G-Invariant Representations and their Lipschitz Properties

Radu Balan

Department of Mathematics and Norbert Wiener Center for Harmonic Analysis and Applications
University of Maryland, College Park, MD

Joint work with: Efstratios Tsoukanis, Matthias Wellershoff

June 14, 2024
More on Harmonic Analysis, Strobl, Austria, June 9-15 2024
1 Introduction

2 G-Invariant Coorbit Representations

3 Injectivity

4 Stability
Table of Contents

1. Introduction
2. G-Invariant Coorbit Representations
3. Injectivity
4. Stability
Motivation

Certain phenomena and systems enjoy invariance to group actions.

In physics: the celebrated Noether theorem asserts that a conservation law exists for any symmetry (i.e., group invariance) of the Hamiltonian.

In data science, certain systems exhibit intrinsic invariance to group actions: in graph deep learning, graph level regression and classification must be invariant to node labeling. Specifically, this means: if \((W, X)\) is a data graph, where \(W \in \text{Sym}(\mathbb{R}^n)\) and \(X \in \mathbb{R}^{n \times d}\), then for any \(n \times n\) permutation matrix \(P\), the regression/classification function \(f, (W, X) \mapsto f(W, X)\) must satisfy \(f(PWP^T, PX) = f(W, X)\).
Problem Formulation

Consider a group $G \subset O(d)$ acting on the Euclidean space $V = \mathbb{R}^d$.

General problem

Construct an embedding map $\Phi : V \rightarrow \mathbb{R}^m$

1. **Invariance**: $\Phi(U_g x) = \phi(x)$ $\forall g \in G, x \in V$

2. **Injectivity**: if $\Phi(x) = \Phi(y)$ then there exists $g \in G$ so that $y = U_g x$.

3. **$\Phi$ is bi-Lipschitz on $(\hat{V} = V/G, d)$, where**

$$d([x], [y]) = \inf_{u \in [x], v \in [y]} \|u - v\|.$$
Approaches

Over the past years, several constructions have been proposed:

1. Invariant Polynomials: Hilbert, Noether, …, Cahill\(^1\), Bandeira\(^2\)
2. Kernels: replace monomials by other kernels, e.g. \(e^{iωx}\), \(e^{-x^2}\), \(σ(⟨x, a⟩)\)^3
3. Sorting: extends the 1-D sorting, \(x \rightarrow \downarrow x\)^4,\(^5\)

1+2: \textit{sum pooling} layer; 3: extension of \textit{max pooling} layer in deep nets\(^6\), \(^7\).

\(^2\) A. Bandeira, B. Blum-Smith, J. Kileel, J. Niles-Weed, A. Perry, A.S. Wein, Estimation under group actions: Recovering orbits from invariants, ACHA 66 (2023)
\(^3\) D. Yarotsky, Universal approximations of invariant maps by neural networks, Constructive Approximation (2021)
\(^4\) R. Balan, N. Haghani, M.Singh, Permutation Invariant Representations with Applications to Graph Deep Learning, arXiv:2203.07546
Existing Results

**Injectivity problem**

Over the past 15 years or so, there have been works that recognized the difference between *generating polynomials* and *separating invariants*.  
A seminal paper that resurfaces results on semi-algebraic sets is [8]. The method goes back to earlier works in phase retrieval [10].  
More recently, in the context of G-invariance, [11, 12], or permutation invariance [13].

---


Existing Results (2)

**Lipschitz and Bi-Lipschitz properties**

Earlier results obtain Lipschitz/bi-Lipschitz properties on compacts, or certain classes of functions. Global L/bi-L are harder to establish and typically rule out polynomial based embeddings. So far only sorting based embeddings showed such global properties \(^{14, 15, 16}\).

---

\(^{14}\) R. Balan, E. Tsoukanis, G-invariant representations using coorbits: Bi-lipschitz properties, arXiv:2308.11784

\(^{15}\) J. Cahill, J. W. Iverson, D. G. Mixon, Bilipschitz group invariants, arXiv:2305.17241

\(^{16}\) D. G. Mixon, Y. Qaddura, Injectivity, stability, and positive definiteness of max filtering, arXiv:2212.11156
Table of Contents

1 Introduction

2 G-Invariant Coorbit Representations

3 Injectivity

4 Stability
Coorbit Representations

Let $V$ be a $d$-dimensional Hilbert space and $G$ a finite group of size $N = |G|$ acting unitarily on $V$, $\{U_g, g \in G\}$. The quotient space $\hat{V} = V/G$ is the set of orbits $[x] = \{U_gx, g \in G\}$ induced by the group action, where for $x, y \in V$, $x \sim y$ iff $y = U_gx$ for some $g \in G$. $(\hat{V}, d)$ becomes a metric space with the natural distance

$$d([x], [y]) = \min_{g \in G} \|x - U_gy\|$$

Fix a generator $w \in V$ (call it, window or template) and consider the nonlinear map induced by sorting its coorbit:

$$\phi_w : V \rightarrow \mathbb{R}^N, \quad \phi_w(x) = \downarrow (\langle x, U_gw \rangle)_{g \in G}.$$ 

where $\downarrow (y) = (y_{\pi(i)})_{i \in [N]}$ is the non-increasing sorting operator:

$y_{\pi(1)} \geq \cdots \geq y_{\pi(N)}$. 

