphasizes that most clouds should be first distinguished by simpler and faster invariants $\text{SRV}, \text{SDV}, \text{PDD}$. The complete NDP is needed only in rare cases but is still essential because any incomplete invariant I has no chance to predict different properties on *false positives* that are molecules $A \not\cong B$ with $I(A) = I(B)$.

Table 2. Invariants and metrics on cloud $A \subset \mathbb{R}^2$ with m unordered points: from the fastest (linear-time) to complete.

INVARIANT	TIME	METRIC	TIME
SRV	$O(m \log m)$	L_∞	$O(m)$
SDV	$O(m^2)$	L_∞	$O(m^2)$
PDD	$O(m^2 \log m)$	EMD	$O(m^3)$
NDP	$O(m^2)$	NBM	$O(m^{3.5} \log m)$

For a fixed atom $p \in A$ and $k < m$, the first k distances to neighbors in the row of p in $\text{PDD}(A)$ is an atomwise version of $\text{SRV}(A)$. This vector $D(A, p; k)$ of k distances was the only input for predicting the chemical element of p. A default network in TensorFlow was trained on clouds with the 80/20 split and achieved 98% accuracy for $k = 4$ in Table 4 despite the unbalanced counts of frequent elements in Table 3. Appendix A has all implementation details.

Table 3. Counts of atoms by chemical elements in QM9 (2,407,753 atoms), GD0 (GEOM_drugs 0th conformers, 12,917,980 atoms).

QM9: H 1,230,122	QM9: C 846,557	QM9: N 139,764	QM9: O 187,996	QM9: F 3,314
GD0: H 5,660,986	GD0: C 5,267,096	GD0: N 842,562	GD0: O 854,400	GD0: F 64,299
GD0: P 1,350	GD0: S 159,648	GD0: Cl 53,404	GD0: Br 14,010	GD0: I 225

Table 4. Accuracies in percentages for predicting chemical elements by a 4-layer network using *only Euclidean distances* from an atomic center to its k nearest neighbors for QM9 and GD0.

data	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
QM9	94.63	98.64	98.24	98.54	98.77
GD0	91.44	96.67	98.05	98.70	98.49

All past attempts by both ML and non-ML in chemistry achieved only 86% on similar size data, see Table 7 summarized in (Vasylenko et al., 2025), because the underlying descriptors were not invariant, e.g. under permutations of atoms, which creates exponentially many representations of the same molecule, incomplete, or their similarities failed the triangle axiom, e.g. see (Steck et al., 2024).

High accuracies of $D(A, p; 4)$ in Table 4 are explained by the following cascade computations. First, split all clouds from Table 3 by the 1st distance (to the nearest neighbor of a central atom p) rounded to 3 decimal places in Å. This is a typical experimental precision, where $1\text{Å} = 10^{-10}m$ is the smallest interatomic distance. Second, split each subset with equal 1st distances by 2nd distances, and so on up to $k = 5$ distances. All clouds of different elements in QM9 and GD0 were separated by $D(A, p; 4)$ and $D(A, p; 5)$, respectively.

We compared full molecules starting with the pseudo-metric L_∞ (max abs difference of corresponding coordinates) on SRVs of all 873,527,974 pairs of 3D atomic clouds having equal numbers of atoms in QM9, then 8,735,279 distances L_∞ on SDVs of the 1% closest pairs, 87,352 EMDs on PDDs of the 1% closest pairs, and NBMs on NDPs for the final 10K closest pairs. In this hierarchical computation, large values of L_∞ (then EMD) guarantee that molecules are distant and cannot be closely matched by rigid motion. Tiny or zero values of pseudo-metrics guarantee nothing because SDV and PDD can coincide for very different clouds, see Fig. 2, Fig. S4 in (Pozdnyakov et al., 2020).

Table 5. Chemically different molecules (given by QM9 ids) are geometrically distinguished by $\mathrm{SRV}, \mathrm{SDV}, \mathrm{PDD}, \mathrm{NDP}$, see Fig. 8.

smallest distances in Å, molecule A $\neq$ molecule B
SRV, $L_\infty = 0.021$, $H_4C_5N_2O(5365) \neq H_3C_4N_3O_2(131923)$
SDV, $L_\infty = 0.055$, $H_3C_4N_5(123533) \neq H_3C_5N_3O(24547)$
$\mathrm{EMD} = 0.051$, $H_3C_4N_5(123533) \neq H_3C_5N_3O(24521)$
$\mathrm{NBM} = 0.148$, $H_3C_4N_3O_2(28141) \neq H_3C_3N_5O(130099)$

Fig. 7 compares the new metric $y = \mathrm{NBM}$ on complete NDPs with the pseudo-metric $x = \mathrm{PDD}$. All pairs A, B with (x, y) close to the vertical axis in Fig. 7 (left) have $\mathrm{EMD} \approx 0$ because they are almost mirror images (indistinguishable by PDD) well distinguished by higher values of NBM. Fig. 8 shows bonds by standard visualization, they were not used for clouds of points without any edges.

For each of 31M+ entries (3D conformers) in the much larger database GD, we took the cloud A of all atoms without chemical elements and computed $\mathrm{SRV}(A; k)$ of up to $k = 10$ largest distances (rounded to 3 decimal places) from the center of mass of A to all atoms. Similar to QM9, cascade comparisons confirmed that $\mathrm{SRV}(A; 7)$ distinguishes all chemically different molecules, while only four pairs have equal $\mathrm{SRV}(A; 6)$ rounded to 3 decimal places. This transparent reconstruction of a full chemical composition from precise enough geometry gives hope to explain other molecular properties in terms of geometric invariants.

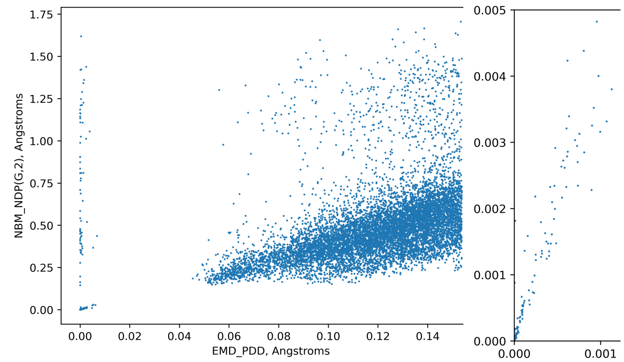

Figure 7. $x = \mathrm{EMD}(\mathrm{PDD}(A), \mathrm{PDD}(B))$ vs $y = \mathrm{NBM}(A, B)$ on complete invariants NDP with zoomed-in comparisons on the right, which all appear only for chemically identical molecules.

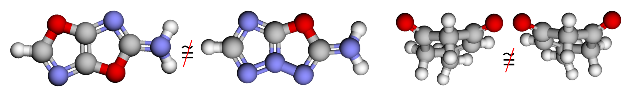

Figure 8. **Left**: chemically different QM9 molecules 28141 and 130099 have the smallest distances $\mathrm{NBM} \approx 0.15$Å. **Right**: molecules 70954 and 74130 are almost mirror images with $\mathrm{EMD} \approx 0.0004$Å but are well distinguished by $\mathrm{NBM} \approx 1.619$Å.

7. Discussion: conclusions and limitations

For clouds with different numbers of points, we can replace the bottleneck distance BD in Definition 4.2 with any metric between fixed clouds of different sizes, e.g. the Hausdorff distance, to get a metric on PRs. Then we can compare NDPs of any clouds as weighted distributions by EMD. The limitation is the proof of Theorem 5.4 in dimension $n = 2$, though the experiments indicate the Lipschitz continuity of NDP^{-1} in $\mathbb{R}^3$. All other conditions in Problem 1.1 are proved in the appendices for any dimension $n \geq 2$.

The experiments imply that mapping any molecule to (the rigid class of) its cloud of atomic centers is *injective* without losing any chemical information, so all chemical elements can be reconstructed from pure geometry. This result confirms our physical intuition that replacing atoms should perturb geometry at least slightly, which was impossible to establish without complete and Lipschitz continuous invariants. Hence all molecules of m atoms live at different locations in the common *Cloud Rigid Space* $\mathrm{CRS}(\mathbb{R}^3; m)$ of SE(3)-classes of all clouds of m unordered points.

Most significantly, a *molecular structure* can now be defined not as a huge collection of vectors under rotations and atom permutations, see Fig. 1 in (Lang et al., 2024), but as a rigid (class of a) cloud of atomic centers (without chemical elements), which is uniquely determined by an efficient hierarchy of invariants from the fastest (linear-time) SRV to the new complete invariant NDP solving Problem 1.1.

Impact Statement

This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of our work, none which we feel must be specifically highlighted here.

References

Akhtar, N. and Mian, A. Threat of adversarial attacks on deep learning in computer vision: A survey. *IEEE Access*, 6:14410–14430, 2018.

Anosova, O., Gorelov, A., Jeffcott, W., Jiang, Z., and Kurlin, V. A complete and bi-continuous invariant of protein backbones under rigid motion. *MATCH Communications in Mathematical and in Computer Chemistry (to appear), arxiv:2410.08203*, 2025.

Antunes, L. M., Grau-Crespo, R., and Butler, K. T. Distributed representations of atoms and materials for machine learning. *npj Computational Materials*, 8(1):44, 2022.

Aronsson, J. Homogeneous vector bundles and g-equivariant convolutional neural networks. *Sampling Theory, Signal Processing, and Data Analysis*, 20(2):10, 2022.

Assaad, S., Downey, C., Al-Rfou, R., Nayakanti, N., and Sapp, B. Vn-transformer: Rotation-equivariant attention for vector neurons. *Transactions on Machine Learning Research*, 2023.

Atz, K., Grisoni, F., and Schneider, G. Geometric deep learning on molecular representations. *Nature Machine Intelligence*, 3(12):1023–1032, 2021.

Axelrod, S. and Gomez-Bombarelli, R. Geom, energy-annotated molecular conformations for property prediction and molecular generation. *Scientific Data*, 9(1):185, 2022.

Boutin, M. and Kemper, G. On reconstructing n-point configurations from the distribution of distances or areas. *Adv. Appl. Math.*, 32(4):709–735, 2004.

Brass, P. and Knauer, C. Testing the congruence of d-dimensional point sets. In *SoCG*, pp. 310–314, 2000.

Brass, P. and Knauer, C. Testing congruence and symmetry for general 3-dimensional objects. *Computational Geometry*, 27(1):3–11, 2004.

Brécheteau, C. A statistical test of isomorphism between metric-measure spaces using the distance-to-a-measure signature. pp. 795–849, 2019.

Brink, C., Kahl, W., and Schmidt, G. *Relational methods in computer science*. Springer Science & Business Media, 1997.

Bronstein, M. M., Bruna, J., Cohen, T., and Veličković, P. Geometric deep learning: grids, groups, graphs, geodesics, and gauges. *arXiv:2104.13478*, 2021.

Cahill, J., Iverson, J. W., and Mixon, D. G. Towards a bilipschitz invariant theory. *Applied and Computational Harmonic Analysis*, 72:101669, 2024.

Chauvin, L., Wells III, W., and Toews, M. Registering image volumes using 3D SIFT and discrete SP-symmetry. *arXiv:2205.15456*, 2022.

Chen, C., Ye, W., Zuo, Y., Zheng, C., and Ong, S. P. Graph networks as a universal machine learning framework for molecules and crystals. *Chemistry of Materials*, 31(9): 3564–3572, 2019.

Chen, H., Liu, S., Chen, W., Li, H., and Hill, R. Equivariant point network for 3D point cloud analysis. In *Computer Vision and Pattern Recognition*, pp. 14514–14523, 2021.

Chew, P. and Kedem, K. Improvements on geometric pattern matching problems. In *Scandinavian Workshop on Algorithm Theory*, pp. 318–325, 1992.

Chew, P., Goodrich, M., Huttenlocher, D., Kedem, K., Kleinberg, J., and Kravets, D. Geometric pattern matching under Euclidean motion. *Computational Geometry*, 7 (1-2):113–124, 1997.

Chew, P., Dor, D., Efrat, A., and Kedem, K. Geometric pattern matching in d-dimensional space. *Discrete & Computational Geometry*, 21(2):257–274, 1999.

Cohen, T. S., Geiger, M., and Weiler, M. A general theory of equivariant cnns on homogeneous spaces. *Advances in Neural Information Processing Systems*, 32, 2019.

Colbrook, M. J., Antun, V., and Hansen, A. C. The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem. *Proc. National Academy of Sciences*, 119(12):e2107151119, 2022.

Cosmo, L., Panine, M., Rampini, A., Ovsjanikov, M., Bronstein, M. M., and Rodola, E. Isospectralization, or how to hear shape, style, and correspondence. In *Proceedings of CVPR*, pp. 7529–7538, 2019.

Dekster, B. V. and Wilker, J. B. Edge lengths guaranteed to form a simplex. *Archiv der Mathematik*, 49(4):351–366, 1987.

Delle Rose, V., Kozachinskiy, A., Rojas, C., Petrache, M., and Barceló, P. Three iterations of (d- 1)-wl test distinguish non isometric clouds of d-dimensional points. *Advances in Neural Information Processing Systems*, 36, 2024.

Deng, C., Litany, O., Duan, Y., Poulenard, A., Tagliasacchi, A., and Guibas, L. J. Vector neurons: A general framework for so(3)-equivariant networks. In *Proceedings of the International Conference on Computer Vision*, pp. 12200–12209, 2021.

Deza, E. and Deza, M. M. *Encyclopedia of distances*. Springer, 2009.

Dong, Y., Liao, F., Pang, T., Su, H., Zhu, J., Hu, X., and Li, J. Boosting adversarial attacks with momentum. In *Computer vision and pattern recognition*, pp. 9185–9193, 2018.

Duxbury, P. M., Granlund, L., Gujarathi, S., Juhas, P., and Billinge, S. J. The unassigned distance geometry problem. *Discrete Applied Mathematics*, 204:117–132, 2016.

Dym, N. and Gortler, S. J. Low-dimensional invariant embeddings for universal geometric learning. *Foundations of Computational Mathematics*, pp. 1–41, 2024.

Efrat, A., Itai, A., and Katz, M. J. Geometry helps in bottleneck matching and related problems. *Algorithmica*, 31(1):1–28, 2001.

Feynman, R. *The Feynman lectures on physics. Chapter 1: atoms in motion*, volume 1. 1971.

Fuchs, F., Worrall, D., Fischer, V., and Welling, M. Se(3)-transformers: 3d roto-translation equivariant attention networks. *Advances in neural information processing systems*, 33:1970–1981, 2020.

Goodrich, M. T., Mitchell, J. S., and Orletsky, M. W. Approximate geometric pattern matching under rigid motions. *Transactions on Pattern Analysis and Machine Intelligence*, 21(4):371–379, 1999.

Gordon, C., Webb, D., and Wolpert, S. Isospectral plane domains and surfaces via riemannian orbifolds. *Inventiones mathematicae*, 110(1):1–22, 1992a.

Gordon, C., Webb, D. L., and Wolpert, S. One cannot hear the shape of a drum. *Bulletin of the American Mathematical Society*, 27(1):134–138, 1992b.

Goyal, A., Law, H., Liu, B., Newell, A., and Deng, J. Revisiting point cloud shape classification with a simple and effective baseline. In *International Conference on Machine Learning*, pp. 3809–3820, 2021.

Guo, C., Gardner, J., You, Y., Wilson, A. G., and Weinberger, K. Simple black-box adversarial attacks. In *International Conference on Machine Learning*, pp. 2484–2493, 2019.

Hausdorff, F. Dimension und äuβeres maβ. *Mathematische Annalen*, 79(2):157–179, 1919.

Hopcroft, J. E. and Karp, R. M. An n^5/2 algorithm for maximum matchings in bipartite graphs. *SIAM Journal on Computing*, 2(4):225–231, 1973.

Hordan, S., Amir, T., Gortler, S. J., and Dym, N. Complete neural networks for euclidean graphs. In *AAAI Conference on Artificial Intelligence*, volume 38 (11), pp. 12482–12490, 2024.

Horn, R. A. and Johnson, C. R. *Matrix analysis*. Cambridge University Press, 2012.

Huang, Q.-X., Flöry, S., Gelfand, N., Hofer, M., and Pottmann, H. Reassembling fractured objects by geometric matching. In *ACM SIGGRAPH*, pp. 569–578. 2006.

Huttenlocher, D. P., Klanderman, G. A., and Rucklidge, W. J. Comparing images using the Hausdorff distance. *Transactions on pattern analysis and machine intelligence*, 15 (9):850–863, 1993.

Kac, M. Can one hear the shape of a drum? *The american mathematical monthly*, 73(4P2):1–23, 1966.

Kapovich, M. and Millson, J. J. The symplectic geometry of polygons in euclidean space. *Journal of Differential Geometry*, 44(3):479–513, 1996.

Kendall, D. G., Barden, D., Carne, T. K., and Le, H. *Shape and shape theory*. John Wiley & Sons, 2009.

Keriven, N. and Peyré, G. Universal invariant and equivariant graph neural networks. *Advances in Neural Information Processing Systems*, 32, 2019.

Kondor, R. and Trivedi, S. On the generalization of equivariance and convolution in neural networks to the action of compact groups. In *International Conference on Machine Learning*, pp. 2747–2755, 2018.

Kruskal, J. B. and Wish, M. *Multidimensional scaling*. Number 11. Sage, 1978.

Kurlin, V. Polynomial-time algorithms for continuous metrics on atomic clouds of unordered points. *MATCH Communications in Mathematical and in Computer Chemistry*, 91:79–108, 2024.

Laidlaw, C. and Feizi, S. Functional adversarial attacks. *Adv. Neural Information Proc. Systems*, 32, 2019.

Lang, L., Cezar, H. M., Adamowicz, L., and Pedersen, T. B. Quantum definition of molecular structure. *Journal of the American Chemical Society*, 146(3):1760–1764, 2024.

Leman, A. and Weisfeiler, B. A reduction of a graph to a canonical form and an algebra arising during this reduction. *Nauchno-Technicheskaya Informatsiya*, 2(9):12–16, 1968.

Li, X., Li, R., Chen, G., Fu, C.-W., Cohen-Or, D., and Heng, P.-A. A rotation-invariant framework for deep point cloud analysis. *IEEE transactions on visualization and computer graphics*, 28(12):4503–4514, 2021.

Liberti, L. and Lavor, C. *Euclidean distance geometry*. Springer, 2017.

Lin, Z. C., Lee, H., and Huang, T. S. Finding 3d point correspondences in transformation estimation. In *Proceedings-International Conference on Pattern Recognition*, pp. 303–305. IEEE, 1986.

Lowe, D. G. Object recognition from local scale-invariant features. In *Proceedings of ICCV*, volume 2, pp. 1150–1157, 1999.

Lowe, D. G. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60: 91–110, 2004.

Maennel, H., Unke, O. T., and Muller, K.-R. Complete and efficient covariants for three-dimensional point configurations with application to learning molecular quantum properties. *The Journal of Physical Chemistry Letters*, 15:12513–12519, 2024.

Majhi, S., Vitter, J., and Wenk, C. Approximating gromov-hausdorff distance in euclidean space. *Computational Geometry*, 116:102034, 2024.

Marin, R., Rampini, A., Castellani, U., Rodolà, E., Ovsjanikov, M., and Melzi, S. Spectral shape recovery and analysis via data-driven connections. *International journal of computer vision*, 129:2745–2760, 2021.

Maron, H., Fetaya, E., Segol, N., and Lipman, Y. On the universality of invariant networks. In *International conference on machine learning*, pp. 4363–4371, 2019.

Mémoli, F. Gromov–Wasserstein distances and the metric approach to object matching. *Foundations of computational mathematics*, 11:417–487, 2011.

Merchant, A., Batzner, S., Schoenholz, S. S., Aykol, M., Cheon, G., and Cubuk, E. D. Scaling deep learning for materials discovery. *Nature*, 624(7990):80–85, 2023.

Morissette, S. L., Soukasene, S., Levinson, D., Cima, M. J., and Almarsson, Ö. Elucidation of crystal form diversity of the hiv protease inhibitor ritonavir by high-throughput crystallization. *Proceedings of the National Academy of Sciences*, 100(5):2180–2184, 2003.

Morris, C., Dym, N., Maron, H., Ceylan, İ. İ., Frasca, F., Levie, R., Lim, D., Bronstein, M., Grohe, M., and Jegelka, S. Future directions in foundations of graph machine learning. *arXiv:2402.02287*, 2024.

Nemec, L. Principal component analysis (pca): A physically intuitive mathematical introduction. https://towardsdatascience.com/principal-component-analysis-pca-8133b02f11bd, 2022.

Nigam, J., Pozdnyakov, S. N., Huguenin-Dumittan, K. K., and Ceriotti, M. Completeness of atomic structure representations. *APL Machine Learning*, 2(1), 2024.

Oliynyk, A. O., Antono, E., Sparks, T. D., Ghadbeigi, L., Gaultois, M. W., Meredig, B., and Mar, A. High-throughput machine-learning-driven synthesis of full-heusler compounds. *Chemistry of Materials*, 28(20): 7324–7331, 2016.

Pozdnyakov, S. N. and Ceriotti, M. Incompleteness of graph convolutional neural networks for points clouds in three dimensions. *arXiv:2201.07136*, 2022.

Pozdnyakov, S. N., Willatt, M. J., Bartók, A. P., Ortner, C., Csányi, G., and Ceriotti, M. Incompleteness of atomic structure representations. *Phys. Rev. Lett.*, 125:166001, 2020. URL arXiv:2001.11696.

Ramakrishnan, R., Dral, P. O., Rupp, M., and Von Lilienfeld, O. A. Quantum chemistry structures and properties of 134 kilo molecules. *Scientific data*, 1(1):1–7, 2014.

Rass, S., König, S., Ahmad, S., and Goman, M. Metricizing the euclidean space towards desired distance relations in point clouds. *IEEE Transactions on Information Forensics and Security*, 19:7304–7319, 2024.

Reuter, M., Wolter, F.-E., and Peinecke, N. Laplace–beltrami spectra as 'shape-dna'of surfaces and solids. *Computer-Aided Design*, 38(4):342–366, 2006.

