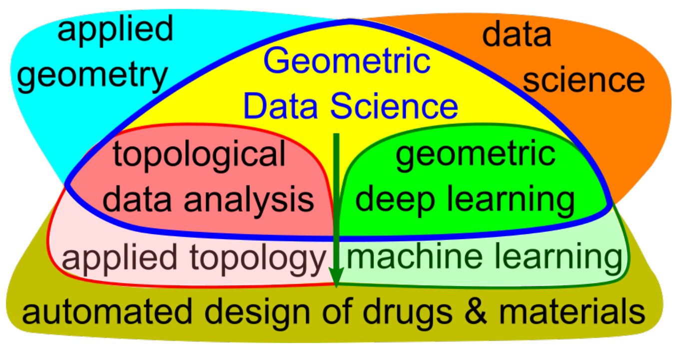
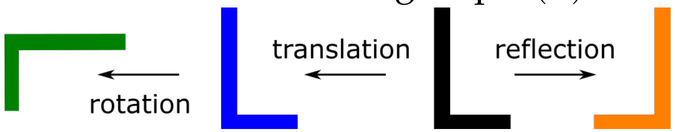
Recognizing rigid patterns of unlabeled point clouds by the complete and continuous isometry invariants with **no false negatives, no false positives** for all possible data

Daniel Widdowson, *Vitaliy Kurlin*'s Data Science group
Materials Innovation Factory, University of Liverpool, UK



A **cloud** consists of m *unlabeled* points. An **isometry** is any map preserving inter-point distances. In any Euclidean space $\mathbb{R}^n$, all isometries are compositions of translations, rotations, and reflections, and form the group $E(n)$.



If reflections are excluded, the resulting *orientation-preserving isometries* are rigid motions and form the group $SE(n)$.

If m points are *unlabeled*, such clouds can be represented by $m!$ distance matrices obtained by $m!$ permutations of given points, which is impractical.

Geometric Deep Learning experimentally outputs *invariants* preserved by the actions of $E(3)$ or $SE(3)$, optimized for specific data without using *stronger* explicitly defined invariants [1,2].

$\frac{1}{2}m(m-1)$ sorted pairwise distances are *generically complete*: distinguish m-point clouds in *general* position in $\mathbb{R}^n$ [1].

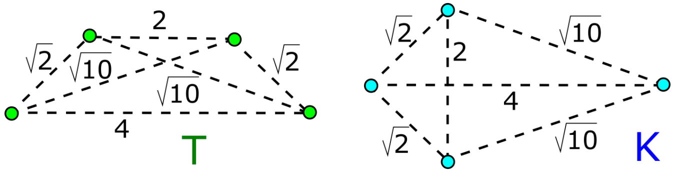
Problem: design a *complete* invariant I for unlabeled point clouds satisfying

(a) **completeness** : any clouds A, B are isometric *if and only if* $I(A) = I(B)$, equivalently I has *no false negatives* and *no false positives* for all possible data;

(b) **Lipschitz continuity** : there is a constant λ such that if any point of A is perturbed within its ε-neighborhood, then $I(A)$ changes by at most $\lambda\varepsilon$ in a metric d satisfying all the metric axioms below:
(1) *coincidence* : $d(I(A), I(B)) = 0$ if and only if the clouds A, B are isometric,
(2) $d(I(A), I(B)) = d(I(B), I(A))$,
(3) triangle inequality $d(I(A), I(B)) + d(I(B), I(C)) \geq d(I(A), I(C))$;

(c) **computability** : I, d are computable in a polynomial time in the number m of points for a fixed dimension of $\mathbb{R}^n$.

For any point $p \in C$, write distances $d_1 \leq \cdots \leq d_{m-1}$ to all points in $C - \{p\}$. The *Pointwise Distance Distribution* [2] is the unordered set of all such distance rows in the $m(m-1)$-matrix $\mathrm{PDD}(C)$.



$$\mathrm{PDD}(T;3) = \left(\begin{array}{c|ccc} 1/2 & \sqrt{2} & 2 & \sqrt{10} \\ 1/2 & \sqrt{2} & \sqrt{10} & 4 \end{array}\right)$$

$$\mathrm{PDD}(K;3) = \left(\begin{array}{c|ccc} 1/4 & \sqrt{2} & \sqrt{2} & 4 \\ 1/2 & \sqrt{2} & 2 & \sqrt{10} \\ 1/4 & \sqrt{10} & \sqrt{10} & 4 \end{array}\right)$$

New invariants: the *Simplexwise Centered Distribution* (SCD) solves the problem in any $\mathbb{R}^n$. Fix the center of C at the origin $p_0 = 0$. $\mathrm{SCD}(C)$ is the unordered set of pairs $[D(A'), M(C; A')]$ for all subsets $A \subset C$ of permutable points $p_1, \ldots, p_{n-1}$, $D(A')$ is the distance matrix on $A' = A \cup \{0\}$, $M(C; A')$ is the $(n+1) \times (m-n+1)$-matrix with permutable columns for $q \in C - A$, each consisting of n distances $|q - p_i|$, sign of determinant on $q - p_i$, $i = 0, \ldots, n-1$.

[1] M.Boutin, G.Kemper. Advances in Applied Math. 32 (2004), 709-735.

[2] Widdowson, Kurlin. NeurIPS 2022.