Radu Balan (UMD)
Invariant Coorbit Representations

For a collection $\mathbf{w} = (w_1, \ldots, w_p) \in V^p$ let

$$\Phi_{\mathbf{w}} : V \to \mathbb{R}^{N \times p} \ , \ \Phi_{\mathbf{w}}(x) = [\phi_{w_1}(x) | \cdots | \phi_{w_p}(x)].$$

For a subset $S \subset [N] \times [p]$ of cardinal $m = |S|$, let

$$\Phi_{\mathbf{w},S} : V \to l^2(S) \sim \mathbb{R}^m \ , \ \Phi_{\mathbf{w},S}(x) = (\Phi_{\mathbf{w}}(x))|_S$$

be the restriction of $\Phi_{\mathbf{w}}$ to $S$. For a linear operator $\mathcal{L} : l^2(S) \to \mathbb{R}^m$, let

$$\Psi_{\mathbf{w},S,\mathcal{L}} : V \to \mathbb{R}^m \ , \ \Psi_{\mathbf{w},\mathcal{L}}(x) = \mathcal{L}(\Phi_{\mathbf{w},S}(x))$$

be the “projection” of $\Phi_{\mathbf{w},S}$ through $\mathcal{L}$ into $\mathbb{R}^m$.

Problems: Construct $(\mathbf{w}, S)$ so that $\Phi_{\mathbf{w},S}$ is a bi-Lipschitz embedding of $\hat{V}$. Construct $(\mathbf{w}, S, \mathcal{L})$ so that $\Psi_{\mathbf{w},S,\mathcal{L}}$ is bi-Lipschitz.
Invariant Coorbit Representations (2)

Special cases:
1. If $G = S_n$ and $V = \mathbb{R}^{n \times d}$ with action $(P, X) \mapsto PX$, then \(^{17}\) introduced the embedding $\beta_A(X) = \downarrow (XA)$, for key $A \in \mathbb{R}^{d \times D}$ and sorting operator acting independently in each column.

Equivalent recasting: Let $w_1 = \delta_1 \cdot a_1^T, \ldots, w_D = \delta_1 \cdot a_D^T$, where $\delta_1 = (1, 0, \ldots, 0)^T$ and $A = [a_1 | \cdots | a_D]$. Then note $\phi_{w_1}(X) = \downarrow (Xa_1) \otimes 1_{(n-1)!}$. Thus $\Phi_w(X) = \beta_A(X) \otimes 1_{(n-1)!}$. Thus $\beta_A(X) = \Phi_{w,S}(X)$ for an appropriate subset $S \subset [n!] \times [D]$ of size $nD$.

2. The max filter introduced in \(^{18}\) for some template $w \in V$ is defined by $\langle \langle \cdot, w \rangle \rangle : V \to \mathbb{R}$, $\langle \langle x, w \rangle \rangle = \max_{g \in G} \langle x, U_g w \rangle$. Equivalent recasting: $\langle \langle x, w \rangle \rangle = \Phi_{w,S}(X)$, for $S = \{1\}$.

\(^{17}\)R. Balan, N. Haghani, M. Singh, Permutation Invariant Representations with Applications to Graph Deep Learning, arXiv:2203.07546 (2022)

Table of Contents

1. Introduction
2. G-Invariant Coorbit Representations
3. Injectivity
4. Stability
Minimal embeddings

Setup: Let $V$ be a $d$-dimensional Hilbert space and $G$ a finite group of size $N = |G|$ acting unitarily on $V$, $\{U_g, g \in G\}$. For a subset $S \subset [N] \times [p]$ of cardinal $m = |S|$, let

$$\Phi_{w,S} : V \to l^2(S) \sim \mathbb{R}^m, \quad \Phi_{w,S}(x) = (\Phi_w(x))|_S$$

be the restriction of $\Phi_w$ to $S$.

A typical injectivity result asserts that for $p \geq p_{\text{min}}$ and a generic $w \in V^p$, for any $S$ of cardinal $m \geq m_{\text{min}}$ that satisfy certain shape conditions, the map $\Phi_{w,S}$ is injective on $\hat{V}$. $(p_{\text{min}}, m_{\text{min}})$ depend on specific rep.
Current injectivity results

Setup: \( \dim(V) = d \), \( G \) finite group of size \( N \). Denote by \( d_G = \dim\{x \in V ; \ U_g x = x \ \forall g \in G\} \).

**Theorem (R.B, E.Tsoukanis ‘23)**

Let \( p \geq 2d - d_G \). Then for a generic \( w \in V^p \) the following holds: for every \( S \subset [N] \times [p] \) that satisfies the property that for every \( i \in [p] \) there is a \( k \in [N] \) so that \( (k, i) \in S \), the map \( \Phi_{w,S} \) is injective on \( \hat{V} \).