Rubner, Y., Tomasi, C., and Guibas, L. The Earth Mover's Distance as a metric for image retrieval. *International Journal of Computer Vision*, 40(2):99–121, 2000.

Rudin, W. et al. *Principles of mathematical analysis*, volume 3. McGraw-hill New York, 1976.

Satorras, V. G., Hoogeboom, E., and Welling, M. E(n) equivariant graph neural networks. In *International conference on machine learning*, pp. 9323–9332, 2021.

Schmiedl, F. Computational aspects of the Gromov–Hausdorff distance and its application in non-rigid shape matching. *Discrete Comp. Geometry*, 57:854–880, 2017.

Schoenberg, I. Remarks to Maurice Frechet's article "Sur la definition axiomatique d'une classe d'espace distances vectoriellement applicable sur l'espace de Hilbert. *Annals of Mathematics*, pp. 724–732, 1935.

Shi, J., Yang, H., and Carlone, L. Robin: a graph-theoretic approach to reject outliers in robust estimation using invariants. In *International Conference on Robotics and Automation (ICRA)*, pp. 13820–13827, 2021.

Spezialetti, R., Salti, S., and Stefano, L. D. Learning an effective equivariant 3d descriptor without supervision. In *ICCV*, pp. 6401–6410, 2019.

Steck, H., Ekanadham, C., and Kallus, N. Is cosine-similarity of embeddings really about similarity? In *Companion Proceedings of the ACM on Web Conference 2024*, pp. 887–890, 2024.

Su, Z., Welling, M., Pietikäinen, M., and Liu, L. Svnet: Where SO(3) equivariance meets binarization on point cloud representation. In *International Conference on 3D Vision*, pp. 547–556, 2022.

Thomas, N., Smidt, T., Kearnes, S., Yang, L., Li, L., Kohlhoff, K., and Riley, P. Tensor field networks: Rotation-and translation-equivariant neural networks for 3d point clouds. *arXiv:1802.08219*, 2018.

Tshitoyan, V., Dagdelen, J., Weston, L., Dunn, A., Rong, Z., Kononova, O., Persson, K. A., Ceder, G., and Jain, A. Unsupervised word embeddings capture latent knowledge from materials science literature. *Nature*, 571(7763):95–98, 2019.

Turner, K., Mukherjee, S., and Boyer, D. M. Persistent homology transform for modeling shapes and surfaces. *Information and Inference: A Journal of the IMA*, 3(4): 310–344, 2014.

Vasylenko, A., Antypov, D., Schewe, S., Daniels, L. M., Claridge, J. B., Dyer, M. S., and Rosseinsky, M. J. Digital features of chemical elements extracted from local geometries in crystal structures. *Digital Discovery*, 2025.

Villar, S., Hogg, D. W., Storey-Fisher, K., Yao, W., and Blum-Smith, B. Scalars are universal: equivariant machine learning, structured like classical physics. *Advances in Neural Information Processing Systems*, 34:28848–28863, 2021.

Wang, Y. and Solomon, J. M. Deep closest point: Learning representations for point cloud registration. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 3523–3532, 2019.

Ward, L., Agrawal, A., Choudhary, A., and Wolverton, C. A general-purpose machine learning framework for predicting properties of inorganic materials. *npj Computational Materials*, 2(1):1–7, 2016.

Weston, L., Tshitoyan, V., Dagdelen, J., Kononova, O., Trewartha, A., Persson, K. A., Ceder, G., and Jain, A. Named entity recognition and normalization applied to large-scale information extraction from the materials science literature. *Journal of chemical information and modeling*, 59(9):3692–3702, 2019.

Weyl, H. *The classical groups: their invariants and representations*. Number 1. Princeton university press, 1946.

Widdowson, D. and Kurlin, V. Resolving the data ambiguity for periodic crystals. *Advances in Neural Information Processing Systems*, 35:24625–24638, 2022.

Widdowson, D. E. and Kurlin, V. A. Recognizing rigid patterns of unlabeled point clouds by complete and continuous isometry invariants with no false negatives and no false positives. In *Computer Vision and Pattern Recognition*, pp. 1275–1284, 2023.

Xu, Y., Lei, J., Dobriban, E., and Daniilidis, K. Unified fourier-based kernel and nonlinearity design for equivariant networks on homogeneous spaces. In *International Conference on Machine Learning*, pp. 24596–24614, 2022.

Yang, H., Shi, J., and Carlone, L. Teaser: Fast and certifiable point cloud registration. *IEEE Transactions on Robotics*, 37(2):314–333, 2020.

Yarotsky, D. Universal approximations of invariant maps by neural networks. *Constructive Approximation*, 55(1): 407–474, 2022.

Zhou, Q., Tang, P., Liu, S., Pan, J., Yan, Q., and Zhang, S.-C. Learning atoms for materials discovery. *Proceedings of the National Academy of Sciences*, 115(28):E6411–E6417, 2018.

Zhu, W., Chen, L., Hou, B., Li, W., Chen, T., and Liang, S. Point cloud registration of arrester based on scale-invariant points feature histogram. *Scientific Reports*, 12 (1):1–13, 2022.

Introduction to appendices

The main contribution is the roadmap for any data challenge through well-motivated Problem 1.1, where clouds and rigid motion can be replaced with any objects and equivalences. The conditions of completeness and Lipschitz continuity of an invariant I cover the *discriminative* challenge. After these conditions 1.1(a,b,e) are satisfied, the invariant I can be inverted in principle and opens the *generative challenge* of its realizability and inverse continuity in 1.1(c,d,e).

Problem 1.1 was stated for unordered clouds under rigid motion but was also solved for *isometry* and compositions of these equivalences with uniform scaling in $\mathbb{R}^n$. For $m = 4$ points, plane quadrilaterals were previously classified in discrete classes in Fig. 1 (right), while appendix C shows the first continuous maps of the invariant space $\mathrm{CRS}(\mathbb{R}^2; 4)$. Conditions 1.1(d,e,f) enable a generation of real clouds in $\mathrm{CRS}(\mathbb{R}^n; m)$ from their invariants. A full answer to the question 'same or different, and how much different' required complete invariants with Lipschitz continuous metrics.

The key contribution is a theoretically justified solution to Problem 1.1. The experiments on the databases QM9 and GEOM_drugs are considered complementary. Example C.1 and its extension in Example C.2 prove that infinitely many pairs of non-isometric clouds $C^+ \not\cong C^-$ (depending on 4 free parameters and having the same 6 pairwise distances) are distinguished by the new invariants. This result is impossible to justify by any finite experiment. Example C.1 demonstrated the non-zero distances between the complete invariants of $C^{\pm}$ in Fig. **??**.

The completeness and bi-Lipschitz continuity of the proposed invariants enabled the new experiments on 130K+ real molecules in section 6, which were not previously possible because all past invariants did not satisfy all conditions of Problem 1.1, especially the realizability condition that provides geographic-style maps on cloud spaces.

The full solution to Problem 1.1 for $n = 2$ is justified by Theorem 3.5 and Lemmas 3.3, 5.1, 5.2, 5.3. Theorem 3.3 enables a visualization of cloud spaces, which were unknown even for $m = 4$ unordered points in $\mathbb{R}^2$.

- The *Cloud Isometry Space* $\mathrm{CIS}(\mathbb{R}^n; m)$ of clouds of m unordered points under isometry in $\mathbb{R}^n$.

- The *Cloud Rigid Space* $\mathrm{CRS}(\mathbb{R}^n; m)$ of clouds of m unordered points under rigid motion in $\mathbb{R}^n$.

- The *Cloud Similarity Space* $\mathrm{CSS}(\mathbb{R}^n; m)$ of clouds of m unordered points under *geometric similarity*, which is a composition of isometry and uniform scaling in $\mathbb{R}^n$.

- The *Cloud Dilation Space* $\mathrm{DCS}(\mathbb{R}^n; m)$ of clouds of m unordered points under orientation-preserving geometric similarity (rigid motion and uniform scaling) in $\mathbb{R}^n$.

Here is a summary of the supplementary materials.

- Appendix A extends section 6 with more details of new invariants and metrics computed on the QM9 database and compared with past pseudo-metrics.

- Appendix C discusses parametrization of $\mathrm{CSS}(\mathbb{R}^2; m)$ and includes Examples C.1 and C.2 computing the new invariants NDP in detail for infinitely many 4-point clouds from Example C.1.

- Appendices B, D, E prove all theoretical results from sections 3, 4, 5, respectively.

- The zip folder with supplementary materials includes the code for computing all invariants and metrics as well as tables with all coordinates of colorful maps of QM9 and distances.

A. Extra details of experiments in section 6

The default 4-layer network from TensorFlow used a "sequential" mode, 3 epochs, and the settings in Table 6.

The only difference between QM9 and GD settings was in the number N of chemical elements in tf.keras.layers.Dense(N), where $N = 5$ for QM9 and $N = 10$ for GD.

The maps of QM9 in Fig. 9 are based on eigenvalues and too dense without clear separation. Even if we zoom in, these incomplete invariants will not separate molecules because 3D clouds have at most 3 eigenvalues. The complete invariants

Table 6. Parameters of the default 4-layer network for predictions in Table 4.

LAYER (TYPE)	OUTPUT SHAPE	NUMBER OF PARAMETERS
DENSE (DENSE)	(NONE, 32)	352
BATCH_NORMALIZATION	(NONE, 32)	128
RE_LU (RELU)	(NONE, 32)	0
DENSE_1 (DENSE)	(NONE, 5)	165

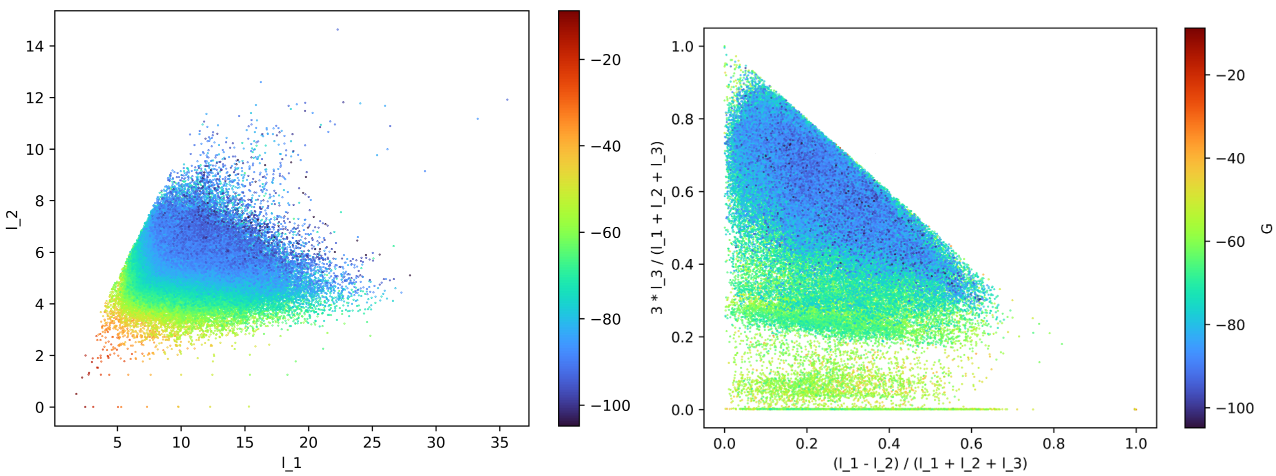

Figure 9. **Left**: each dot represents one QM9 molecule whose atomic cloud has two largest roots $l_1 \geq l_2$ of eigenvalues (moments of inertia (Nemec, 2022) or elongations in principal directions) in Angstroms ($1\mathring{A} = 10^{-10}m \approx$ smallest interatomic distance). The color represents the free energy G characterizing molecular stability. **Right**: each dot represents one QM9 molecule whose atomic cloud has coordinates x, y expressed via the roots $l_1 \geq l_2 \geq l_3 \geq 0$ of three eigenvalues.

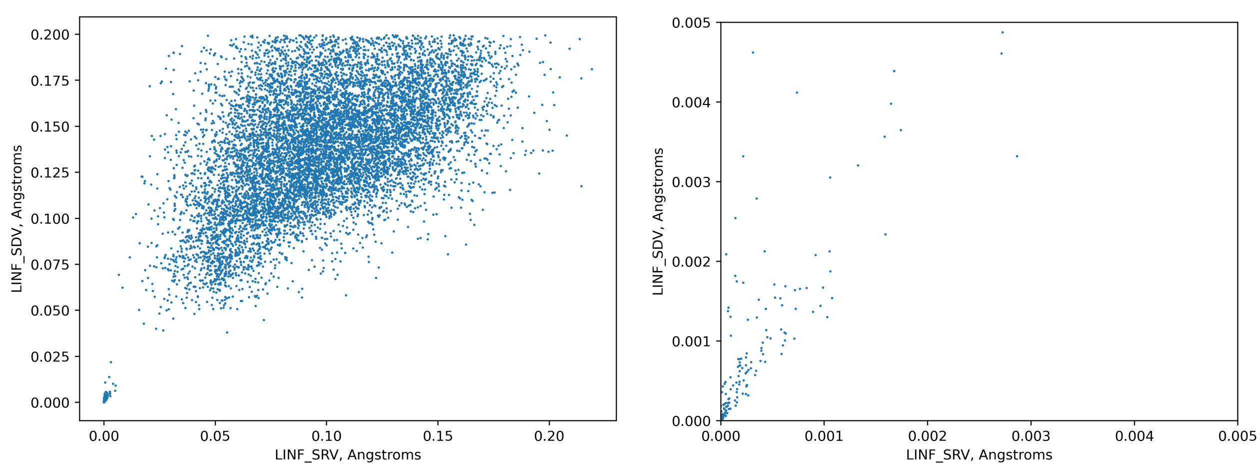

Figure 10. **Left**: each dot is a comparison of closest atomic clouds A, B from QM9 by the distances L_∞ on SRV vs L_∞ on SDV. **Right**: zoomed-in comparisons for very small distances.

NDP contain much more geometric information. Fig. 10 and 11 show that distances on stornger invariants have larger values and hence better separate molecules, though all these distances have the same Lipschitz constant 2.

Fig. 12 (left) shows the simplest projections of the atomic clouds from QM9, see the familiar molecules such as H_2O (water). Any small region on such a map can be zoomed in and displayed in other invariants from Table 2, see Fig. 12 (right).

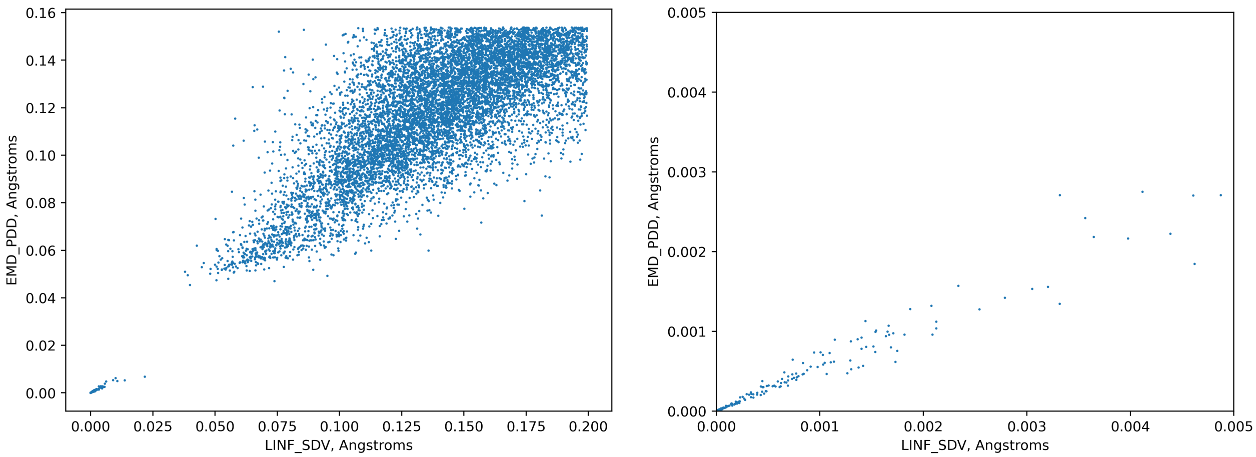

Figure 11. **Left**: each dot is a comparison of closest atomic clouds A, B from QM9 by the distances L_∞ on SDV vs EMD on PDD. **Right**: zoomed-in comparisons for very small distances.

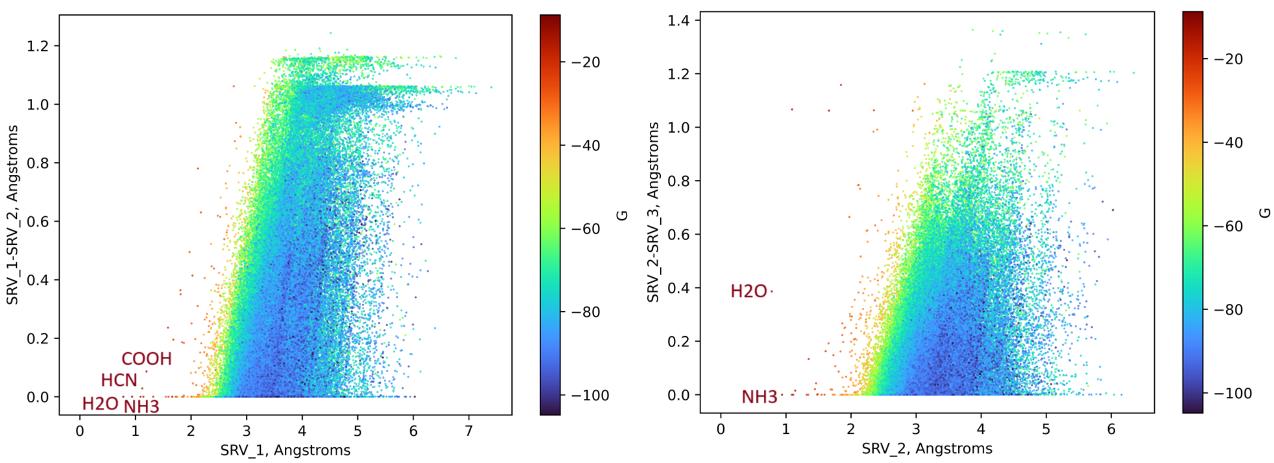

Figure 12. QM9 maps: each dot colored by the free energy G represents an atomic cloud. **Left**: $x = \mathrm{SRV}_1$, $y = \mathrm{SRV}_1 - \mathrm{SRV}_2$. **Right**: all molecules with $\mathrm{SRV}_1 = \mathrm{SRV}_2$ (two equidistant atoms from the center of mass) are projected to $x = \mathrm{SRV}_2$, $y = \mathrm{SRV}_2 - \mathrm{SRV}_3$.

Table 7. Past ML and non-ML predictions of chemical elements have lower accuracies than by distance invariants in Table 4.

Method	Description	Accuracy	Reference
LEAF	Local coordination geometry	86%	(Vasylenko et al., 2025)
MatScholar	ML-derived from literature	81%	(Weston et al., 2019)
Mat2Vec	ML-derived from literature	80%	(Tshitoyan et al., 2019)
Atom2Vec	ML-derived from compositional content	79%	(Zhou et al., 2018)
GNoME	Frequency of elements at the same atomic sites	79%	(Merchant et al., 2023)
Magpie	Elemental physical characteristics	78%	(Ward et al., 2016)
Oliynyk	Elemental physical characteristics	75%	(Oliynyk et al., 2016)
MEGNet	ML-derived from atom, bond and graph attributes	73%	(Chen et al., 2019)
SkipAtom	ML-derived from atom connectivity graphs	68%	(Antunes et al., 2022)

B. Generalization of section 3 and all proofs in dimensions $n \geq 2$

This appendix extends all concepts from section 3 to dimensions $n \geq 2$, extends Theorem 3.3 to Theorem B.7, which is proved with Theorem B.9 for any $n \geq 2$.

Lemma B.1 (vector $p_n^\perp$ orthogonal to $p_1,\dots,p_{n-1}$ in $\mathbb{R}^n$)**.** *Let $e_1,\dots,e_n$ be an orthonormal basis of $\mathbb{R}^n$, so $|e_i|=1$ and $e_i\cdot e_j=0$ for $i\neq j$. For any $n-1$ vectors $p_1,\dots,p_{n-1}\in\mathbb{R}^n$, there is a vector $p_n^\perp$ that is orthogonal to all $p_1,\dots,p_{n-1}$ and has coordinates that are degree $n-1$ polynomials in the coordinates of $p_1,\dots,p_{n-1}$.*

Proof of Lemma B.1. Below the 'unusual determinant' with the $n-1$ vector columns $p_1,\dots,p_{n-1}$ and the last column of the n vectors $e_1,\dots,e_n$ is only a short notation for the following expansion by the last column: $\begin{vmatrix} | & \dots & | & e_1 \\ p_1 & \dots & p_{n-1} & \vdots \\ | & \dots & | & e_n \end{vmatrix} =$ $\sum\limits_{i=1}^{n}(-1)^{n+i}\det(i)e_i$, where $\det(i)$ is the usual $(n-1)\times(n-1)$ determinant obtained from the $n-1$ vector columns $p_1,\dots,p_{n-1}$ by removing the i-th row, so we set $p_n^\perp=\sum\limits_{i=1}^{n}(-1)^{n+i}\det(i)e_i$.

For example, if $n=2$ then $p_1=(x_1,x_2)$ has the vector $p_2^\perp=\begin{vmatrix} x_1 & e_1 \\ x_2 & e_2\end{vmatrix}=x_1e_2-x_2e_1=(-x_2,x_1)\perp p_1$ If $n=3$, $p_1=(x_1,x_2,x_3)$ and $p_2=(y_1,y_2,y_3)$, then $p_3^\perp=\begin{vmatrix} x_1 & y_1 & e_1 \\ x_2 & y_2 & e_2 \\ x_3 & y_3 & e_3\end{vmatrix}=\begin{vmatrix} x_2 & y_2 \\ x_3 & y_3\end{vmatrix}e_1-\begin{vmatrix} x_1 & y_1 \\ x_3 & y_3\end{vmatrix}e_2+\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2\end{vmatrix}e_3=$ $p_1\times p_2$ is the *vector* product of p_1,p_2.

To show that $p_n^\perp$ is orthogonal to each p_i, we compute the scalar product $p_n^\perp\cdot p_i=\sum\limits_{i=1}^{n}(-1)^{n+1}\det(i)e_i\cdot p_i$. Since $e_i\cdot p_i$ equals the i-th coordinate of the vector p_i, the last sum is the expansion of the $n\times n$ determinant obtained from the original $p_n^\perp$ above by replacing the last column with p_i. Since the resulting determinant contains two identical columns equal to p_i, we conclude that $p_n^\perp\cdot p_i=0$. □

Lemma B.1 holds when given vectors $p_1,\dots,p_{n-1}\in\mathbb{R}^n$ are linearly dependent, even if some $p_j=0$. Then $p_n^\perp=0$ is orthogonal to each p_j so that $p_n^\perp\cdot p_j=0$.

Definition B.2 extends a point-based representation from Definition 3.1 to dimensions $n\geq 2$. The key idea is to represent any m-point cloud $A\subset\mathbb{R}^n$ relative to (a simplex of) any base sequence of ordered points $p_1,\dots,p_{n-1}\in A$. If the vectors $p_1,\dots,p_{n-1}$ are linearly independent, they form with the vector $p_n^\perp$ from Lemma B.1 a (not necessarily orthogonal) basis in $\mathbb{R}^n$. Below we represent any point $p\in A$ by normalized scalar products, which are valid even if $p_1,\dots,p_{n-1}$ are linearly dependent.

Definition B.2 (point-based representation PR for $n\geq 2$)**.** *For any cloud $A\subset\mathbb{R}^n$ of m unordered points, the* center of mass *is $O(A)=\dfrac{1}{m}\sum\limits_{p\in A}p$. Shift A so that $O(A)$ is the origin $0\in\mathbb{R}^n$. The* radius *of A is $R(A)=\max\limits_{p\in A}|p|$. For any* basis *sequence of points $p_1,\dots,p_{n-1}\in A$, the* squared distance matrix $\mathrm{SD}(p_1,\dots,p_{n-1})$ *consists of $|p_i-p_j|^2$ for $i,j=0,\dots,n-1$, where $p_0=0$. Let $p_n^\perp$ be the vector in Lemma B.1. For any point $q\in A\setminus\{p_1,\dots,p_{n-1}\}$, the $n\times(m-n+1)$ matrix $M(A;p_1,\dots,p_{n-1})$ has a column of scalar products $q\cdot p_1,\dots,q\cdot p_n$. The* point-based representation $\mathrm{PR}(A;p_1,\dots,p_{n-1})$ *is the pair*

$$\big[\mathrm{SD}(p_1,\dots,p_{n-1}),\ M(A;p_1,\dots,p_{n-1})\big].$$

The normalized *representation* $\mathrm{NPR}(A;p_1,\dots,p_{n-1})$ *is obtained by dividing all components of $\mathrm{PR}(A;p_1,\dots,p_{n-1})$ by $R^2(A)$, except the last row of $M(A;p_1,\dots,p_{n-1})$, which is divided by $R^n(A)$.*

Lemma B.3 (PR under isometry)**.** *Let a point cloud $A\subset\mathbb{R}^n$ have a base sequence $(p_1,\dots,p_{n-1})$.*

(a) *Any rigid motion f of $\mathbb{R}^n$ respects point-based representations from Definition B.2 so that*

$$\mathrm{PR}(A;p_1,\dots,p_{n-1})=\mathrm{PR}(f(A);f(p_1),\dots,f(p_{n-1})).$$

(b) *For any orientation-reversing isometry f of $\mathbb{R}^n$, the representation $\mathrm{PR}(f(A);f(p_1),\dots,f(p_{n-1})$ differs from $\mathrm{PR}(A;p_1,\dots,p_{n-1})$ by reversing all signs in the last row of the matrix $M(A;p_1,\dots,p_{n-1})$.*

(c) *The normalized point-based representation* $\text{NPR}(A; p_1, \dots, p_{n-1})$ *in Definition B.2 is preserved by any composition of rigid motion and uniform scaling.*

***Proof of Lemma** B.3.* **(a)** Since rigid motion preserves distances and scalar products, all components of the point-based representation $\text{PR}(A; p_1, \dots, p_{n-1})$ are invariant.

(b) Using a composition with a suitable orientation-preserving isometry (rigid motion), one can assume that f is the mirror reflection in a linear hyperspace H containing the origin 0 and the base sequence $p_1, \dots, p_{n-1}$ of A. Since f preserves distances, $R(A)$ and $\text{SD}(A; p_1, \dots, p_{n-1})$ are invariant. Then f fixes all points from H including $p_1, \dots, p_{n-1}$, hence the vector p_n from Lemma B.1. Any point $q \in A \setminus p_1, \dots, p_{n-1}$ keeps its scalar product $q \cdot p_i$ for $i = 1, \dots, n-1$ and changes the sign of $q \cdot p_n$, because q and its mirror image $f(q)$ have opposite projections to p_n. The above arguments hold even if the base sequence $p_1, \dots, p_{n-1}$ is degenerate, not generating an $(n-1)$-dimensional subspace in $\mathbb{R}^n$. Then there are infinitely many choices of H above and $p_n = 0$, so the last row of $M(A; p_1, \dots, p_{n-1})$ consists of zeros.

(c) Under uniform scaling by a factor s, all squared distances and scalar products $q \cdot p_i$, $i = 1, \dots, n-1$, are multiplied by s^2. The vector $p_n^\perp$ from Lemma B.1 is multiplied by s^{n-1}, hence all scalar products $q \cdot p_n$ in the last row of $M(A; p_1, \dots, p_{n-1})$ are divided by $R^n(A)$. □

The *affine dimension* $0 \leq \text{aff}(A) \leq n$ of a cloud $A = \{p_1, \dots, p_m\} \subset \mathbb{R}^n$ is the maximum dimension of the vector space generated by all inter-point vectors $p_i - p_j$, $i, j \in \{1, \dots, m\}$. Then $\text{aff}(A)$ is an isometry invariant and is independent of an order of points of A. Any cloud A of 2 distinct points has $\text{aff}(A) = 1$. Any cloud A of 3 points that are not in the same straight line has $\text{aff}(A) = 2$.

Lemma B.4 provides a simple criterion for a matrix to be realizable by squared distances of a point cloud in $\mathbb{R}^n$.

Lemma B.4 (realization of distances)**.** ***(a)*** *A symmetric* $m \times m$ *matrix of* $s_{ij} \geq 0$ *with* $s_{ii} = 0$ *is realizable as a matrix of squared distances between points* $p_0 = 0, p_1, \dots, p_{m-1} \in \mathbb{R}^n$ if and only if *the* $(m-1) \times (m-1)$ *matrix* $g_{ij} = \dfrac{s_{0i} + s_{0j} - s_{ij}}{2}$ *has only non-negative eigenvalues.*

(b) *If the condition in (a) holds,* $\text{aff}(0, p_1, \dots, p_{m-1})$ *equals the number* $k \leq m-1 \leq n$ *of positive eigenvalues. Also in this case,* $g_{ij} = p_i \cdot p_j$ *define the* Gram matrix GM *of the vectors* $p_1, \dots, p_{m-1} \in \mathbb{R}^n$*, which are uniquely determined in time* $O(m^3)$ *up to an orthogonal map in* $\mathbb{R}^n$*.*

***Proof of Lemma** B.4.* **(a)** We extend Theorem 1 from (Dekster & Wilker, 1987) to the case $m < n+1$ and also justify the reconstruction of $p_1, \dots, p_{m-1}$ in time $O(m^3)$ uniquely in $\mathbb{R}^n$ up to an orthogonal map from the group $\text{O}(n)$.

The part *only if* $\Rightarrow$. Let a symmetric matrix S consist of squared distances between points $p_0 = 0, p_1, \dots, p_{m-1} \in \mathbb{R}^n$. For $i, j = 1, \dots, m-1$, the matrix with the elements

$$g_{ij} = \frac{s_{0i} + s_{0j} - s_{ij}}{2} = \frac{p_i^2 + p_j^2 - |p_i - p_j|^2}{2} = p_i \cdot p_j$$

is the Gram matrix, which can be written as $\text{GM} = P^T P$, where the columns of the $n \times (m-1)$ matrix P are the vectors $p_1, \dots, p_{m-1}$. For any vector $v \in \mathbb{R}^{m-1}$, we have

$$0 \leq |Pv|^2 = (Pv)^T (Pv) = v^T (P^T P) v = v^T \text{GM} v.$$

Since the quadratic form $v^T \text{GM} v \geq 0$ for any $v \in \mathbb{R}^{m-1}$, the matrix GM is positive semi-definite meaning that GM has only non-negative eigenvalues, see Theorem 7.2.7 in (Horn & Johnson, 2012).

The part *if* $\Leftarrow$. For any positive semi-definite matrix GM, there is an orthogonal matrix Q such that $Q^T \text{GM} Q = D$ is the diagonal matrix, whose $m-1$ diagonal elements are non-negative eigenvalues of GM. The diagonal matrix $\sqrt{D}$ consists of the square roots of eigenvalues of GM.

(b) The number of positive eigenvalues of GM equals the dimension $k = \text{aff}(\{0, p_1, \dots, p_{m-1}\})$ of the subspace in $\mathbb{R}^n$ linearly spanned by $p_1, \dots, p_{m-1}$. We may assume that all $k \leq n$ positive eigenvalues of GM correspond to the first k

coordinates of $\mathbb{R}^n$. Since $Q^T = Q^{-1}$, the given matrix $\mathrm{GM} = QDQ^T = (Q\sqrt{D})(Q\sqrt{D})^T$ becomes the Gram matrix of the columns of $Q\sqrt{D}$. These columns become the reconstructed vectors $p_1, \dots, p_{m-1} \in \mathbb{R}^n$.

If there is another diagonalization $\tilde{Q}^T \mathrm{GM} \tilde{Q} = \tilde{D}$ for $\tilde{Q} \in \mathrm{O}(n)$, then $\tilde{D}$ differs from D by a permutation of eigenvalues, which is realized by an orthogonal map, so we set $\tilde{D} = D$. Then $\mathrm{GM} = \tilde{Q}D\tilde{Q}^T = (\tilde{Q}\sqrt{D})(\tilde{Q}\sqrt{D})^T$ is the Gram matrix of the columns of $\tilde{Q}\sqrt{D}$.

The new columns differ from the previously reconstructed vectors $p_1, \dots, p_{m-1} \in \mathbb{R}^n$ by the orthogonal map $Q\tilde{Q}^T$. Hence the reconstruction is unique up to $\mathrm{O}(n)$-transformations. Computing eigenvectors $p_1, \dots, p_{m-1}$ needs a diagonalization of GM in time $O(m^3)$, see (**?**)section 11.5]press2007numerical. □

Though Lemma B.4 gives a two-sided criterion for realizability of distances by points $p_1, \dots, p_m \in \mathbb{R}^n$, the space of distance matrices is highly singular and cannot be easily sampled. Even $m = 4$ points in $\mathbb{R}^2$ have 6 distances that should satisfy a polynomial equation saying that the tetrahedron with these 6 edge lengths has volume 0.

So a randomly sampled matrix of potential distances for $m > n + 1$ is unlikely to be realizable by a cloud of m ordered points in $\mathbb{R}^n$. Hence Lemma B.4 for $m \le n + 1$ is complemented by Theorem B.7 describing the much more practical realizabilty of a point-based representation.

Chapter 3 in (Liberti & Lavor, 2017) discusses realizations of a complete graph given by a distance matrix in $\mathbb{R}^n$.

Lemma B.5(a) and later results hold for all clouds including degenerate ones, e.g. for 3 points in a straight line.

Any points $p_1, \dots, p_{n-1} \in A$ have $\mathrm{aff}(p_1, \dots, p_{n-1}) \le n - 2$. For example, any two distinct points in $A \subset \mathbb{R}^3$ generate a straight line. Lemma B.5(c) proves that $\mathrm{PR}(A; p_1, \dots, p_{n-1})$ suffices to reconstruct a cloud $A \subset \mathbb{R}^n$ for a suitable sequence $p_1, \dots, p_{n-1}$. In $\mathbb{R}^2$, any point $p_1 \neq O(A)$ forms a suitable $\{p_1\}$. In $\mathbb{R}^3$, one can choose any distinct points $p_1, p_2 \in A$ so that the infinite straight line via p_1, p_2 avoids $O(A)$.

If there are no such p_1, p_2, then $A \subset \mathbb{R}^3$ is contained in a straight line L, so $\mathrm{aff}(A) = 1$. In this degenerate case, the stronger condition $\mathrm{aff}(O(A) \cup \{p_1, \dots, p_{n-1}\}) = \mathrm{aff}(A)$ will help reconstruct $A \subset L$ by using any point $p_1 \neq O(A)$. The first step is to reconstruct any ordered sequence from its distance matrix in Lemma B.5(a).

Lemma B.5 improves Lemma E.5 in (Widdowson & Kurlin, 2023) by justifying a time for a point cloud reconstruction based on Lemma B.4.

Lemma B.5 (reconstruction)**.** ***(a)*** *Any sequence of ordered points $p_1, \dots, p_m$ in $\mathbb{R}^n$ can be reconstructed (uniquely up to isometry) from the matrix of the Euclidean distances $|p_i - p_j|$ in time $O(m^3)$. If all distances are divided by $R = \max\limits_{i=1,\dots,m} |p_i|$, the reconstruction of $p_1, \dots, p_m$ is unique up to isometry and uniform scaling in $\mathbb{R}^n$.*

(b) *If $m \le n$, the uniqueness of reconstructions in part (a) remains true if we replace isometry by rigid motion in $\mathbb{R}^n$.*

(c) *Any cloud $A \subset \mathbb{R}^n$ of m unordered points can be reconstructed (uniquely up to rigid motion in $\mathbb{R}^n$) from a point-based representation $\mathrm{PR}(A; p_1, \dots, p_{n-1})$ in time $O(m^3)$ for any $p_1, \dots, p_{n-1} \in A$ with $\mathrm{aff}(O(A) \cup \{p_1, \dots, p_{n-1}\}) = \mathrm{aff}(A)$. If $\mathrm{aff}(A) = n$, then $\mathrm{aff}(O(A) \cup \{p_1, \dots, p_{n-1}\}) = n - 1$ suffices. Any cloud $A \subset \mathbb{R}^n$ has a suitable sequence $p_1, \dots, p_{n-1}$ in all cases.*

***Proof of Lemma** B.5.* **(a)** By translation, we can put p_1 at the origin $0 \in \mathbb{R}^n$. Let G be the $(m-1) \times (m-1)$ matrix $G_{ij} = \dfrac{p_i^2 + p_j^2 - |p_i - p_j|^2}{2} = p_i \cdot p_j$ constructed from squared distances between $p_1 = 0, \dots, p_m$ for $i, j = 2, \dots, m$. By Lemma B.4 if G has $k \le n$ positive eigenvalues, then $p_1 = 0, \dots, p_m$ can be uniquely determined up to isometry in $\mathbb{R}^k \subset \mathbb{R}^n$ in time $O(m^3)$. If all distances are divided by the same radius $R(p\{m\})$, the above construction guarantees uniqueness up to isometry and uniform scaling.

(b) If $m \le n$, any mirror images of $p\{m\} \subset \mathbb{R}^n$ after a suitable rigid motion in $\mathbb{R}^n$ can be assumed to belong to an $(n-1)$-dimensional hyperspace $H \subset \mathbb{R}^n$, where they are matched by a mirror reflection $H \to H$ with respect to an $(n-2)$-dimensional subspace $S \subset H$, which is realized by the 180° orientation-preserving rotation of $\mathbb{R}^n$ around S.

(c) We will reconstruct a cloud $A \subset \mathbb{R}^n$ so that the center of mass $O(A)$ is the origin $0 \in \mathbb{R}^n$. If $\text{aff}(A) = k < n$, the cloud $A \subset \mathbb{R}^n$ is contained in an affine k-dimensional subspace, which can be rigidly moved to the linear subspace $\mathbb{R}^k \subset \mathbb{R}^n$ for the first k of n coordinates in $\mathbb{R}^n$.

It suffices to reconstruct $A \subset \mathbb{R}^k$ up to rigid motion in $\mathbb{R}^k$. Since $\text{aff}(0, p_1, \dots, p_{n-1}) = k$, some k vectors (say) $p_1, \dots, p_k$ from $p_1, \dots, p_{n-1}$ form a linear basis of $\mathbb{R}^k$. The k points $p_1, \dots, p_k$ are uniquely reconstructed up to rigid motion in $\mathbb{R}^k$ by part (b). Any other point $q \in A \setminus \{p_1, \dots, p_k\}$ is uniquely determined by its projections $(q \cdot p_i)/|p_i|$, which can be found from the first $k < n$ rows of the matrix $M(A; p_1, \dots, p_{n-1})$ for the point q, see Definition B.2.

In the generic case $\text{aff}(A) = n$, the condition $\text{aff}(0, p_1, \dots, p_{n-1}) = n-1$ means that $p_1, \dots, p_{n-1}$ are linearly independent and hence form a linear basis of $\mathbb{R}^n$ with the extra vector $p_n^\perp$ from Lemma B.1. The sequence $(0, p_1, \dots, p_{n-1})$ of n points can be uniquely reconstructed up to rigid motion in $\mathbb{R}^n$ by part (b). Any other point $q \in A \setminus \{p_1, \dots, p_{n-1}\}$ is uniquely determined by its projections $\dfrac{q \cdot p_i}{|p_i|}$ to the n basis vectors $p_1, \dots, p_{n-1}, p_n^\perp$, which can be found from the column of $M(A; p_1, \dots, p_{n-1})$ for q. □

Lemma B.5(b) for $m = n = 3$ implies that any triangle is determined by its sides up to rigid motion in $\mathbb{R}^3$. For example, the sides $3, 4, 5$ define a right-angled triangle whose mirror images are not related by rigid motion inside a plane $H \subset \mathbb{R}^3$, but are matched by composing a suitable rigid motion in H and a $180°$ rotation of $\mathbb{R}^3$ around a line in H.

Lemma B.6 (smoothness of PR)**.** *For any cloud $A \subset \mathbb{R}^n$ and a base sequence $p_1, \dots, p_{n-1} \in A$, all components of $\text{PR}(A; p_1, \dots, p_{n-1})$ have continuous partial derivatives (of any order) with respect to all (coordinates of) points of A as long as $R(A) > 0$, so some points of A remain distinct.*

***Proof of Lemma** B.6.* The point-based representation $\text{PR}(A; p\{n-1\})$ consists of squared distances in the matrix $\text{SD}(p\{n-1\})$ and scalar products in the matrix $M(A; p\{n-1\})$ of all points $q \in A \setminus p\{n-1\}$ with the vectors $p_1, \dots, p_{n-1}$ from the base sequence $p\{n-1\}$ and the vector $p_n \perp p_1, \dots, p_{n-1}$ from Lemma B.1. All these components are polynomials in the coordinates of the points of A, so have all continuous partial derivatives. □

Theorem B.7 extends Theorem 3.3 to dimensions $n \geq 2$.

Theorem B.7 (realizability of abstract PR)**.** *Let S be a symmetric $n \times n$ matrix of $s_{ij} \geq 0$ with $s_{ii} = 0$. Let M be any $n \times (m-n+1)$ matrix for $m \geq n$. The pair $[S, M]$ is realizable as a point-based representation $\text{PR}(A; p_1, \dots, p_{n-1})$ for a cloud $A \subset \mathbb{R}^n$ of m points with $O(A) = 0$ and a base sequence $p_1, \dots, p_{n-1}$ if and only if (1) the $(n-1) \times (n-1)$ matrix $G_{ij} = \dfrac{1}{2}(s_{1i} + s_{1j} - s_{ij})$ has only positive eigenvalues, which uniquely determines $p_1, \dots, p_{n-1}$ up to isometry, and (2) $\sum\limits_{j=1}^{n-1} (p_i \cdot p_j) + \sum\limits_{j=1}^{m-n+1} M_{ij} = 0$ for $i = 1, \dots, n$, where $p_n = p_n^\perp$ is the orthogonal vector from Lemma B.1.*

***Proof of Theorem** B.7.* The realizability of S as a matrix of squared distances between n points $0, p_1, \dots, p_{n-1}$ from the base sequence $p_1, \dots, p_{n-1}$ follows from Lemma B.4. The orthogonal vector $p_n^\perp$ (also denoted by p_n here for uniformity) from Lemma B.1 complements $p_1, \dots, p_{n-1}$ to a linear basis of $\mathbb{R}^n$. By Definition B.2, every element M_{ij} of the matrix $M = M(A; p_1, \dots, p_{n-1})$ equals $p_i \cdot q$ for some $q \in A \setminus \{p_1, \dots, p_{n-1}\}$, where $i = 1, \dots, n$.

Hence $\sum\limits_{j=1}^{n-1} (p_i \cdot p_j) + \sum\limits_{j=1}^{m-n+1} M_{ij} = 0$ can be rewritten as $p_i \cdot (\sum\limits_{p \in A} p) = 0$ for $i = 1, \dots, n$. These n equations mean that $O(A) = \dfrac{1}{m} \sum\limits_{p \in A} p$ is at the origin $0 \in \mathbb{R}^n$.

Conversely, for any M satisfying condition (2), we interpret every column $(M_{1j}, \dots, M_{nj})^T$ as a vector of scalar products $(q \cdot p_1, \dots, q \cdot p_n)$, which determine a position of a point $q \in A \setminus \{p_1, \dots, p_{n-1}\}$ in the basis $p_1, \dots, p_n$. □

In Theorem B.7, condition (2) is equivalent to $O(A) = 0 \in \mathbb{R}^n$ and implies that $m - n$ columns of M consist of free parameters, which determine the remaining column.

For $n = 2$, condition (1) means only that $s_{12} > 0$, so the distance between the points $p_0 = 0$ and p_1 is positive.