**Theorem (R.B, E.Tsoukanis ‘23)**

If \( S \subset [N] \times [p] \) satisfies a stronger intersection property, namely, for each \( i \in [p] \) there are distinct \( k_1, \ldots, k_n \in [N] \) so that \( (k_1, i), \ldots, (k_n, i) \in S \), then the lower bound \( 2d - d_G \) can be replaced by
\[
p_{n,\min} = \max_{q \in [n], g, h \in G^n} \frac{1}{q}(\gamma_{q}^{g,h} - d_G - 1)
\]
where
\[
\gamma_{q}^{g,h} = \operatorname{semialg} - \dim\{(x, y) \in V \times V : \dim(\operatorname{span}\{ U_{g(1)} x - U_{h(1)} y, \ldots, U_{g(n)} x - U_{h(n)} y \}) \leq q\}.
\]
Table of Contents

1. Introduction
2. G-Invariant Coorbit Representations
3. Injectivity
4. Stability
Injectivity implies (bi-Lipschitz) Stability

**Theorem**

For fixed $w \in V^p$ and $S \subset [N] \times [p]$, where $|S| = m$, suppose that the map $\Phi_{w,S} : V \to \mathbb{R}^m$, is injective on $V/G$. Then, $\exists 0 < a \leq b < \infty$ such that $\forall (x, y) \in V$, $x \approx y$

$$a d([x], [y]) \leq \|\Phi_{w,S}(x) - \Phi_{w,S}(y)\|_2 \leq b d([x], [y]).$$
Injectivity implies (bi-Lipschitz) Stability

**Theorem**

For fixed \( w \in V^p \) and \( S \subset [N] \times [p] \), where \( |S| = m \), suppose that the map \( \Phi_{w,S} : V \to \mathbb{R}^m \), is injective on \( V/G \). Then, \( \exists 0 < a \leq b < \infty \) such that \( \forall (x, y) \in V, x \preceq y \)

\[
ad([x], [y]) \leq \|\Phi_{w,S}(x) - \Phi_{w,S}(y)\|_2 \leq b d([x], [y]).
\]

**Corollary**

For max filter bank \( \Phi : \mathbb{R}^d/G \to \mathbb{R}^m \), injectivity implies stability.
Lemma

Let $w \in V^p$, $S \subset [N] \times [p]$ and

$$B = \max_{\sigma_1, \ldots, \sigma_p \subset G} \lambda_{\max} \left( \sum_{i=1}^{p} \sum_{g \in \sigma_i} g \cdot w_i w_i^T U_g^T \right)$$

where $S_i = \{j \in [N], (i, j) \in S\}$ and $m_i = |S_i|$. Then $\Phi_{w, S} : \hat{V} \to \mathbb{R}^m$ is Lipschitz with constant upper bounded by $\sqrt{B}$.
The proof of the main Theorem is by contradiction.

1. If lower Lipschitz constant vanishes, then it must vanish locally: there are \((x_n)_n, (y_n)_n\) such that

\[
\lim_{n \to \infty} \frac{\|\Phi_{\omega, S}(x_n) - \Phi_{\omega, S}(y_n)\|^2}{d([x_n], [y_n])^2} = 0
\]

and

\[
\lim_{n \to \infty} x_n = \lim_{n \to \infty} y_n = z_1, \quad \|x_n\| = 1, \quad \|y_n\| \leq 1, \quad \|z_1\| = 1
\]

and they are aligned with one another:

\[
\|x_n - y_n\| = \min_{g \in G} \|x_n - U_g y_n\|
\]

(1)

\[
\|x_n - z_1\| = \min_{g \in G} \|x_n - U_g z_1\|
\]

(2)

\[
\|y_n - z_1\| = \min_{g \in G} \|y_n - U_g z_1\|
\]

(3)
Lower Lipschitz bound

2. We construct inductively $z_2, z_3, \ldots, z_d$ such that for all $1 \leq k \leq d - 1$:

\[ \|z_{k+1}\| \ll \|z_k\|, \quad \dim(\text{span}(z_1, \ldots, z_k)) = k \]

and the local lower Lipschitz constant vanishes in a convex set
\[ \{\sum_{r=1}^{k} a_r z_r , |a_r - 1| < \epsilon\}. \]

3. For $k = d$ this construction defines a non-empty open set
\[ \{\sum_{r=1}^{k} a_r z_r , |a_r - 1| < \epsilon\} \]
where the local lower Lipschitz constant vanishes.

4. Finally, we can construct $u, v \neq 0$, so that $x = u + \sum_{r=1}^{d} z_r$ and $y = v + \sum_{r=1}^{d} z_r$ satisfy $x \neq y$ and yet

\[ \Phi_{w,S}(x) = \Phi_{w,S}(y). \]

This contradicts the injectivity hypothesis.
Happy Birthday Charly!

Thank you!