For $n = 3$, condition (1) about positive eigenvalues of the 2×2 matrix G means that 3 distances $a \le b \le c$ between points $0, p_1, p_2$ in $\mathbb{R}^3$ satisfy $a > 0$ and $a + b > c$, so the triangle on $0, p_1, p_2$ is non-degenerate. By the cosine theorem $p_1 \cdot p_2 = \frac{1}{2}(a^2 + b^2 - c^2)$, so the matrix $G = \begin{pmatrix} a^2 & \frac{1}{2}(a^2+b^2-c^2) \\ \frac{1}{2}(a^2+b^2-c^2) & b^2 \end{pmatrix}$ has $a^2 > 0$ and a positive determinant:

$$\begin{aligned}&4\det G = 4a^2b^2 - (a^2+b^2-c^2)^2 = \\ &(c^2 - (a^2 - 2ab + b^2))((a^2 + 2ab + b^2) - c^2) = \\ &(c^2 - (a-b)^2)((a+b)^2 - c^2) > 0.\end{aligned}$$

Assuming that $0 < a \le b \le c$, the last inequality is equivalent to one triangle inequality $a + b > c$.

Now we extend a point-based representation from Definition B.2 to a complete invariant of a point cloud A under rigid motion in $\mathbb{R}^n$. In applications, A can have distinguished points, for example, heavy atoms in atomic clouds, which can be used to minimize choices for $p_1, \dots, p_{n-1}$.

Definition B.8 will extend Definition 3.4 to $n > 2$ by combining all $\mathrm{PR}(A; p_1, \dots, p_{n-1})$ in a nested invariant by dropping points $p_1, \dots, p_{n-1} \in A$ one at a time. This invariant is needed only for comparisons (metric computations), while any cloud A can be stored in computer memory as a single $\mathrm{PR}(A; p_1, \dots, p_{n-1})$ due to Theorem B.7.

Definition B.8 (NDP : Nested Distributed Projection)**.** *Let $A \subset \mathbb{R}^n$ be any cloud of m unordered points. For any ordered points $p_1, \dots, p_{n-2} \in A$, let $\mathrm{NDP}(A; p_1, \dots, p_{n-2})$ be the unordered collection of $\mathrm{PR}(A; p_1, \dots, p_{n-1})$ for all points $p_{n-1} \in A \setminus \{p_1, \dots, p_{n-2}\}$. Similarly, for any $1 \le k \le n-2$, let $\mathrm{NDP}(A; p_1, \dots, p_{k-1})$ be the unordered collection of $\mathrm{NDP}(A; p_1, \dots, p_k)$ for all points $p_k \in A \setminus \{p_1, \dots, p_{k-1}\}$. For $k = 1$, the full* Nested Distributed Projection $\mathrm{NDP}(A)$ *depends only on A.*

For $n = 2$ and any cloud $A \subset \mathbb{R}^2$, the Nested Distributed Projection $\mathrm{NDP}(A)$ in Definition B.8 is the same as in Definition 3.4, i.e. $\mathrm{NDP}(A)$ is the unordered collection of point-based representations $\mathrm{PR}(A; p_1)$ for all $p_1 \in A$.

For $n = 3$ and any $A \subset \mathbb{R}^3$, the Nested Distributed Projection $\mathrm{NDP}(A)$ is the unordered collection of $\mathrm{NDP}(A; p_1)$ for all $p_1 \in A$. Each $\mathrm{NDP}(A; p_1)$ is the unordered collection of $\mathrm{PR}(A; p_1, p_2)$ for all $p_2 \in A \setminus \{p_1\}$.

Similarly to Definition 3.4, if a cloud A has internal symmetries as in Example 3.2, one can collapse identical objects to a single one with a weight to speed up computations. We avoid collapsing only to simplify arguments for $n > 2$.

Lemma B.5(c) implies that any cloud $A \subset \mathbb{R}^n$ of m unordered points can be reconstructed from $\mathrm{NDP}(A)$ uniquely up to rigid motion. Indeed, $\mathrm{NDP}(A)$ contains (nested) PRs depending on all possible $n-1$ points $p_1, \dots, p_{n-1} \in A$. At least one $\mathrm{PR}(A; p_1, \dots, p_{n-1})$ satisfies Lemma B.5(c) and suffices to reconstruct A uniquely up to rigid motion.

In Theorem B.9 for $n > 2$, the equality $\mathrm{NDP}(A) = \mathrm{NDP}(B)$ means a bijection $\beta : \mathrm{NDP}(A) \to \mathrm{NDP}(B)$ respecting the nested structure of all PRs in Definition B.8.

In detail, for any $1 \le k \le n-1$ and points $p_1, \dots, p_k$, the bijection β matches $\mathrm{NDP}(A; p_1, \dots, p_k)$ with a unique $\mathrm{NDP}(B; q_1, \dots, q_k)$ for some $q_1, \dots, q_k \in B$.

If $n = 3$, then β matches every $\mathrm{NDP}(A; p_1)$ with a unique $\mathrm{NDP}(B; q_1)$ in the sense that this bijection $\mathrm{NDP}(A; p_1) \to \mathrm{NDP}(B; q_1)$ matches $\mathrm{PR}(A; p_1, p_2)$ for every $p_2 \in A \setminus \{p_1\}$ with $\mathrm{PR}(B; q_1, q_2)$ for a unique $q_2 \in B - \{q_1\}$.

Theorem B.9 (completeness of NDP)**.** *The Nested Distributed Projection is complete in the sense that any clouds $A, B \subset \mathbb{R}^n$ of m unordered points are related by rigid motion in $\mathbb{R}^n$ if and only if $\mathrm{NDP}(A) = \mathrm{NDP}(B)$ so that there is a bijection $\mathrm{NDP}(A) \to \mathrm{NDP}(B)$ matching all PRs.*

***Proof of Theorem** B.9.* The part *only if* : we will prove that any rigid motion f moving the cloud A to $B = f(A)$ implies that $\mathrm{NDP}(A) = \mathrm{NDP}(B)$. By Lemma B.3(a) the rigid motion f matches every $\mathrm{PR}(A; p_1, \dots, p_{n-1})$ from $\mathrm{NDP}(A)$ with $\mathrm{PR}(B; f(p_1), \dots, f(p_{n-1}))$. Then, for any $1 \le k \le n-2$ and $p_1, \dots, p_k \in A$, we get a bijection $\mathrm{NDP}(A; p_1, \dots, p_k) \to \mathrm{NDP}(B; f(p_1), \dots, f(p_k))$ Hence f induces a bijecton $\mathrm{NCP}(A) \to \mathrm{NCP}(B)$ between all PRs respecting the nested structure in Definition B.8.

The part *if* : $\mathrm{NDP}(A) = \mathrm{NDP}(B)$ will guarantee a rigid motion f moving the cloud A to $B = f(A)$. Choose any base sequence $p_1, \dots, p_{n-1} \in A$ that suffices for a unique reconstruction of $A \subset \mathbb{R}^n$ up to rigid motion in Lemma B.5(c). The given bijection $\mathrm{NDP}(A) \to \mathrm{NDP}(B)$ matches $\mathrm{PR}(A; p_1, \dots, p_{n-1})$ with an equal $\mathrm{PR}(B; q_1, \dots, q_{n-1})$ for some $q_1, \dots, q_{n-1} \in B$.

Lemma B.5(c) implies that a reconstruction of A, B from $\mathrm{PR}(A; \sigma(p_1, \dots, p_{n-1})) = \mathrm{PR}(B; q_1, \dots, q_{n-1})$ is unique up to rigid motion in $\mathbb{R}^n$ so that A, B are matched by a rigid motion f as required. If $\mathrm{aff}(A) = \mathrm{aff}(B) < n$, this motion f may not be unique. For example, any clouds $A, B \subset \mathbb{R}^3$ that are contained in a straight line $L \subset \mathbb{R}^3$ are pointwise fixed by any rotation around the line L. □

C. Maps of cloud spaces and explicit computations of invariants

This section explains how cloud spaces can be visualized by considering the previously known and new types of 4-point clouds (quads) in $\mathbb{R}^2$. This geographic-style approach extends to any number m of points in $\mathbb{R}^n$.

For any cloud $A \subset \mathbb{R}^n$, the center $O(A) = 0 \in \mathbb{R}^n$ is the origin. For $n = 2$, let $p\{1\}$ consist of a single point $p_1 \in A$ with $|p_1| = R(A) = R$. We can fix $p_1 = (R, 0)$ in $\mathbb{R}^2$. Then all points $p_2, \dots, p_m$ are in the disk $D = \{x^2 + y^2 \le R^2\}$. Since $\sum\limits_{i=2}^{m} p_i = -p_1 = (-R, 0)$, p_m is determined from $p_2, \dots, p_{m-1} \in D$ that satisfy only one equation

$$R^2 \ge |p_m|^2 = |(R,0)^T + \sum_{i=2}^{m-1} p_i|^2 = (R + x)^2 + y^2,$$

where (x, y) are the coordinates of $s = \sum\limits_{i=2}^{m-1} p_i$. The domain of s is the intersection $J = D \cap \{(R + x)^2 + y^2 \le R^2\}$.

For $m = 3$, we have $s = (x, y) = p_2$. The symmetry $p_2 \leftrightarrow p_3$ allows us to choose any p_2 in the left half (yellow) D_3 of the intersection J in Fig. 13 (left). Then the Rigid Cloud Space $\mathrm{CRS}(\mathbb{R}^n; 3)$ is parametrized by any radius $R > 0$ and $p_2 \in D_3$. All equilateral triangles have $p_2 = (-\frac{1}{2}R, \pm\frac{\sqrt{3}}{2}R)$. All isosceles triangles have p_2 in the boundary ∂D_3 whose points should be identified under $(x, y) \mapsto (x, -y)$. All $p_2 = (x, 0)$ with $-R \le x \le -\frac{1}{2}R$ represent degenerate triangles with the vertices $(R, 0)$, $(x, 0)$, $(-R - x, 0)$ in the same line.

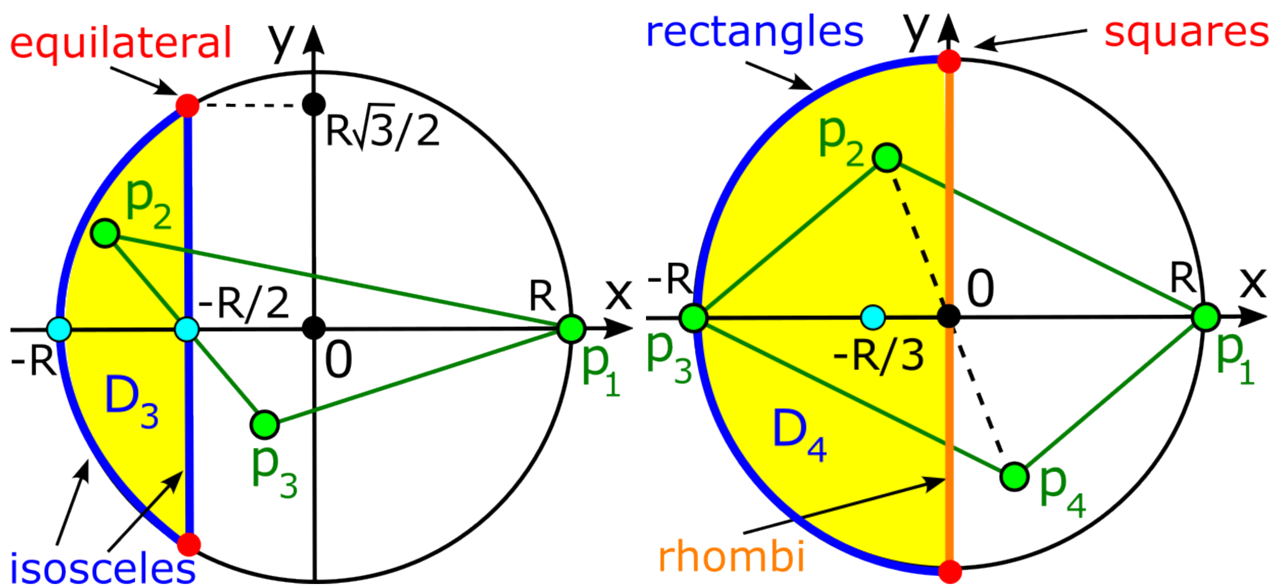

Figure 13. The spaces in yellow for triangles (D_3) and parallelograms (D_4) under rigid motion and uniform scaling in $\mathbb{R}^2$.

For $m = 4$, we can choose $s = p_2 + p_3 \in J$, then any p_3 in the disk with the radius R and center s so that $|p_2| = |p_3 - s| \le R$. For any parallelogram in $\mathbb{R}^2$, its vertex cloud A has a longest diagonal between (say) p_1, p_3 that should be at $(\pm R, 0)$. All possible $s = p_2 + (-R, 0) \in J$ mean that p_2 can be anywhere in D. Due to the symmetry $p_2 \leftrightarrow p_4$, the left half D_4 of D in Fig. 13 (right) is the subspace of all parallelograms in $\text{DCS}(\mathbb{R}^2; 4) = \text{CRS}(\mathbb{R}^2; 4)/\text{scaling}$.

Similarly for $m > 4$, $n \geq 2$, we can sequentially sample points $p_2, \dots, p_{m-1}$ from allowed disks (high-dimensional for $n > 2$) to get a unique representation of A under rigid motion. The symmetry $f : (x, y) \mapsto (x, -y)$ on D identifies mirror images of A. $\text{CIS}(\mathbb{R}^n; m)$ is the quotient of $\text{CRS}(\mathbb{R}^n; m)$ under $(x, y) \sim (x, -y)$, take the upper halves of D_3, D_4 for triangles and parallelograms, respectively.

We expand Fig. 13 above to illustrate severak important subspaces in the Isometry Cloud Space $\text{CIS}(\mathbb{R}^2; m)$ and the Similarity Cloud Space $\text{CSS}(\mathbb{R}^2; m)$ for $m = 3, 4$. For simplicity, we call all clouds of 3 and 4 unordered points triangles and quadrilaterals, respectively.

However, all these polygons are considered equivalent when we re-order their vertices. If all m points are ordered, parametrizations of the resulting shape spaces were studied in geometry (Kapovich & Millson, 1996) and shape theory (Kendall et al., 2009). We focus on the much harder quotient spaces of m unordered points.

Theorem B.7 explicitly describes all realizable Point-based Representations. Though the same point cloud $A \subset \mathbb{R}$ can have many $\text{PR}(A; p\{n-1\})$ depending on a base sequence $p\{n-1\} \subset A$, we can easily sample any of them and always reconstruct A, while random sampling distance-based invariants doesn't guarantee the existence of A because of extra relations between inter-point distances.

Though $\text{PR}(A; p\{n-1\})$ consists of scalar products $q \cdot p_i$ with basis vectors $p_1, \dots, p_n$, it is easier to visualize the isometry spaces by directly using some points $q \in A$ as parameters instead of their projections.

Case $m = 3$ of triangles is the same in all dimensions $n \geq 2$. We consider $\mathbb{R}^2$ for simplicity. Fig. 13 (left) showed the Dilation Cloud Space $\text{DCS}(\mathbb{R}^2; 3)$ of triangles A modulo rigid motion and uniform scaling in $\mathbb{R}^2$. We assume that the center of mass is at the origin: $C(A) = 0$ in $\mathbb{R}^2$. After the radius $R = 1$ of A is fixed up to scaling, we also fix the first vertex at $p_1 = (R, 0)$. Then $\text{DCS}(\mathbb{R}^2; 3)$ is parametrized by the second vertex $p_2 \in D_3$, because the vertex p_3 is uniquely determined by $p_1 + p_2 + p_3 = 0$.

The blue boundary of $\text{DCS}(\mathbb{R}^2; 3)$ consists of points p_2 that define isosceles triangles. The vertical part of the blue boundary in Fig. 14 (left) represents all isosceles triangles with a unique angle (not equal to two equal ones) less than $60°$. The round part of the blue boundary in Fig. 14 (right) represents all isosceles triangles with a unique angle greater than $60°$. These boundary parts meet at the red points $(-\frac{R}{2}, \pm\frac{\sqrt{3}}{2}R)$ representing all equilateral triangles.

If $p_2 = (x, 0)$ for $-R \leq x \leq -\frac{R}{2}$, then $p_3 = (-R - x, 0)$, so the triangle generates to three points in the line. In the yellow space $D_3 = \text{CSS}^o(\mathbb{R}^2; 3)$, the mirror reflection $(x, y) \mapsto (x, -y)$ maps every isosceles triangle to itself, more exactly, to an equivalent triangle under rigid motion. Hence all points of the blue boundary of D_3 should be identified under $(x, y) \mapsto (x, -y)$. Then the space D_3 of all triangles (including degenerate ones) under rigid motion and uniform scaling can be visualized as a topological sphere S^2 whose the northern and southern hemispheres are obtained from the upper and lower halves of D_3.

Case $m = 4$ of quadrilaterals in $\mathbb{R}^2$. Fix the center of mass $O(A) = 0 \in \mathbb{R}^2$ at the origin, the radius $R(A) = R$, and a most distant (from 0) point p_1 at $(R, 0)$. The other vertices p_2, p_3, p_4 belong to the disk $D = \{x^2 + y^2 \leq R^2\}$ and have the shifted center of mass $\frac{p_2+p_3+p_4}{3} = (-\frac{R}{3}, 0)$. Hence, for a fixed radius R, the space $\text{CSS}(\mathbb{R}^2; 4)$ is 4-dimensional.

The subspace of parallelograms in $\text{CSS}(\mathbb{R}^2; 4)$ is 2-dimensional. For any parallelogram A, its other most distant vertex is $p_3 = (-R, 0)$ opposite to p_1 with respect to 0. Then $p_2 + p_4 = 0$ and the symmetry $p_2 \leftrightarrow p_4$ allows us to consider only p_2 in the yellow half-disk D_4, which uniquely determines its symmetric image p_4 in Fig. 13 (left).

The round (blue) boundary of D_4 in Fig. 15 (left) represents all rectangles inscribed in the circle $x^2 + y^2 = R^2$. The vertical (orange) boundary of D_4 in Fig. 15 (right) represents all rhombi with equal sides. The reflection $(x, y) \mapsto (x, -y)$ maps any parallelogram to its mirror image and preserves the equivalence class (up to rigid motion) of any rectangle or rhombus, which are mirror-symmetric. Hence all points on the boundary of D_4 should be identified under $(x, y) \mapsto (x, -y)$. The resulting quotient is a topological sphere S^2 as D_3 for all triangles, unsurprisingly because a parallelogram can be considered as a double triangle.

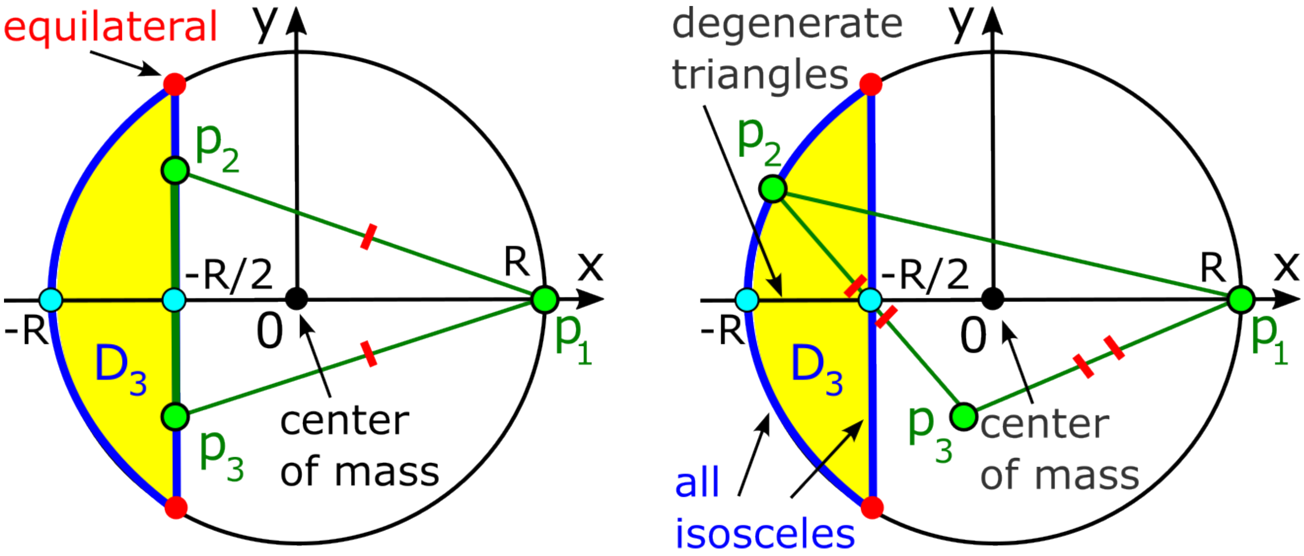

Figure 14. The (blue) subspace of all isosceles triangles in $\mathrm{CSS}(\mathbb{R}^2;3)$. **Left**: isosceles triangles with $|p_1-p_2|=|p_1-p_3|$. **Right**: isosceles triangles with $|p_3-p_1|=|p_3-p_2|$.

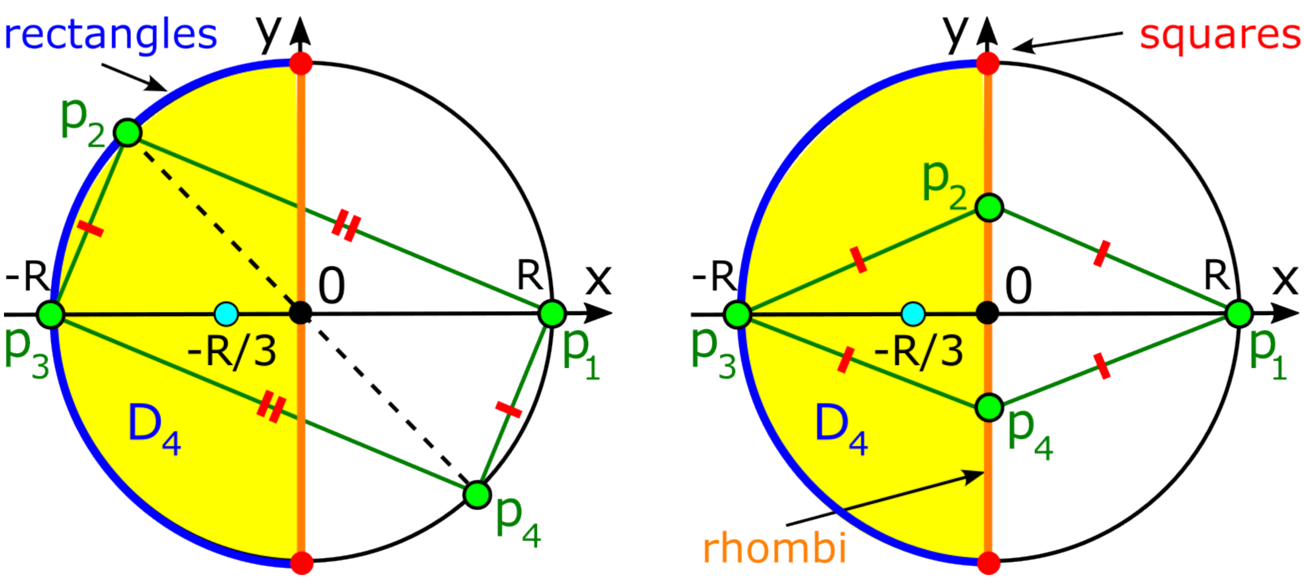

Figure 15. The (yellow) subspace D_4 of all parallelograms with $p_1=(R,0)$ and $p_3=(-R,0)$ in $\mathrm{CSS}(\mathbb{R}^2;4)$. **Left**: the (blue) subspace of rectangles. **Right**: the (orange) subspace of rhombi.

Another interesting case is when one of the vertices $p_3=(x,0)$ belongs to the x-axis for $x\in[-R,R]$. Then the (horizontal line passing through) diagonal joining p_1,p_3 intersects another diagonal at its mid-point $\frac{p_2+p_4}{2}=(x_{2,4},0)$ for $x_{2,4}=-\frac{x+R}{2}\in[-R,0]$. The resulting cloud A can be called a *quadrilateral with a median diagonal*, briefly *qmed*. If a qmed A is also symmetric with respect to its median diagonal, the A has two pairs of equal sides and is often called a *kite*, see the kite K in Fig. 2 (right).

Since any kite is mirror-symmetric, the points $p_2=(x,y)$ and $p_4=(x,-y)$ represents the same kite up to rigid motion. Hence the (yellow) subspace of all kites in $\mathrm{CSS}(\mathbb{R}^2;4)$ is the upper half K_4 of the disk D in Fig. 16 (left). For points p_2 in the vertical line $x=-\frac{R}{3}$, we get a degenerate kites whose vertices p_2,p_3,p_4 are in the same straight line. If $p_2=(x,0)$,

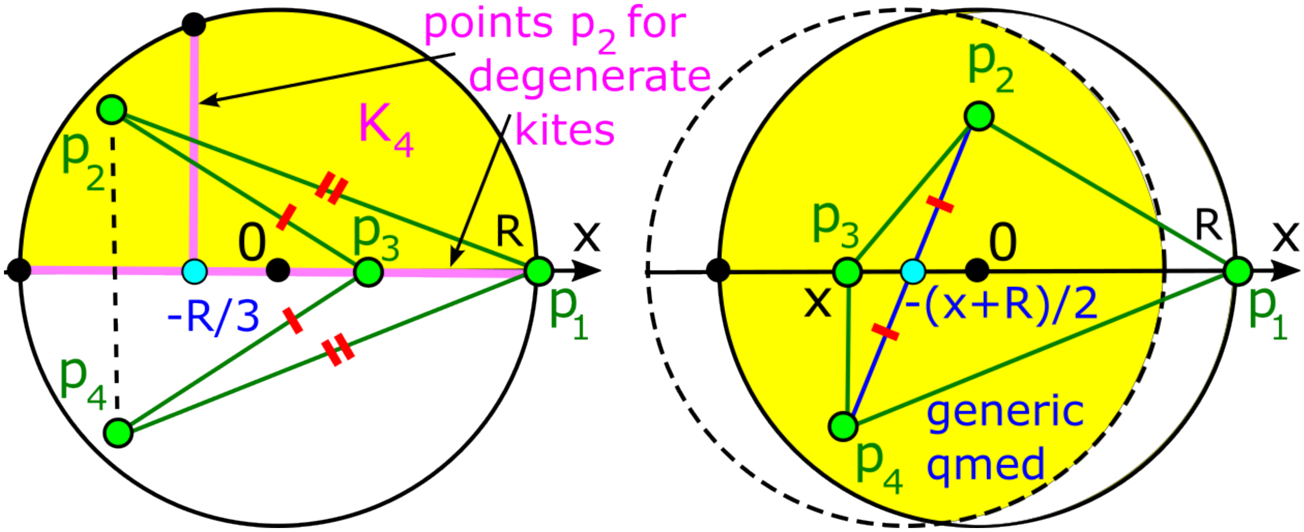

Figure 16. **Left**: the (yellow) subspace of kites in $\mathrm{CSS}(\mathbb{R}^2;4)$ parametrized by $p_2 \in K_4$. **Right**: the subspace of qmeds is parametrized by $x \in [-R,R]$ and p_2 in the yellow region.

the kite degenerates even further to the case of identical vertices $p_2 = p_4$.

So the subspace K_4 of kites in $\mathrm{CSS}(\mathbb{R}^2;4)$ is 2-dimensional, while the larger subspace of qmeds is 3-dimensional, parametrized by $x \in [-R,R]$ and a point p_2 that can take any position in the intersection of the disk $D = \{x^2+y^2 \leq R^2\}$ and its symmetric image with respect to the diagonal mid-point $(x_{2,4},0) = (-\frac{x+R}{2},0)$.

The full space $\mathrm{CSS}(\mathbb{R}^2;4)$ is parametrized by the sum $s = p_2 + p_3$ in the intersection $J = D \cap \{(R+x)^2+y^2 \leq R^2\}$ and then taking p_2 in the disk with the radius R and center s to guarantee that $|p_3| = |p_2 - s| \leq R$.

Case $m=4$ of tetrahedra in $\mathbb{R}^3$. In $\mathbb{R}^3$, we similarly fix the center of mass at the origin and the most distant points p_1 at $(R,0,0)$. The second most distant point p_2 (if not in the line through 0 and p_1) forms a base sequence p_1,p_2 and can be fixed at $(x,y,0)$ with $x^2+y^2 \leq R^2$, which determines the mid-point $p_{3,4}\frac{p_3+p_4}{2} = (-\frac{x+R}{2},-\frac{y}{2},0)$. Due to the symmetry $p_3 \leftrightarrow p_4$ around $p_{3,4}$, it remains to choose p_3 in the upper half ball with the center $p_{3,4}$ and radius $\sqrt{x^2+y^2}$.

The clouds in Example C.1 are instances of $C^{\pm}$ from Example 4.5: $K = C^+, T = C^-$ for $4a = b = c = 4d = 2\sqrt{2}$ and are easy enough to write their NDPs below.

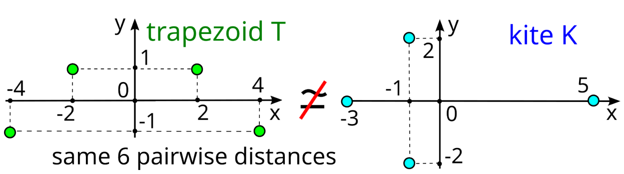

Figure 17. Non-isometric clouds of 4 points with the same 6 pairwise distances. **Left**: the trapezoid T has points $(\pm 2,1)$, $(\pm 4,-1)$. The kite K has $(5,0)$, $(-3,0)$, $(-1,\pm 2)$.

Example C.1 (4-point clouds T,K in Fig. 17)**.** *Both clouds $T,K \subset \mathbb{R}^2$ in Fig. 17 have the center of mass at the origin.*

(T) *The cloud T has the points $p_1 = (2,1)$, $p_2 = (-2,1)$, $p_3 = (-4,-1)$, $p_4 = (4,-1)$. For the basis point $p_1 = (2,1)$ with $|p_1|^2 = 5$ and orthogonal vector $p_1^\perp = (-1,2) \perp p_1$ from Lemma B.1, the point-based representation is* $\mathrm{PR}(T;p_1) = \left[5, \begin{pmatrix} -3 & -9 & 7 \\ 4 & 2 & -6 \end{pmatrix}\right]$.

For the second point $p_2 = (-2,1)$ with $|p_2|^2 = 5$, $p_2^\perp = (-1,-2)$, we have $\mathrm{PR}(T;p_2) = \left[5, \begin{pmatrix} -3 & 7 & -9 \\ -4 & 6 & -2 \end{pmatrix}\right]$*, which* *differs from* $\mathrm{PR}(T;p_1)$ *by the sign of the last row (up to a permutation of columns). The symmetries under* $p_1 \leftrightarrow p_2$ *(above) and* $p_3 \leftrightarrow p_4$ *(below) are explained by the reflection* $(x,y)\mapsto(-x,y)$ *mapping* T *to itself.*

For $p_3=(-4,-1)$ *with* $|p_3|^2=17$*,* $p_3^\perp=(1,-4)$*, we have* $\mathrm{PR}(T;p_3)=\left[17,\begin{pmatrix} -9 & 7 & -15 \\ -2 & -6 & 8 \end{pmatrix}\right]$.

For the fourth point $p_4=(4,-1)$ *with* $|p_4|^2=17$*,* $p_4^\perp=(1,4)$*, we have* $\mathrm{PR}(T;p_4)=\left[17,\begin{pmatrix} 7 & -9 & -15 \\ 6 & 2 & -8 \end{pmatrix}\right]$.

So $\mathrm{NDP}(T)$ *is the unordered set of the four* PR*s above.*

(K) *The cloud* K *has the points* $p_1=(5,0)$*,* $p_2=(-1,2)$*,* $p_3=(-3,0)$*,* $p_4=(-1,-2)$*.*

For the basis point $p_1=(5,0)$ *with* $|p_1|^2=25$ *and* $p_1^\perp=(0,5)\perp p_1$*, the point-based representation is* $\mathrm{PR}(K;p_1)=\left[25,\begin{pmatrix} -5 & -15 & -5 \\ 10 & 0 & -10 \end{pmatrix}\right]$.

For the second point $p_2=(-1,2)$ *with* $|p_2|^2=5$ *and* $p_2^\perp=(-2,-1)$*, we have* $\mathrm{PR}(K;p_2)=\left[5,\begin{pmatrix} -5 & 3 & 1 \\ -10 & 6 & 4 \end{pmatrix}\right]$.

For the third point $p_3=(-3,0)$ *with* $|p_3|^2=9$ *and* $p_3^\perp=(0,-3)$*, we have* $\mathrm{PR}(K;p_3)=\left[9,\begin{pmatrix} -15 & 3 & 3 \\ 0 & -6 & 6 \end{pmatrix}\right]$.

For the point $p_4=(-1,-2)$ *with* $|p_4|^2=5$ *and* $p_4^\perp=(2,-1)$*, we have* $\mathrm{PR}(K;p_4)=\left[5,\begin{pmatrix} -5 & 1 & 3 \\ 10 & -4 & -6 \end{pmatrix}\right]$.

So $\mathrm{NDP}(K)$ *is the unordered set of the four* PR*s above.*

$T\not\cong K$ *are distinguished by (unordered) squared distances to their centers:* $5,5,17,17$ *for* T*, and* $25,5,9,5$ *for* K*.*

Example C.2 finishes the computations of the Nested Distributed Projection (NDP) for the 4-point clouds $C^\pm\subset\mathbb{R}^2$ in Fig. 2, which we started in Example C.1.

The simultaneous swapping $a\leftrightarrow d$, $b\leftrightarrow c$ maps each cloud $C^\pm$ to its mirror image in the diagonal $x=y$ in $\mathbb{R}^2$, hence the metric between $C^\pm$ remains the same, which explains the symmetry of the top two plots in Fig. 18, 19, 20.

Example C.2 (4-point clouds $C^\pm$ in Fig. 2)**.** *In* $\mathbb{R}^2$*, consider the 4-point clouds* $C^\pm=\{p_1,p_2,p_3,p_4^\pm\}$*, where* $p_1=(4a,0)$*,* $p_2=(b,c)$*,* $p_3=-p_2=(-b,-c)$*,* $p_4^+=(0,4d)$*, and* $p_4^-=(0,-4d)$ *for parameters* $a,b,c,d\geq 0$*.*

After shifting the center $O(C^+)=(a,d)$ *to the origin* $(0,0)$*, the points of* C^+ *become* $p_1^+=(3a,-d)$*,* $p_2^+=(b-a,c-d)$*,* $p_3^+=(-a-b,-c-d)$*,* $\hat p_4^+=(-a,3d)$*.*

Each matrix $\mathrm{SD}(C^+;p)$ *is one squared distance* $|p|^2$*.*

$$\begin{aligned}
&\mathrm{SD}(C^+;p_1^+)=9a^2+d^2,\\
&\mathrm{SD}(C^+;p_2^+)=(a-b)^2+(c-d)^2,\\
&\mathrm{SD}(C^+;p_3^+)=(a+b)^2+(c+d)^2,\\
&\mathrm{SD}(C^+;\hat p_4^+)=a^2+9d^2.
\end{aligned}$$

For the second cloud C^-*, after shifting the center* $O(C^-)=(a,-d)$ *to the origin* $(0,0)$*, the points become* $p_1^-=(3a,d)$*,* $p_2^-=(b-a,d+c)$*,* $p_3^-=(-a-b,d-c)$*,* $\hat p_4^-=(-a,-3d)$*.*

Hence C^- *has the following squared distances to its center:*

$$\begin{aligned}
&\mathrm{SD}(C^-;p_1^-)=9a^2+d^2,\\
&\mathrm{SD}(C^-;p_2^-)=(a-b)^2+(c+d)^2,\\
&\mathrm{SD}(C^-;p_3^-)=(a+b)^2+(c-d)^2,\\
&\mathrm{SD}(C^-;\hat p_4^-)=a^2+9d^2.
\end{aligned}$$

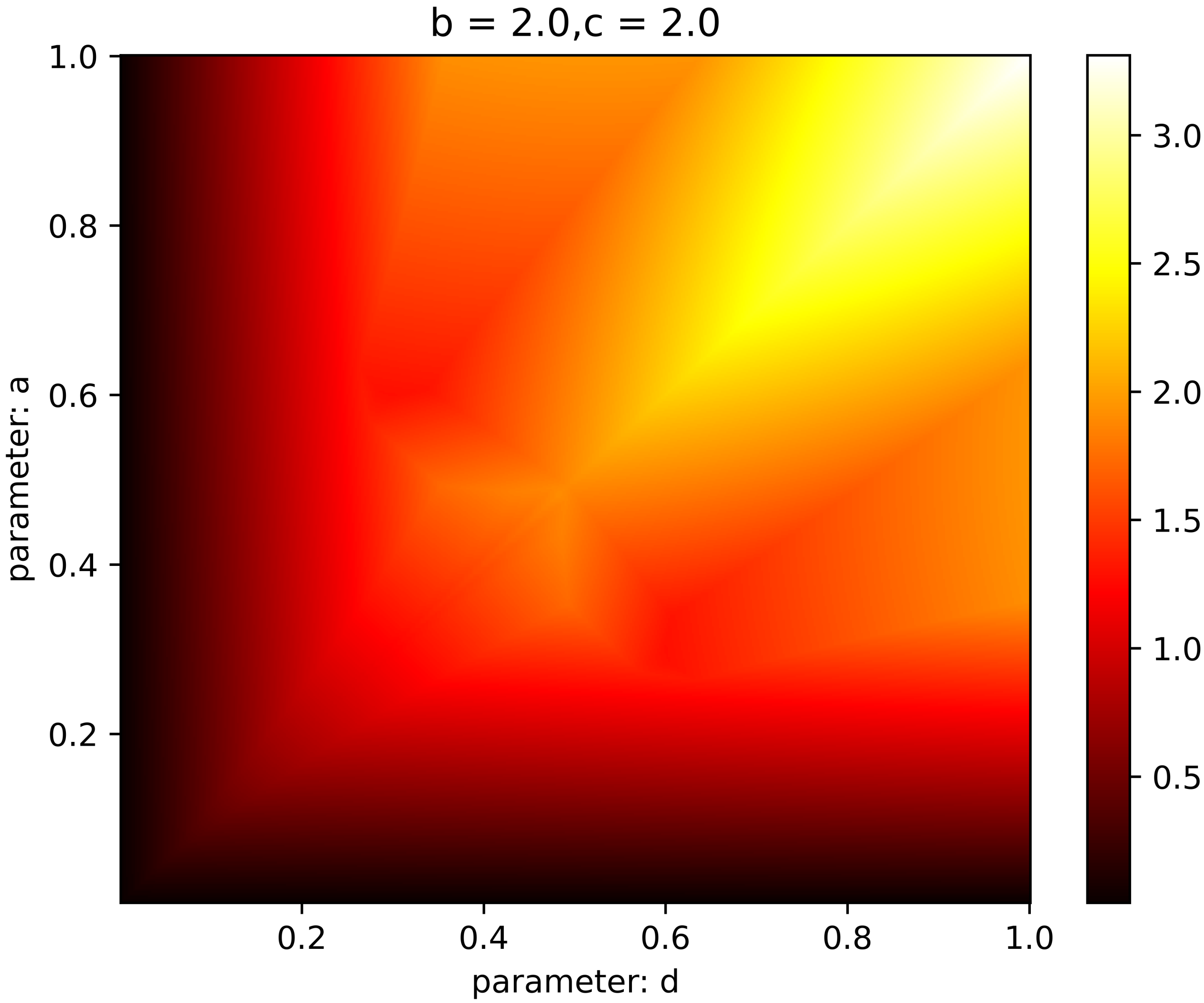

Figure 18. The Nested Bottleneck Metric NBM from Definition 4.4 for the 4-point clouds $C^{\pm} \subset \mathbb{R}^2$ with variable parameters a, d, see details in Example C.1.

The (unordered) collections of squared distances above differ unless at least one of a, b, c, d is zero. Indeed, the squared distances $9a^2+d^2$ and a^2+9d^2 are shared by $C^{\pm}$ but $\mathrm{SD}(C^+;p_2^+)$ is unique and cannot equal $\mathrm{SD}(C^-;p_2^-)$ or $\mathrm{SD}(C^-;p_3^-)$. Indeed, if all $a, b, c, d \neq 0$, then

$(a-b)^2+(c-d)^2 \neq (a-b)^2+(c+d)^2$ *or* $cd \neq 0$,
$(a-b)^2+(c-d)^2 \neq (a+b)^2+(c-d)^2$ *or* $ab \neq 0$.

If $d = 0$, then $p_4^{\pm} = (0,0)$, so the clouds $C^{\pm}$ are identical.

If $a = 0$, then $p_1 = (0,0)$ and $C^{\pm}$ are related by the $180°$ rotation around the origin: $(x,y) \mapsto (-x,-y)$.

If $b = 0$ or $c = 0$, then $C^{\pm}$ are related by the reflection $(x,y) \mapsto (x,-y)$, so distances cannot distinguish these mirror images. We compute $\mathrm{NDP}(C^{\pm})$ below to distinguish all non-rigidly equivalent $C^+ \not\cong C^-$, see Fig. **??**.

For the basis point p_1^+, the matrix $\mathrm{SD}(C^+;p_1^+) = 9a^2+d^2$ is the single squared distance. Lemma B.1 gives the orthogonal vector $q_1^+ = (d, 3a) \perp p_1^+$. $M(C^+;p_1^+)$ consists of the 3 unordered columns

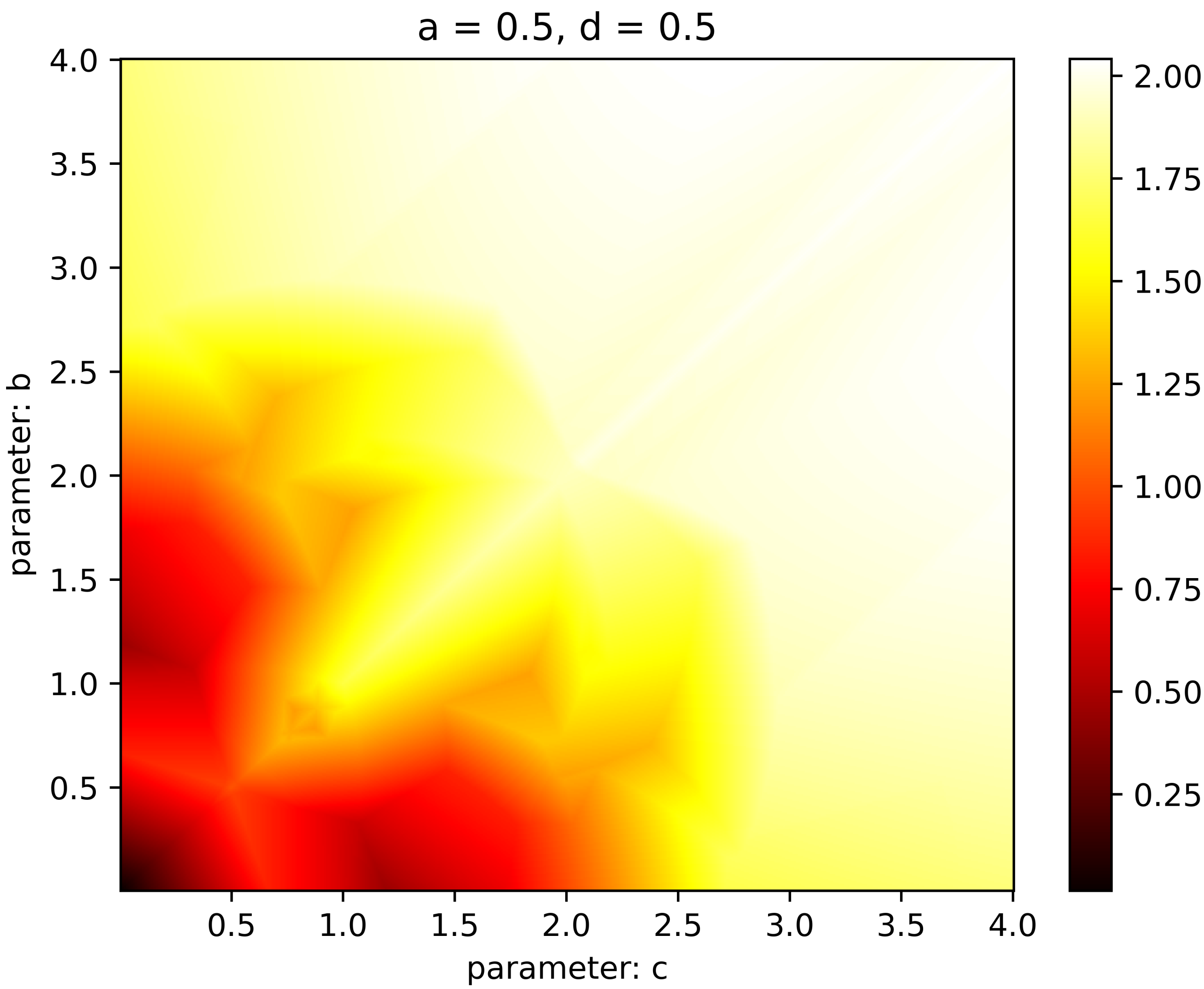

Figure 19. The Nested Bottleneck Metric NBM from Definition 4.4 for the 4-point clouds $C^{\pm} \subset \mathbb{R}^2$ with variable parameters b, c, see details in Example C.1.

$\begin{pmatrix} p_2^+ \cdot p_1^+ \\ p_2^+ \cdot q_1^+ \end{pmatrix} = \begin{pmatrix} 3a(b-a)+d(d-c) \\ d(b-a)+3a(c-d) \end{pmatrix}$,
$\begin{pmatrix} p_3^+ \cdot p_1^+ \\ p_3^+ \cdot q_1^+ \end{pmatrix} = \begin{pmatrix} -3a(a+b)+d(c+d) \\ -d(a+b)-3a(c+d) \end{pmatrix}$,
$\begin{pmatrix} \hat{p}_4^+ \cdot p_1^+ \\ \hat{p}_4^+ \cdot q_1^+ \end{pmatrix} = \begin{pmatrix} -3(a^2+d^2) \\ 8ad \end{pmatrix}$. *The second point* $p_2^+ = (b-a, c-d)$ *has the orthogonal vector* $q_2^+ = (d-c, b-a) \perp p_2^+$, $\mathrm{SD}(C^+; p_2^+) = (a-b)^2 + (c-d)^2$ *and* $M(C^+; p_2^+)$ *consisting of the 3 unordered columns*
$\begin{pmatrix} p_1^+ \cdot p_2^+ \\ p_1^+ \cdot q_2^+ \end{pmatrix} = \begin{pmatrix} 3a(b-a)+d(d-c) \\ 3a(d-c)+d(a-b) \end{pmatrix}$,
$\begin{pmatrix} p_3^+ \cdot p_2^+ \\ p_3^+ \cdot q_2^+ \end{pmatrix} = \begin{pmatrix} a^2-b^2-c^2+d^2 \\ 2(ac-bd) \end{pmatrix}$,
$\begin{pmatrix} \hat{p}_4^+ \cdot p_2^+ \\ \hat{p}_4^+ \cdot q_2^+ \end{pmatrix} = \begin{pmatrix} a(a-b)+3d(c-d) \\ a(c-d)+3d(b-a) \end{pmatrix}$. *The third point* $p_3^+ = (-a-b, -c-d)$ *has the vector* $q_3^+ = (c+d, -a-b) \perp p_3^+$, $\mathrm{SD}(C^+; p_3^+) = (a+b)^2 + (c+d)^2$ *and* $M(C^+; p_3^+)$ *consisting of the 3 unordered columns*

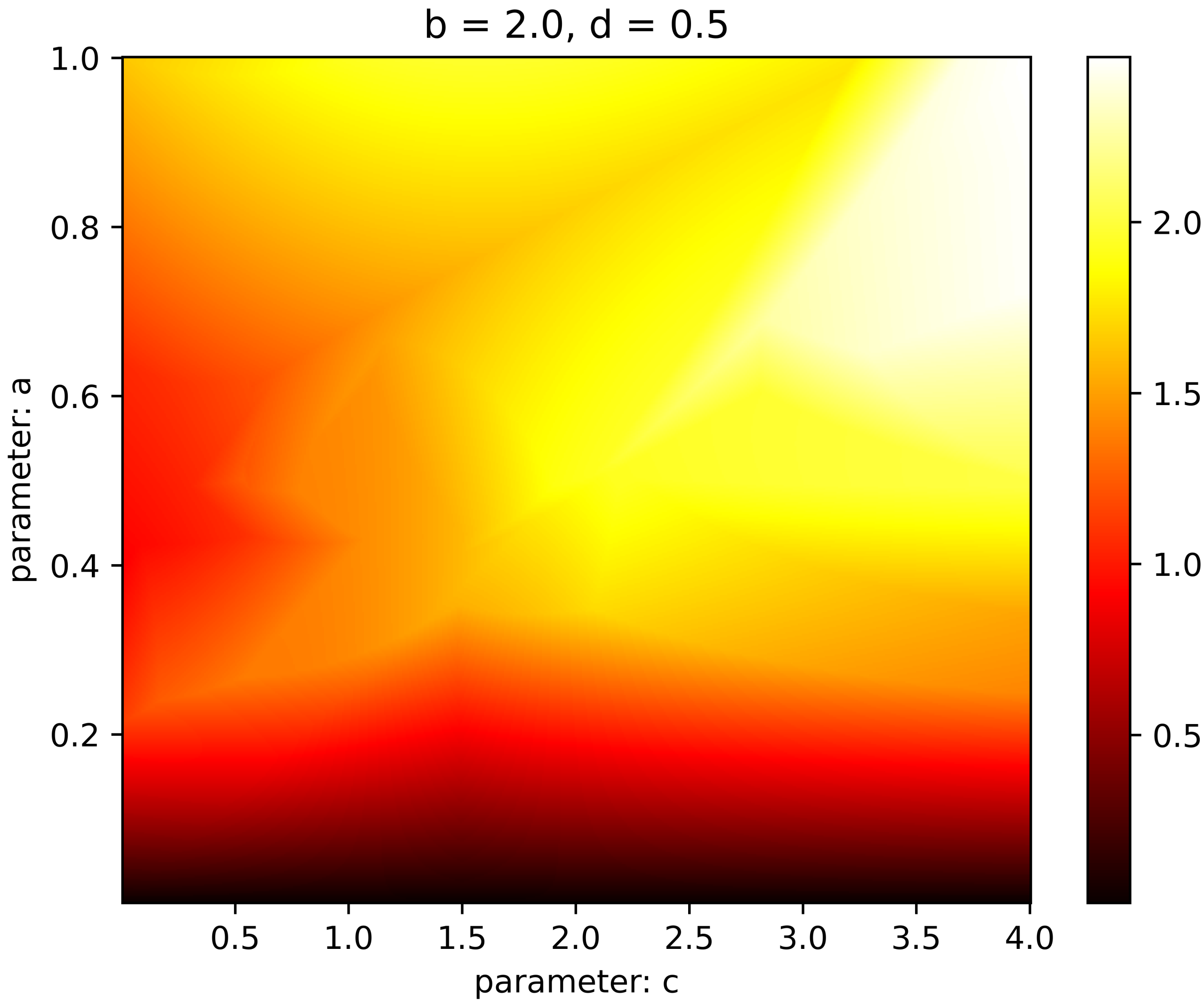

Figure 20. The Nested Bottleneck Metric NBM from Definition 4.4 for the 4-point clouds $C^{\pm} \subset \mathbb{R}^2$ with variable parameters a, c, see details in Example C.1.

$$\begin{pmatrix} p_1^+ \cdot p_3^+ \\ p_1^+ \cdot q_3^+ \end{pmatrix} = \begin{pmatrix} -3a(a+b) + d(c+d) \\ 3a(c+d) + d(a+b) \end{pmatrix},$$
$$\begin{pmatrix} p_2^+ \cdot p_3^+ \\ p_2^+ \cdot q_3^+ \end{pmatrix} = \begin{pmatrix} a^2 - b^2 - c^2 + d^2 \\ 2(bd - ac) \end{pmatrix},$$
$$\begin{pmatrix} \hat{p}_4^+ \cdot p_3^+ \\ \hat{p}_4^+ \cdot q_3^+ \end{pmatrix} = \begin{pmatrix} a(a+b) - 3d(c+d) \\ -a(c+d) - 3d(a+b) \end{pmatrix}.$$
The fourth point $\hat{p}_4^+ = (-a, 3d)$ *has the vector* $q_4^+ = (-3d, -a) \perp p_4^+$, $\mathrm{SD}(C^+; \hat{p}_4^+) = a^2 + 9d^2$, $M(C^+; \hat{p}_4^+)$ *has the columns*
$$\begin{pmatrix} p_1^+ \cdot \hat{p}_4^+ \\ p_1^+ \cdot q_4^+ \end{pmatrix} = \begin{pmatrix} -3(a^2 + d^2) \\ -8ad \end{pmatrix},$$
$$\begin{pmatrix} p_2^+ \cdot \hat{p}_4^+ \\ p_2^+ \cdot q_4^+ \end{pmatrix} = \begin{pmatrix} a(a-b) + 3d(c-d) \\ 3d(a-b) + a(d-c) \end{pmatrix},$$
$$\begin{pmatrix} p_3^+ \cdot \hat{p}_4^+ \\ p_3^+ \cdot q_4^+ \end{pmatrix} = \begin{pmatrix} a(a+b) - 3d(c+d) \\ 3d(a+b) + a(c+d) \end{pmatrix}.$$
The Nested Distributed Projection $\mathrm{NDP}(C^+)$ *consists of the four pairs (of a squared distance and* 2×3 *matrix) above.*

For C^-, after shifting the center $O(C^-) = (a, -d)$ to the origin $(0,0)$, the points of C^- become $p_1^- = (3a, d)$, $p_2^- = (b-a, d+c)$, $p_3^- = (-a-b, d-c)$, $\hat{p}_4^- = (-a, -3d)$. The first point p_1^- has the vector $q_1^- = (-d, 3a) \perp p_1^-$, $\mathrm{SD}(C^-; p_1^-) = 9a^2 + d^2$, $M(C^-; p_1^-)$ has the columns $\begin{pmatrix} p_2^- \cdot p_1^- \\ p_2^- \cdot q_1^- \end{pmatrix} = \begin{pmatrix} 3a(b-a) + d(d+c) \\ d(a-b) + 3a(d+c) \end{pmatrix}$, $\begin{pmatrix} p_3^- \cdot p_1^- \\ p_3^- \cdot q_1^- \end{pmatrix} = \begin{pmatrix} -3a(b+a) + d(d-c) \\ d(b+a) + 3a(d-c) \end{pmatrix}$, $\begin{pmatrix} \hat{p}_4^- \cdot p_1^- \\ \hat{p}_4^- \cdot q_1^- \end{pmatrix} = \begin{pmatrix} -3(a^2+d^2) \\ -8ad \end{pmatrix}$. *The second point $p_2^- = (b-a, d+c)$ has the vector $q_2^- = (-d-c, b-a) \perp p_2^-$, $\mathrm{SD}(C^-; p_2^-) = (a-b)^2 + (c+d)^2$, $M(C^-; p_2^-)$ of* $\begin{pmatrix} p_1^- \cdot p_2^- \\ p_1^- \cdot q_2^- \end{pmatrix} = \begin{pmatrix} 3a(b-a) + d(d+c) \\ -3a(c+d) + d(b-a) \end{pmatrix}$, $\begin{pmatrix} p_3^- \cdot p_2^- \\ p_3^- \cdot q_2^- \end{pmatrix} = \begin{pmatrix} a^2 - b^2 - c^2 + d^2 \\ 2(ac+bd) \end{pmatrix}$, $\begin{pmatrix} \hat{p}_4^- \cdot p_2^- \\ \hat{p}_4^- \cdot q_2^- \end{pmatrix} = \begin{pmatrix} a(a-b) - 3d(c+d) \\ a(c+d) + 3d(a-b) \end{pmatrix}$. *The third point $p_3^- = (-a-b, d-c)$ has $q_3^- = (c-d, -a-b) \perp p_3^-$, $\mathrm{SD}(C^-; p_3^-) = (a+b)^2 + (c-d)^2$, $M(C^-; p_3^-)$ of* $\begin{pmatrix} p_1^- \cdot p_3^- \\ p_1^- \cdot q_3^- \end{pmatrix} = \begin{pmatrix} -3a(a+b) + d(d-c) \\ 3a(c-d) - d(a+b) \end{pmatrix}$, $\begin{pmatrix} p_2^- \cdot p_3^- \\ p_2^- \cdot q_3^- \end{pmatrix} = \begin{pmatrix} a^2 - b^2 - c^2 + d^2 \\ -2(ac+bd) \end{pmatrix}$, $\begin{pmatrix} \hat{p}_4^- \cdot p_3^- \\ \hat{p}_4^- \cdot q_3^- \end{pmatrix} = \begin{pmatrix} a(a+b) + 3d(c-d) \\ a(d-c) + 3d(a+b) \end{pmatrix}$. *The fourth point $\hat{p}_4^- = (-a, -3d)$ has $q_4^- = (3d, -a) \perp \hat{p}_4^-$, $\mathrm{SD}(C^-; \hat{p}_4^-) = a^2 + 9d^2$, $M(C^-; \hat{p}_4^-)$ consisting of* $\begin{pmatrix} p_1^- \cdot \hat{p}_4^- \\ p_1^- \cdot q_4^- \end{pmatrix} = \begin{pmatrix} -3(a^2+d^2) \\ 8ad \end{pmatrix}$, $\begin{pmatrix} p_2^- \cdot \hat{p}_4^- \\ p_2^- \cdot q_4^- \end{pmatrix} = \begin{pmatrix} a(a-b) - 3d(d+c) \\ 3d(b-a) - a(d+c) \end{pmatrix}$, $\begin{pmatrix} p_3^- \cdot \hat{p}_4^- \\ p_3^- \cdot q_4^- \end{pmatrix} = \begin{pmatrix} a(a+b) + 3d(c-d) \\ -3d(a+b) + a(c-d) \end{pmatrix}$. *The Nested Distributed Projection $\mathrm{NDP}(C^-)$ consists of the four pairs (of a squared distance and 2×3 matrix) above.*

Shorter Example C.1 justified that $C^+ \not\cong C^-$ unless at least of the parameters a, b, c, d is 0. If $a = 0$ or $d = 0$, then $C^+ \cong C^-$ are isometric. In the remaining cases $b = 0$ and $c = 0$, the clouds $C^\pm$ are mirror images, which can be distinguished by matrices M above, not by any distances.

***Case** $b = 0$. We write down the above matrices $M(C^+; p_i^+)$ with unordered columns after substituting $b = 0$.*

$$\begin{pmatrix} -3a^2 + d(d-c) & -3a^2 + d(d+c) & -3(a^2+d^2) \\ a(3c-4d) & -a(3c+4d) & 8ad \end{pmatrix}$$

$$\begin{pmatrix} -3a^2 + d(d-c) & a^2 - c^2 + d^2 & a^2 + 3d(c-d) \\ a(4d-3c) & 2ac & a(c-4d) \end{pmatrix}$$

$$\begin{pmatrix} -3a^2 + d(d+c) & a^2 - c^2 + d^2 & a^2 - 3d(c+d) \\ a(3c+4d) & -2ac & -a(c+4d) \end{pmatrix}$$

$$\begin{pmatrix} -3(a^2+d^2) & a^2 + 3d(c-d) & a^2 - 3d(c+d) \\ -8ad & a(4d-c) & a(c+4d) \end{pmatrix}$$

The mirror image C^- has the following matrices:

$$\begin{pmatrix} -3a^2 + d(d+c) & -3a^2 + d(d-c) & -3(a^2+d^2) \\ a(3c+4d) & a(4d-3c) & -8ad \end{pmatrix}$$

$$\begin{pmatrix} -3a^2+d(d+c) & a^2-c^2+d^2 & a^2-3d(c+d)) \\ -a(3c+4d) & 2ac & a(c+4d) \end{pmatrix}$$

$$\begin{pmatrix} -3a^2+d(d-c) & a^2-c^2+d^2 & a^2+3d(c-d) \\ a(3c-4d) & -2ac & a(4d-c) \end{pmatrix}$$

$$\begin{pmatrix} -3(a^2+d^2) & a^2-3d(c+d) & a^2+3d(c-d) \\ 8ad & -a(c+4d) & a(c-4d) \end{pmatrix}$$

By Lemma B.3(b), the reflection $C^+ \to C^-$ changes the sign of the last row in the matrix M from any point-based representation PR*. Indeed, changing the sign of the last row in each matrix M from $\mathrm{NDP}(C^+)$ makes this matrix identical to one of the matrices from $\mathrm{NDP}(C^-)$, up to a permutation of columns as always. However, with all signs kept, the above unordered collections of four matrices are different unless all elements in the last row vanish, which happens only for a=0, when $C^+ = C_-$ are identical.*

***Case** $c = 0$ is symmetric to the case $c = 0$ under the reflection $(x, y) \mapsto (y, x)$, which swaps $b \leftrightarrow c$ and $a \leftrightarrow d$.*

We have considered only non-negative values of a, b, c, d because all other cases are obtained by symmetries. For example, the reflection $y \mapsto -y$ maps the cloud $C^+(a, b, c, d)$ to $C^-(a, -b, c, d) = C^-(a, b, -c, d)$.

Example C.2 importantly demonstrates that the invariant NDP is simple enough for manual computations.

A numerical experiment can only illustrate but not prove the conclusion of Example C.2 that all (infinitely many) non-rigidly equivalent clouds $C^{\pm}$ are distinguished by NDP.

D. Generalization of section 4 and all proofs in dimensions $n \geq 2$

This appendix extends the metrics to dimensions $n \geq 2$ and proves all metric results from section 4 in full generality.

The point-based representation in Definition B.2 included the matrix $\mathrm{SD}(p_1, \dots, p_{n-1})$ of squared distances, which can be rewritten as a vector row-by-row.

Below we can take any norm on matrices and choose the simplest max norm below for consistency with the bottleneck distance and for Lipschitz constant 2 in Theorem E.5.

Definition D.1 (max norm and metric on matrices)**.** *The* max norm $||D||_\infty = \max\limits_{i,j} |D_{ij}|$ *of a matrix is the maximum absolute value of its elements D_{ij}. The* max metric *between matrices M, M' of the same size is $d_\infty = ||M - M'||_\infty$.*

Definition D.2 will extend Definition 4.2 to dimensions $n \geq 2$. Below the notation SD/R means that all elements of a matrix SD are divided by R. The radius of a base sequence $p\{n-1\} = (p_1, \dots, p_{n-1}) \subset A$ is defined as $R(p\{n-1\}) = \max\limits_{i=1,\dots,n-1} |p_i|$ in the same way as $R(A)$ of a full cloud A. The notation M/R means that all elements in the first $n-1$ rows of a matrix M are divided by R, and by R^{n-1} in the n-th row, because $p_n^\perp$ in Lemma B.1 is a polynomial of degree $n-1$. Then PRM and further metrics have units of original points. One more division by R makes all metrics invariant under scaling.

Definition D.2 (Point-Based Representation Metric)**.** *Let clouds $A, B \subset \mathbb{R}^n$ of m unordered points have base sequences $p\{n-1\} = (p_1, \dots, p_{n-1})$, $q\{n-1\} = (q_1, \dots, q_{n-1})$ of ordered points, from Definition B.2. The* Point-Based Representation Metric *between the* PR*s above is*

$$\mathrm{PRM} = \max\{\, |R(p\{n-1\}) - R(q\{n-1\})|,\, w_D,\, |R(A) - R(B)|,\, w_M \,\}, \text{ where}$$

$$w_D = d_\infty\left(\frac{\mathrm{SD}(p\{n-1\})}{R(p\{n-1\})}, \frac{\mathrm{SD}(q\{n-1\})}{R(q\{n-1\})}\right), \text{ and } w_M = \mathrm{BD}\left(\frac{M(A; p\{n-1\})}{R(A)}, \frac{M(B; q\{n-1\})}{R(B)}\right).$$

Lemma D.3 (axioms for PRM)**.** PRM *in Definition D.2 satisfies all metric axioms from Problem (1.1b) on any point-based representations from Definition B.8.*

Proof of Lemma D.3. The first axiom means that $\mathrm{PRM}(\mathrm{PR}(A;p\{n-1\}),\mathrm{PR}(B;q\{n-1\}))=0$ if and only if these PRs are identical. The part *if*: by Lemma B.5(c), equal PRs guarantee that the clouds A,B are rigidly equivalent, so $R(p\{n-1\})=R(q\{n-1\})$, $R(A)=R(B)$, $\mathrm{SD}(p\{n-1\})=\mathrm{SD}(q\{n-1\})$, and $M(A;p\{n-1\})=M(B;q\{n-1\})$, so $\mathrm{PRM}=0$.

The part *only if*: by Definition D.2 the equality $\mathrm{PRM}=0$ means that $R(A)=R(B)$ and $w_D=0=w_M$. The coincidence axioms for the max metric and bottleneck distance together with $R(p\{n-1\})=R(q\{n-1\})$ and $R(A)=R(B)$ imply that $\mathrm{SD}(p\{n-1\})=\mathrm{SD}(q\{n-1\})$ and $M(A;p\{n-1\})=M(B;q\{n-1\})$. Then the point-based representations become identical: $\mathrm{PR}(A;p\{n-1\})=\mathrm{PR}(B;q\{n-1\})$.

The symmetry axiom for PRM follows from the symmetry axiom for the bottleneck distance and max metric d_∞. Since each of the distances $|R(A)-R(B)|$, w_D, w_M satisfies the triangle inequality, then so does their maximum, see metric transforms in section 4.1 of (Deza & Deza, 2009). □

Definition D.4 extends Definition 4.4 to all dimensions $n>2$.

Definition D.4 (NBM : Nested Bottleneck Metric)**.** *Let $A,B\subset\mathbb{R}^n$ be any clouds of m unordered points. For any ordered points $p_1\dots,p_{n-2}\in A$ and $q_1\dots,q_{n-2}\in B$, the complete bipartite graph $\Gamma(A;p_1,\dots,p_{n-2};B;q_1,\dots,q_{n-2})$ has $m-n+2$ white vertices and $m-n+2$ black vertices representing $\mathrm{PR}(A;p_1,\dots,p_{n-1})$ and $\mathrm{PR}(B;q_1,\dots,q_{n-1})$ for all $m-n+1$ variable points $p_{n-1}\in A\setminus\{p_1,\dots,p_{n-2}\}$ and $q_{n-1}\in B-\{q_1,\dots,q_{n-2}\}$, respectively.*

Set the weight $w(e)$ *of an edge e joining the vertices represented by $\mathrm{PR}(A;p_1,\dots,p_{n-1})$ and $\mathrm{PR}(B;q_1,\dots,q_{n-1})$ as PRM between these PRs, see Definition D.2. Then Definition 4.3 gives us the bottleneck matching distance $\mathrm{BMD}(\Gamma(A;p_1,\dots,p_{n-2};B;q_1,\dots,q_{n-2}))$. We continue dropping points iteratively. For any $1\le k\le n-2$ and ordered points $p_1\dots,p_{k-1}\in A$ and $q_1\dots,q_{k-1}\in B$, the complete bipartite graph $\Gamma(A;p_1,\dots,p_{k-1};B;q_1,\dots,q_{k-1})$ has $m-k+1$ white vertices and $m-k+1$ black vertices representing $\mathrm{NDP}(A;p_1,\dots,p_k)$ and $\mathrm{NDP}(B;q_1,\dots,q_k)$ for all $m-k+1$ variable points $p_k\in A\setminus\{p_1,\dots,p_{k-1}\}$ and $q_k\in B-\{q_1,\dots,q_{k-1}\}$, respectively.*

Set the weight $w(e)$ *of an edge e joining the vertices represented by $\mathrm{NDP}(A;p_1,\dots,p_k)$ and $\mathrm{NDP}(B;q_1,\dots,q_k)$ as $\mathrm{BMD}(\Gamma(A;p_1,\dots,p_k;B;q_1,\dots,q_k))$ obtained above. Then Definition 4.3 gives us the bottleneck matching distance $\mathrm{BMD}(\Gamma(A;p_1,\dots,p_{k-1};B;q_1,\dots,q_{k-1}))$. Finally, for $k=1$, we get the* Nested Bottleneck Metric $\mathrm{NBM}(A,B)=\mathrm{BMD}(\Gamma(A,B))$.

Lemma D.5 (metric axioms for the bottleneck matching distance BMD)**.** *Let S,Q be any unordered distributions of the same number of objects with a base metric d. Define the complete bipartite graph $\Gamma(S,Q)$ whose every edge e joining objects $R_S\in S$ and $R_Q\in Q$ has the weight $w(e)=d(R_S,R_Q)$. Then the bottleneck matching distance $\mathrm{BMD}(\Gamma(S,Q))$ from Definition 4.3 satisfies all metric axioms on such unordered distributions.*

Proof of Lemma D.5. The coincidence axiom means that $\mathrm{NBM}(S,Q)=0$ if and only if the weighted distributions S,Q are equal in the sense that there is a bijection $g:S\to Q$ so that $d(g(R),R)=0$ for any $R\in S$.

Indeed, if the weighted distributions S,Q can be matched by a bijection, we get a vertex matching E of $\Gamma(S,Q)$ whose all edges have weights $w(e)=0$. Definition 4.3 implies that $\mathrm{BMD}(\Gamma(S,Q))=0$ as required.

Conversely, if $\mathrm{BMD}(\Gamma(S,Q))=0$, there is a vertex matching E in $\Gamma(S,Q)$ with all $w(e)=0$. This matching E defines a required bijection $S\to Q$. The symmetry $\mathrm{BMD}(\Gamma(S,Q))=\mathrm{BMD}(\Gamma(Q,S))$ follows from Definition 4.3 and the symmetry of the base metric d.

To prove the triangle inequality

$$\mathrm{BMD}(\Gamma(S,Q))+\mathrm{BMD}(\Gamma(Q,T))\ge\mathrm{BMD}(\Gamma(S,T)),$$

let E_{SQ},E_{QT} be optimal vertex matchings in the graphs $\Gamma(S,Q),\Gamma(Q,T)$, respectively, such that

$$\mathrm{BMD}(\Gamma(S,Q))=W(E_{SQ}),\mathrm{BMD}(\Gamma(Q,T))=W(E_{QT}),$$

see Definition 4.3. The composition $E_{SQ}\circ E_{QT}$ is a vertex matching in $\Gamma(S,T)$, so $W(E_{SQ}\circ E_{QT})\ge\mathrm{BMD}(\Gamma(S,T))$. It suffices to prove that

$$W(E_{SQ})+W(E_{QT})\ge W(E_{SQ}\circ E_{QT}).$$

Let e_{ST} be an edge with a largest weight from $E_{SQ} \circ E_{QT}$, so $W(E_{SQ} \circ E_{QT}) = w(e_{ST})$. The edge e_{ST} can be considered the union of edges $e_{SQ} \in E_{SQ}$, $e_{QT} \in E_{QT}$.

By the triangle inequality for the base metric d,

$$w(e_{SQ}) + w(e_{QT}) \geq w(e_{ST}) = W(E_{SQ} \circ E_{QT})$$

implies that

$$W(E_{SQ}) + W(E_{QT}) \geq W(E_{SQ} \circ E_{QT})$$

because both terms on the left-hand side are maximized for all edges (not only e_{SQ}, e_{QT}) from E_{SQ}, E_{QT}. □

Lemma D.6 (metric axioms for NBM between NDPs)**.** *The Nested Bottleneck Metric* NBM *from Definition D.4 satisfies all metric axioms on Nested Distributed Projections.*

Proof of Lemma D.6. Induction on $k = n-2, \dots, 1$. The inductive base $k = n-2$ follows from the metric axioms in Lemma D.3 for PRM in Definition D.2. The inductive step from $1 < k < n-2$ to $k-1$ follows from Lemma D.5 and the metric axioms in the inductive hypothesis for k. □

E. Generalization of section 5 and all proofs

This appendix proves Theorems E.5, E.8, and E.9 extending Lemmas 5.1, 5.2, and 5.3, respectively to dimensions $n \geq 2$ by using auxiliary Lemmas E.1, E.2, E.4, and Proposition E.3.

Lemma E.1 (orthogonal vector length)**.** *For any sequence $p_1, \dots, p_{n-1} \in \mathbb{R}^n$, set $R = \max\limits_{i=1,\dots,n-1} |p_i|$. Then the orthogonal vector $p_n^\perp \perp p_1, \dots, p_{n-1}$ from Lemma B.1 has a length satisfying $|p_2^\perp| = R$, $|p_3^\perp| \leq R^2$, and $|p_n^\perp| \leq \sqrt{n} R^{n-1}$ for any $n > 3$.*

Proof of Lemma E.1. For $n = 2$, the explicit formula $p_2^\perp = (-y, x)$ for $p_1 = (x, y)$ gives the exact equality $|p_2^\perp| = |p_1| = R$. For $n = 3$, $p_3^\perp$ equals the vector product $p_1 \times p_2$ whose length is $|p_3^\perp| \leq |p_1| \cdot |p_2| \leq R^2$. For > 3, the expansion of the $n \times n$ determinant $p_n^\perp = \begin{vmatrix} | & \dots & | & e_1 \\ p_1 & \dots & p_{n-1} & \vdots \\ | & \dots & | & e_n \end{vmatrix}$ along the last column gives $p_n^\perp = \sum\limits_{i=1}^{n} (-1)^{n+i} \det(i) e_i$, where $\det(i)$ is the $(n-1) \times (n-1)$ determinant obtained from the $n-1$ vector columns $p_1, \dots, p_{n-1}$ by removing the row of all i-th coordinates. Any determinant on vectors $v_1, \dots, v_{n-1} \in \mathbb{R}^{n-1}$ equals the signed volume of the parallelepiped on $v_1, \dots, v_{n-1}$, which has the upper bound $|v_1| \cdots |v_{n-1}|$.

Since each vector v_i is obtained from p_i by removing one coordinate, we get $|v_i| \leq |p_i|$. So each coordinate of $p_n^\perp$ in the orthonormal basis $e_1, \dots, e_n$ has the upper bound $|p_1| \cdots |p_{n-1}| \leq R^{n-1}$. Then the Euclidean length has the upper bound $|p_n^\perp| \leq \sqrt{n (R^{n-1})^2} = \sqrt{n} R^{n-1}$. □

Lemma E.2 (vector perturbations)**.** *Let points $q_1, \dots, q_{n-1}$ be ε-perturbations of $p_1, \dots, p_{n-1} \in \mathbb{R}^n$ so that $|p_i - q_i| \leq \varepsilon$ for any $i = 1, \dots, n-1$. Set $R = \max\limits_{i=1,\dots,n-1} \{|p_i|, |q_i|\}$. The orthogonal vectors $p_n^\perp \perp p_1, \dots, p_{n-1}$ and $q_n^\perp \perp q_1, \dots, q_{n-1}$ from Lemma B.1 satisfy $|p_2^\perp - q_2^\perp| \leq \varepsilon$ for $n = 2$, $|p_3^\perp - q_3^\perp| \leq \varepsilon 2\sqrt{6} R$ for $n = 3$, and $|p_n^\perp - q_n^\perp| \leq \varepsilon n(n-1) R^{n-2}$ for any $n > 3$.*

Proof of Lemma E.2. If $n = 2$, then $p_2^\perp = (-y, x)$ for $p_1 = (x, y)$, so $|p_2^\perp - q_2^\perp| = |p_1 - q_1| \leq \varepsilon$.

Let $x_i(v_j)$ be the i-th coordinate of a variable vector $v_j \in \mathbb{R}^n$ moving from p_j to its ε-perturbation q_j for $i, j = 1, \dots, n$ in the given orthonormal basis $e_1, \dots, e_n$, where we set $p_n = p_n^\perp$ and $q_n = q_n^\perp$ for brevity. For each $k = 1, \dots, n$, the coordinate $x_k(v_n)$ is the scalar function $f_k(v_1, \dots, v_{n-1})$ of the $(n-1)^2$ variables $x_i(v_j)$ for $i, j = 1, \dots, n-1$.

The upper bound for $|p_n - q_n|$ will follow from the Mean Value Theorem 5.10 from (Rudin et al., 1976) for the functions $f_1, \dots, f_n$ because the coordinates of the vector $q_n^\perp$ are $f_k(q_1, \dots, q_{n-1})$ evaluated at close (coordinates of the) vectors $q_1, \dots, q_{n-1}$ so that $|p_j - q_j| \leq \varepsilon$ for $i, j = 1, \dots, n-1$.

First we estimate the gradient ∇f_k of f_k at any intermediate point in the line segment between $(p_1,\dots,p_{n-1})$ and $(q_1,\dots,q_{n-1})$ with respect to the $(n-1)^2$ variables $x_i(v_j)$ for $i,j=1,\dots,n-1$. For $k=i$, the k-th coordinate of $v_n = \begin{vmatrix} | & \dots & | & e_1 \\ v_1 & \dots & v_{n-1} & \vdots \\ | & \dots & | & e_n \end{vmatrix}$ is $(-1)^{n+k}\det(k)$, where $\det(k)$ is the $(n-1)\times(n-1)$ determinant obtained from the $n-1$ vector columns $v_1,\dots,v_{n-1}$ by removing the row of all k-th coordinates. Then $\dfrac{\partial f_k}{\partial x_i(v_j)} = (-1)^{n+k}\dfrac{\partial \det(k)}{\partial x_i(v_j)}$, which equals 0 for $k=i$ because f_k is independent of the coordinate $x_k(v_j)$ for $j=1,\dots,n-1$.

After expanding the determinant $\det(k)$ along the i-th row, the only terms containing the factor $x_i(v_j)$ form the smaller $(n-2)\times(n-2)$ determinant $\det(k,i)$ obtained from the $n-2$ vector columns $v_1,\dots,v_{j-1},v_{j+1},\dots,v_{n-1}$ after removing the rows of all k-th and i-th coordinates.

Then $|v_j| \le R = \max\limits_{i=1,\dots,n-1}\{|p_i|,|q_i|\}$ for any points $(v_1,\dots,v_{n-1})$ in the line segment between $(p_1,\dots,p_{n-1})$ and $(q_1,\dots,q_{n-1})$. The $(n-2)\times(n-2)$ determinant $\det(k,i)$ equals the signed volume on $n-2$ vectors of maximum length R and hence has the upper bound R^{n-2}, so $\left|\dfrac{\partial f_k}{\partial x_i(v_j)}\right| = |\det(k,i)| \le R^{n-2}$. The gradient ∇f_k is the vector of $(n-1)^2$ partial derivatives and can be considered a vector $(\nabla_1 f_k,\dots,\nabla_{n-1}f_k)$, where $\nabla_j f_k = \left(\dfrac{\partial f_k}{x_1(v_j)},\dots,\dfrac{\partial f_k}{x_{n-1}(v_j)}\right)$ has

$$|\nabla_j f_k| \le \sqrt{n-1}\max_{i=1,\dots,n-1}\left|\frac{\partial f_k}{\partial x_i(v_j)}\right| \le \sqrt{n-1}R^{n-2}.$$

We consider the k-th coordinate f_k of v_n as a function depending on one parameter $t\in[0,1]$ when the point $(v_1,\dots,v_{n-1})$ moves along the line segment from $(p_1,\dots,p_{n-1})$ to $(q_1,\dots,q_{n-1})$. Then Theorem 5.10 from (Rudin et al., 1976) implies for some intermediate point $(v_1,\dots,v_{n-1})$ that

$$|f_k(p_1,\dots,p_{n-1}) - f_k(q_1,\dots,q_{n-1})| = |\nabla f_k(v_1,\dots,v_{n-1})\cdot(p_1-q_1,\dots,p_{n-1}-q_{n-1})| =$$

$$= \left|\sum_{i,j=1}^{n-1}\frac{\partial f_k}{\partial x_i(v_j)}\cdot\big(x_i(p_j)-x_i(q_j)\big)\right| = \left|\sum_{j=1}^{n-1}\nabla_j f_k\cdot(p_j-q_j)\right| \le \sum_{j=1}^{n-1}|\nabla_j f_k|\cdot|p_j-q_j| \le$$

$$\le \varepsilon(n-1)\max_{j=1,\dots,n-1}|\nabla_j f_k| \le \varepsilon(n-1)\sqrt{n-1}R^{n-2}.$$

Since $e_1,\dots,e_n$ form an orthonormal basis, we get

$$|p_n^\perp - q_n^\perp| = \sqrt{\sum_{k=1}^{n}|f_k(p_1,\dots,p_{n-1}) - f_k(q_1,\dots,q_{n-1})|^2}$$

$$\le \sqrt{n}\max_{k=1,\dots,n}|f_k(p_1,\dots,p_{n-1}) - f_k(q_1,\dots,q_{n-1})| \le \sqrt{n}\varepsilon(n-1)\sqrt{n-1}R^{n-2} \le \varepsilon n(n-1)R^{n-2}$$

for any $n\ge 3$. If $n=3$, the final upper bound can be improved to $\varepsilon 2\sqrt{6}R$. □

Proposition E.3 (Lipschitz continuity of PR under perturbations of a cloud)**.** *Let $B\subset\mathbb{R}^n$ and a base sequence $q\{n-1\}\subset B$ be obtained from a cloud $A\subset\mathbb{R}^n$ and a base sequence $p\{n-1\}\subset A$, respectively, by perturbing every point in its Euclidean ε-neighborhood. Then*

(a) $|O(A)-O(B)|\le\varepsilon$, $|R(p\{n-1\}-R(q\{n-1\})|\le 2\varepsilon$, *and* $|R(A)-R(B)|\le 2\varepsilon$*;*

(b) $\mathrm{PRM}\big(\mathrm{PR}(A;p\{n-1\}),\mathrm{PR}(B;q\{n-1\})\big)\le\lambda_n\varepsilon$ *for* $\lambda_2=6$, $\lambda_3=16$, $\lambda_n=3n^2$, $n>3$.

Proof of Proposition E.3. **(a)** Let $p_1\dots,p_m$ be all points of A so that the first $n-1$ points $p_1,\dots,p_{n-1}$ form the base sequence $p\{n-1\}$. Let $q_i\in B$ be an ε-perturbation of p_i, so $q_1\dots,q_m$ are all points of B and the first $n-1$ points

$q_1, \dots, q_{n-1}$ form the base sequence $q\{n-1\}$. The radius of A is $R(A) = \max\limits_{p \in A} |p - O(A)|$, where $O(A) = \dfrac{1}{m} \sum\limits_{p \in A} p$ is the center of mass. Then

$$|O(A) - O(B)| = \frac{1}{m} \left| \sum_{i=1}^{m} p_i - \sum_{i=1}^{m} q_i \right| \leq \frac{1}{m} \sum_{i=1}^{m} |p_i - q_i| \leq \varepsilon.$$

If the radius $R(A)$ is attained at a point $p_i \in A$, then $R(A) = |p_i - O(A)| \leq$

$$\leq |p_i - q_i| + |q_i - O(B)| + |O(B) - O(A)| \leq \varepsilon + \max_{i=1,\dots,m} |q_i - O(B)| + \varepsilon = 2\varepsilon + R(B).$$

Swapping the clouds A, B gives the opposite inequality $R(B) \leq 2\varepsilon + R(A)$, so $|R(A) - R(B)| \leq 2\varepsilon$. The radii of the base sequences also differ by at most 2ε, i.e. $|R(p\{n-1\}) - R(q\{n-1\})| \leq 2\varepsilon$.

(b) All corresponding points of the given clouds A, B are ε-close so that $|p_i - q_i| \leq \varepsilon$ for all $i = 1, \dots, m$. Any distance $|p_i - p_j|$ changes by at most 2ε under perturbation, because

$$|p_i - p_j| \leq |p_i - q_i| + |q_i - q_j| + |q_j - p_j| \leq |q_i - q_j| + 2\varepsilon,$$
$$|q_i - q_j| \leq |q_i - p_i| + |p_i - p_j| + |p_j - q_j| \leq |p_i - p_j| + 2\varepsilon.$$

Hence $\big|\, |p_i - p_j| \; - \; |q_i - q_j| \,\big| \leq 2\varepsilon$ for all $i, j = 1, \dots, m$.

To estimate the max metric d_∞ in (D.2), we rewrite the difference between the corresponding elements in the matrices SD/R of squared distances normalized by the radii in the notations $r(A) = R(p\{n-1\})$ and $r(B) = R(q\{n-1\})$. Without loss of generality, assume that $r(A) \geq r(B)$.

$$\text{Then } \left| \frac{|p_i - p_j|^2}{r(A)} - \frac{|q_i - q_j|^2}{r(B)} \right| \leq \frac{|\,|p_i - p_j|^2 - |q_i - q_j|^2\,|}{r(A)} + |q_i - q_j|^2 \frac{|r(B) - r(A)|}{r(A) r(B)}$$

for $i, j = 0, \dots, n-1$, where $p_0 = O(A)$ and $q_0 = O(B)$ are centers of mass. In the first term above, we estimate the difference of squares by factorizing:

$$|\,|p_i - p_j|^2 - |q_i - q_j|^2\,| = |\,|p_i - p_j| - |q_i - q_j|\,| \cdot (|p_i - p_j| + |q_i - q_j|) \leq 2\varepsilon(2r(A) + 2r(B)).$$

Using $r(A) \geq r(B)$, the bounds $\dfrac{|\,|p_i - p_j|^2 - |q_i - q_j|^2\,|}{r(A)} \leq 4\varepsilon \dfrac{r(A) + r(B)}{r(A)} \leq 8\varepsilon$, $|q_i - q_j|^2 \dfrac{|r(B) - r(A)|}{r(A) r(B)} \leq \dfrac{(2r(B))^2 \cdot 2\varepsilon}{r(A) r(B)} \leq 8\varepsilon$ give $d_\infty \left(\dfrac{\mathrm{SD}(p\{n-1\})}{r(A)}, \dfrac{\mathrm{SD}(q\{n-1\})}{r(B)} \right) \leq 16\varepsilon$.

To estimate the bottleneck distance BD between the matrices M/R in (D.2), which involve scalar products, we shift both clouds A, B so that their centers $O(A)$ and $O(B)$ coincide with the origin $0 \in \mathbb{R}^n$. We keep the same notation p_i, q_i for all points for simplicity. Since $|O(A) - O(B)| \leq \varepsilon$ by part (a), the relative shift by a vector of a maximum length ε guarantees all corresponding points are now 2ε-close, i.e. $|p_i - q_i| \leq 2\varepsilon$. Below we estimate the difference between scalsr products involving any 2ε-close points $p \in A \setminus p\{n-1\}$ and $q \in B - q\{n-1\}$ for $i = 1, \dots, n-1$ (indexing points from the base sequences) and $i = n$ for the orthogonal vectors $p_n = p_n^\perp$, $q_n = q_n^\perp$.

Case $i = 1, \dots, n-1$. The bottleneck distance BD has the upper bound obtained from estimating the differences below in the M/R matrices for any point $p \in A \setminus p\{n-1\}$ matched with its 2ε-perturbation $q \in B - q\{n-1\}$. Without loss of generality, assume that $R(A) \geq R(B)$. Then

$$\left| \frac{p \cdot p_i}{R(A)} - \frac{q \cdot q_i}{R(B)} \right| \leq \frac{|p \cdot p_i - q \cdot q_i|}{R(A)} + |q \cdot q_i| \frac{|R(B) - R(A)|}{R(A) R(B)}.$$

Due to $|q \cdot q_i| \leq |q| \cdot |q_i| \leq R^2(B)$, the second term above has the upper bound $\dfrac{R^2(B) \cdot 2\varepsilon}{R(A) R(B)} \leq 2\varepsilon$. Estimate the difference of products in the first term above:

$$|p \cdot p_i - q \cdot q_i| \leq |(p - q) \cdot p_i + q \cdot (p_i - q_i)| \leq |p - q| \cdot |p_i| + |q| \cdot |p_i - q_i| \leq 2\varepsilon(R(A) + R(B)).$$

Then $\dfrac{|p\cdot p_i - q\cdot q_i|}{R(A)} \leq 2\varepsilon\dfrac{R(A)+R(B)}{R(A)} = 4\varepsilon$. For every $i=1,\dots,n-1$, we get $\left|\dfrac{p\cdot p_i}{R(A)} - \dfrac{q\cdot q_i}{R(B)}\right| \leq 6\varepsilon$ for every point $p\in A\setminus p\{n-1\}$ and its 2ε-perturbation $q\in B-q\{n-1\}$.

Case $i=n$ is for the n-th row of the matrices M/R in (D.2), where the scalar products with the orthogonal vectors $p_n^{\perp}, q_n^{\perp}$ from Lemma B.1 are divided by R^{n-1} instead of R.

Subcase $i=n=2$ coincides with the case $i<n$ above because $R^{n-1}=R$. Combining the upper bounds above, we get $\mathrm{BD}\left(\dfrac{M(A;p\{n-1\})}{R(A)}, \dfrac{M(B;q\{n-1\})}{R(B)}\right) \leq 6\varepsilon$ By Definition 4.2, the Point-based Representation Metric PRM equals the maximum of the bounds $d_\infty = |R(p_1)-R(q_1)| = \big||p_1|-|q_1|\big| \leq 2\varepsilon$, $|R(A)-R(B)|\leq 2\varepsilon$, and BD above, so $\mathrm{PRM}\big(\mathrm{PR}(A;p_1),\mathrm{PR}(B;q_1)\big) \leq 6\varepsilon$, which finishes the proof of part (b) for $n=2$.

Subcase $i=n=3$. Without loss of generality, we can assume that $R(A)\geq R(B)$. The upper bounds of Lemmas E.1 and E.2 imply that

$$|p_3^{\perp}| \leq R^2(A),\quad |q_3^{\perp}| \leq R^2(B),\quad |p_3^{\perp}-q_3^{\perp}| \leq 2\varepsilon\cdot 2\sqrt{6}R(A).$$

We start estimating similarly to the case $i<n$ above:

$$|p\cdot p_3^{\perp} - q\cdot q_3^{\perp}| \leq |(p-q)\cdot p_3^{\perp} + q\cdot(p_3^{\perp}-q_3^{\perp})| \leq |p-q|\cdot|p_3^{\perp}| + |q|\cdot|p_3^{\perp}-q_3^{\perp}| \leq$$
$$2\varepsilon R^2(A) + R(B)\cdot 2\varepsilon\cdot 2\sqrt{6}R(A) = 2\varepsilon R(A)(R(A)+4\sqrt{6}R(B)).$$

$$\text{Then }\left|\frac{p\cdot p_3^{\perp}}{R^2(A)} - \frac{q\cdot q_3^{\perp}}{R^2(B)}\right| \leq \frac{|p\cdot p_3^{\perp} - q\cdot q_3^{\perp}|}{R^2(A)} + |q\cdot q_3^{\perp}|\frac{|R^2(B)-R^2(A)|}{R^2(A)R^2(B)} \leq$$

$$\leq 2\varepsilon\frac{R(A)+2\sqrt{6}R(B)}{R(A)} + |q|\cdot|q_3^{\perp}|\frac{R^2(A)-R^2(B)}{R^2(A)R^2(B)} \leq 2\varepsilon(1+2\sqrt{6}) + R^3(B)\left(\frac{1}{R^2(B)} - \frac{1}{R^2(A)}\right).$$

We use $R(A)\leq R(B)+2\varepsilon$ to bound last term:

$$R(B)\left(1-\frac{R^2(B)}{R^2(A)}\right) \leq R(B)\left(1-\frac{R^2(B)}{(R(B)+2\varepsilon)^2}\right) \leq \frac{R(B)}{(R(B)+2\varepsilon)^2}4\varepsilon(R(B)+\varepsilon) \leq 4\varepsilon.$$

Then $\left|\dfrac{p\cdot p_3^{\perp}}{R^2(A)} - \dfrac{q\cdot q_3^{\perp}}{R^2(B)}\right| \leq 2\varepsilon(1+2\sqrt{6})+4\varepsilon < 16\varepsilon$. By Definition D.2, the Point-based Representation Metric PRM equals the maximum of

$$d_\infty = |R(p\{2\}) - R(q\{2\})| \leq 2\varepsilon,\quad |R(A)-R(B)|\leq 2\varepsilon,\quad d_\infty \leq 16\varepsilon,\quad \mathrm{BD} < 16\varepsilon,$$

so $\mathrm{PRM}\big(\mathrm{PR}(A;p\{2\}),\mathrm{PR}(B;q\{2\})\big) \leq 16\varepsilon$ which finishes the proof of part (b) for $n=3$.

Final subcase $i=n>3$. Assuming again that $R(A)\geq R(B)$, Lemmas E.1 and E.2 give

$$|p_n^{\perp}| \leq \sqrt{n}R^{n-1}(A),\quad |q_n^{\perp}| \leq \sqrt{n}R^{n-1}(B),\quad |p_n^{\perp}-q_n^{\perp}| \leq 2\varepsilon n(n-1)R^{n-2}(A) \text{ for any } n>3.$$

We start estimating similarly to the case $i<n$.

$$|p\cdot p_n^{\perp} - q\cdot q_n^{\perp}| \leq |(p-q)\cdot p_n^{\perp} + q\cdot(p_n^{\perp}-q_n^{\perp})| \leq |p-q|\cdot|p_n^{\perp}| + |q|\cdot|p_n^{\perp}-q_n^{\perp}| \leq$$
$$2\varepsilon\cdot\sqrt{n}R^{n-1}(A) + R(B)\cdot 2\varepsilon n(n-1)R^{n-2}(A).$$

$$\text{Then }\left|\frac{p\cdot p_n^{\perp}}{R^{n-1}(A)} - \frac{q\cdot q_n^{\perp}}{R^{n-1}(B)}\right| \leq \frac{|p\cdot p_n^{\perp} - q\cdot q_n^{\perp}|}{R^{n-1}(A)} + |q\cdot q_n^{\perp}|\cdot\left|\frac{R^{n-1}(B)-R^{n-1}(A)}{R^{n-1}(A)R^{n-1}(B)}\right| \leq$$

$$\leq \frac{2\varepsilon\sqrt{n}R^{n-1}(A) + 2\varepsilon n(n-1)R^{n-2}(A)R(B)}{R^{n-1}(A)} + |q|\cdot|q_n^{\perp}|\cdot\left|\frac{1}{R^{n-1}(A)} - \frac{1}{R^{n-1}(B)}\right| \leq$$

$$\leq 2\sqrt{n}\varepsilon + 2\varepsilon n(n-1) + \sqrt{n}R^n(B)\left(\frac{1}{R^{n-1}(B)} - \frac{1}{R^{n-1}(A)}\right).$$

We use $R(A) \leq R(B) + 2\varepsilon$ and the simpler notation $R = R(B)$ to bound last term after factorizing the difference of the $(n-1)$-st powers as follows:

$$R(B)\left(1 - \frac{R^{n-1}(B)}{R^{n-1}(A)}\right) \leq R\left(1 - \frac{R^{n-1}}{(R+2\varepsilon)^{n-1}}\right) = R\frac{(R+2\varepsilon)^{n-1} - R^{n-1}}{(R+2\varepsilon)^{n-1}} =$$

$$= \frac{R(R+2\varepsilon-R)}{(R+2\varepsilon)^{n-1}} \sum_{j=0}^{n-2} (R+2\varepsilon)^j R^{n-2-j} \leq \frac{2\varepsilon R}{(R+2\varepsilon)^{n-1}} \sum_{j=0}^{n-2} (R+2\varepsilon)^{n-2} \leq 2\varepsilon(n-1).$$

$$\text{Then BD}\left(\frac{M(A;p\{n-1\})}{R(A)}, \frac{M(B;q\{n-1\})}{R(B)}\right) \leq \left|\frac{p \cdot p_n}{R^{n-1}(A)} - \frac{q \cdot q_n}{R^{n-1}(B)}\right| \leq$$

$$2\varepsilon(\sqrt{n} + n(n-1) + \sqrt{n}(n-1)) = 2\varepsilon\sqrt{n}(1 + \sqrt{n}(n-1) + n - 1) \leq 2\varepsilon\sqrt{n}(\sqrt{n}(n-1) + n) =$$

$2\varepsilon n(n + \sqrt{n} - 1) \leq 3\varepsilon n^2$ because $\sqrt{n} - 1 \leq \dfrac{n}{2}$. For $n = 4$, the upper bound above is $3\varepsilon(4)^2 > 6\varepsilon \geq d_\infty$. Hence the final upper bound is $\mathrm{PRM}\big(\mathrm{PR}(A;p\{n-1\}), \mathrm{PR}(B;q\{n-1\})\big) \leq 3\varepsilon n^2$. □

Lemma E.4 (Lipschitz continuity of BMD)**.** *Let Γ be a complete bipartite graph with a vertex matching E such that any $e \in E$ has a weight $w(e) \leq \varepsilon$. Then* $\mathrm{BMD}(\Gamma) \leq \varepsilon$.

***Proof of Lemma** E.4.* By Definition 4.3, the vertex matching E has the weight $W(E) = \max\limits_{e \in E} w(e) \leq \varepsilon$. Since $\mathrm{BMD}(\Gamma) = \min\limits_{E} W(E)$ is minimized for all matchings, $\mathrm{BMD}(\Gamma) \leq \varepsilon$. □

The Lipschitz continuity of NDP in Theorem E.5 extends Theorem 5.1 to any $n \geq 2$ by using Proposition E.3 and Lemma E.4.

Theorem E.5 (Lipschitz continuity of NBM)**.** *Let a cloud $B \subset \mathbb{R}^n$ be obtained from a cloud $A \subset \mathbb{R}^n$ by perturbing every point of A within its Euclidean ε-neighborhood. Then* $\mathrm{NBM}(A, B) \leq \lambda_n \varepsilon$*, where the Lipschitz constants are $\lambda_2 = 6$, $\lambda_3 = 16$, $\lambda_n = 3n^2$ for $n > 3$ as in Proposition E.3.*

***Proof of Theorem** E.5.* Order all vertices of the given clouds A, B so that every point $p_i \in A$ has the same index as its ε-perturbation $q_i \in B$.

In Definition D.4, for any ordered points $p_1, \dots, p_{n-1} \in A$, there are points $q_1, \dots, q_{n-1} \in B$, which are ε-perturbations of $p_1, \dots, p_{n-1}$, respectively, such that $\mathrm{PRM}(\mathrm{PR}(A; p_1, \dots, p_{n-1}), \mathrm{PR}(B; q_1, \dots, q_{n-1})) \leq \lambda_n \varepsilon$ by Proposition E.3. These PRMs are weights of edges in the index-preserving vertex matching E of the complete bipartite graph $\Gamma(A; p_1, \dots, p_{n-1}; B; q_1, \dots, q_{n-1})$ for any $p_1, \dots, p_{n-1}$ and their ε-perturbations $q_1, \dots, q_{n-1}$. Then $\mathrm{BMD}(\Gamma(A; p_1, \dots, p_{n-1}; B; q_1, \dots, q_{n-1})) \leq \lambda_n \varepsilon$ by Lemma E.4. Since this conclusion holds for all (choices of) $p_1, \dots, p_{n-1} \in C$, we iteratively apply this argument for the bipartite graphs $\Gamma(A; p_1, \dots, p_k; B; q_1, \dots, q_k)$ for $1 \leq k \leq n-2$ and finally conclude that $\mathrm{NBM}(A, B) \leq \lambda_n \varepsilon$. □

The upper bounds are higher than the real ratios NBM/BD in practical examples, see Fig. **??**.

Lemma E.6 (time of PR)**.** *For any cloud $A \subset \mathbb{R}^n$ of m unordered points, any point-based representation* $\mathrm{PR}(A; p\{n-1\})$ *in Definition B.2 needs $O(n^3 + mn)$ time.*

***Proof of Lemma** E.6.* We find the center $O(A)$ and translate the cloud A of m points so that $O(A)$ becomes the origin $0 \in \mathbb{R}^n$ in time $O(m)$. We compute the $n \times n$ matrix $\mathrm{SD}(p_1, \dots, p_{n-1})$ of squared distances between $p_0 = 0, p_1, \dots, p_{n-1}$ in time $O(n^2)$. The vector $p_n^\perp$ from Lemma B.1 needs the $n \times n$ determinant computable in time $O(n^3)$. For any point $q \in A \setminus \{p_1, \dots, p_{n-1}\}$, the column of scalar products $q \cdot p_1, \dots, q \cdot p_n$ needs $O(n)$ time. The $n \times (m - n + 1)$ matrix $M(A; p\{n-1\})$ needs $O(mn)$ time. The point-based representation $\mathrm{PR}(A; p_1, \dots, p_{n-1})$ in Definition B.2 needs $O(n^3 + mn)$ time. □

Lemma E.7 (time of PRM)**.** *For any clouds $A, B \subset \mathbb{R}^n$ of m unordered points with base sequences $p\{n-1\}$ and $q\{n-1\}$, respectively, the point-based representation Metric on the equivalences classes of* $\mathrm{PR}(A; p\{n-1\})$ *and* $\mathrm{PR}(B; q\{n-1\})$ *is found in time $O(n^2 + m^{1.5} \log^n m)$ with space $O(n^2 + m \log^{n-2} m)$.*

Proof of Lemma E.7. The centers of masses $O(A), O(B)$ and radii $R(A), R(B)$ are computed in time $O(m)$.

The max metric w_D between the $n \times n$ matrices in (D.2) needs time $O(n^2)$ and space $O(n^2)$. For the bottleneck distance $w_M(\sigma)$, the $n \times (m-n+1)$ matrices of unordered columns are interpreted as fixed (not under isometry) clouds of $(m-n+1)$ points in $\mathbb{R}^n$. Then w_M can be computed in time $O(m^{1.5}\log^n m)$ with space $O(m\log^{n-2} m)$ by Theorem 6.5 in (Efrat et al., 2001). □

Theorems E.8, E.9 extend Theorems 5.2, 5.3 for $n \geq 2$.

Theorem E.8 (time of NDP)**.** *For any cloud $A \subset \mathbb{R}^n$ of m unordered points, the Nested Distributed Projection* $\mathrm{NDP}(A)$ *in Definition B.8 is computable in time $O(n^2m^n)$ with space $O(nm^n)$.*

Proof of Theorem E.8. The given cloud A has $\emptyset(m^{n-1})$ base sequences of $n-1$ ordered points $p_1,\dots,p_{n-1} \in A$. Lemma E.6 computes each $\mathrm{PR}(A;p_1,\dots,p_{n-1})$ in time $O(n^3+mn)$ with space $O(n^2+mn)$ needed to store $O(n^2)$ pairwise distances between the points $p_1,\dots,p_{n-1}$ and $O(mn)$ distances from $p_1,\dots,p_{n-1}$ to other points of A. By Definition B.8, the invariant $\mathrm{NDP}(A)$ consisting of $O(m^{n-1})$ point-based representations can be computed in time $O(n^2m^n)$ with space $O(nm^n)$ because $n \leq m$. □

Theorem E.9 (time of NBM)**.** *For any clouds $A, B \subset \mathbb{R}^n$ of m unordered points, the Nested Bottleneck Metric* $\mathrm{NBM}(A,B)$ *in Definition D.4 can be computed in time $O(m^{2n-2}(n^2+m^{1.5}\log^n m))$ with space $O(m^2(n^2+m\log^{n-2} m))$. If $n=2$, the time is $O(m^2(n^2+m^{1.5}\log m))$.*

Proof of Theorem E.9. In Definition D.4, for any fixed $1 \leq k \leq n-1$ and ordered points $p_1\dots,p_{k-1} \in A$ and $q_1\dots,q_{k-1} \in B$, the bipartite graph $\Gamma(A;p_1,\dots,p_{k-1};B;q_1,\dots,q_{k-1})$ has $V = 2(m-k+1) = O(m)$ vertices and $E = (m-k+1)^2 = O(m^2)$ edges, hence $O(m^2)$ space.

For $k = n-1$, the weight $w(e)$ of each edge e equals PRM, which needs time $O(n^2+m^{1.5}\log^n m)$ and space $O(n^2+m\log^{n-2} m)$ by Lemma E.7. For all $O(m^2)$ edges of $\Gamma(A;p_1,\dots,p_{n-2};B;q_1,\dots,q_{n-2})$, the time is $O(m^2(n^2+m^{1.5}\log^n m))$, the space is $O(m^2(n^2+m\log^{n-2} m))$. The bottleneck matching distance BMD for such a graph is computed by (Hopcroft & Karp, 1973) in time $O(E\sqrt{V}) = O(m^{2.5})$, which is dominated by the above time preparing the weights.

For all $O(m^{n-2})$ choices of ordered points $p_1,\dots,p_{n-2} \in A$ and all $O(m^{n-2})$ choices of $q_1,\dots,q_{n-2} \in B$, the Bottleneck Matching Distances for all graphs $\Gamma(A;p_1,\dots,p_{n-2};B;q_1,\dots,q_{n-2})$ are computed in time $O(m^{2n-2}(n^2+m^{1.5}\log^n m))$ with space $O(m^2(n^2+m\log^{n-2} m))$.

For any next iteration $k = n-2,\dots,1$ in Definition D.4, the parameter k goes down by 1 and the exponent of m drops by 2 each time. The sum over $k = n-1,\dots,1$ is dominated by the time and space of the first iteration.

For $n=2$, the bottleneck distance between fixed m-point clouds in $\mathbb{R}^2$ can be computed in time $O(m^{1.5}\log m)$ without an extra logarithm by Theorem 6.5 from (Efrat et al., 2001), which simplifies the time to $O(m^2(n^2+m^{1.5}\log m))$. □

Theorem E.9 improves the time $O(m^{3(n-1)}\log m)$ of another metric on rigid classes of unordered point clouds from Theorem 4.7(b) in (Widdowson & Kurlin, 2023).

Proof of Theorem 5.4. As usual, we shift both centers of mass $O(A), O(B)$ to the origin $0 \in \mathbb{R}^2$. By Definition 4.4, the distance $d = \mathrm{NBM}(A,B)$ is the Bottleneck Matching Distance $\mathrm{BMD}(\Gamma(A,B))$ computed in time $O(m^{3.5}\log m)$ by Theorem 5.3. Here $\Gamma(A,B)$ is the complete bipartite graph on $m+m$ vertices represented by $\mathrm{PR}(A;p)$ and $\mathrm{PR}(B;q)$ for all points $p \in A$ and $q \in B$.

By Definition 4.3, $\mathrm{BMD}(\Gamma(A,B))$ equals the maximum weight $w(e)$ of an edge e in a vertex matching E of $\Gamma(A,B)$, which can be considered a bijection between the m-point clouds $A \to B$. For any pair $e = (p,p')$ of matched points, the weight $w(e)$ is $\mathrm{PRM}(\mathrm{PR}(A;p),\mathrm{PR}(B;p'))$.

The distance $\mathrm{NBM}(A,B) = \delta \geq w(e)$ is an upper bound for $|R(A)-R(B)|$, where $R(A) = \max\limits_{p\in A}|p|$ and $R(B) = \max\limits_{p'\in B}|p'|$. Choose a point $p \in A$ with $|p| = R(A)$ and the positive x-axis in $\mathbb{R}^2$ through $p' \in B$ matched with p via E. Let f be the rotation of $\mathbb{R}^2$ around 0 such that $f(p)$ is also in the positive x-axis. By Definition 4.2, $f(p), p'$ in the x-axis have lengths satisfying $|p| = |f(p)|$, $|\,|p| - |p'|\,| \leq d$ and hence are d-close: $|f(p) - p'| \leq d$.

It suffices to show that the image $f(q)$ of any other point $q \in A \setminus \{p\}$ is $3\sqrt{2}d$-close to a unique point $q' \in B$ that we will find below. Since all distances and scalar products are preserved under f, we use the matrix $M(f(A); f(p))$ instead of $M(A;p)$ in computing PRM. Each column of $\dfrac{M(f(A); f(p))}{R(A)}$ consists of $\dfrac{f(q) \cdot f(p)}{|R(A)|}, \dfrac{f(q) \cdot f(p^\perp)}{|R(A)|}$, where $f(p) = (|p|, 0)$, $f(p^\perp) = (0, |p|)$, $R(A) = |p|$.

The distance $\mathrm{BD}\left(\dfrac{M(f(A); f(p))}{R(A)}, \dfrac{M(B;q)}{R(B)}\right) \leq d$ guarantees that the above column is d-close to the column of $\dfrac{q' \cdot p'}{|R(B)|}$, $\dfrac{q' \cdot p'^\perp}{|R(B)|}$ for a point $q' \in B$ determined by computing the bottleneck distance BD above. For the first scalar products involving p, p', we have $\left|\dfrac{f(q) \cdot f(p)}{R(A)} - \dfrac{q' \cdot p'}{R(B)}\right| \leq \delta$, where the first fraction is the x-coordinate of $f(q)$.

To get the x-coordinate $\dfrac{q' \cdot p'}{|p'|}$ of the point $q' \in B$, where $|p'|$ is δ-close to $R(A) = |p|$, use the triangle inequality:

$$\left|\frac{f(q) \cdot f(p)}{R(A)} - \frac{q' \cdot p'}{|p'|}\right| \leq \left|\frac{f(q) \cdot f(p)}{R(A)} - \frac{q' \cdot p'}{R(B)}\right| +$$

$$+ \left|\frac{q' \cdot p'}{R(B)} - \frac{q' \cdot p'}{|p'|}\right| \leq d + \frac{|q' \cdot p'|}{R(B)|p'|}|\,R(B) - |p'|\,| \leq$$

$$d + \frac{|q'| \cdot |p'|}{R(B)|p'|}|\,R(B) - |p'|\,| = d + \frac{|q'|}{R(B)}|\,R(B) - |p'|\,| \leq$$

$$d + |R(B) - |p'|| \leq d + |\,R(B) - |p|\,| + |\,|p| - |p'|\,| \leq$$

$$2d + |\,R(B) - |p|\,| = 2d + |R(B) - R(A)| \leq 3d.$$

Then the x-coordinates of $f(q) \in f(A)$ and $q' \in B$ differ by at most $3d$. Applying the same arguments to the scalar products involving the orthogonal vectors $p^\perp, p'^\perp$, which have the same lengths as p, p', respectively, conclude that the y-coordinates of $f(q), q'$ also differ by at most $3d$. So $|f(q) - q'| \leq \sqrt{(3d)^2 + (3d)^2} = 3\sqrt{2}d$, set $\beta(q) = q'$. □