

G-Invariant Representations using Coorbits: Bi-Lipschitz Properties

Radu Balan, Efstratios Tsoukanis

Department of Mathematics

University of Maryland

College Park, MD 20742

email: {rvbalan,etsoukan}@umd.edu

August 22, 2023

Abstract

Consider a real vector space \mathcal{V} and a finite group G acting unitary on \mathcal{V} . We study the general problem of constructing a stable embedding, whose domain is the quotient of the vector space modulo the group action, and whose target space is an Euclidean space. First, we construct an embedding Ψ , which is well defined and injective in the quotient vector space. Then, we show that for the map Ψ injectivity implies stability. The embedding scheme we introduce is based on taking a fixed subset out of sorted orbit $\downarrow \langle U_g w_i, x \rangle_{g \in G}$, where w_i are appropriate vectors.

1 Introduction

Machine learning techniques have impressive results when we feed them with large sets of data. In some cases, our training set can be small but we know that there are some underlying symmetries in the data structure. For example, in graph theory problems each graph is being represented as an adjacent matrix of the labeled nodes of the graph; any relabeling of the nodes shouldn't change the output of our classification or regression algorithm.

A possible solution for this problem is to increase our training set by adding, for each data point of the set, the whole orbit generated by the group action. One problem that arises is that it is computationally costly to find such highly symmetric function.

Another solution is to embed our data into an Euclidean space \mathbb{R}^m with a symmetry-invariant embedding Ψ and then use \mathbb{R}^m as our feature space. It is not enough for our embedding to be symmetric invariant, it should also separate data orbits. Finally, we require certain stability conditions so that small perturbations don't affect our predictions. This problem is an instance of *invariant machine learning* [19, 15, 3, 11, 8, 16, 22, 10, 12, 17].

The most common group action in invariant machine learning are permutations [21, 9, 7] reflections [18] and translations [14]. Also, there are very interesting results in the case of equivariant embeddings [20, 16].

Our work is influenced by [11] where it is shown that $m \approx 2d$ separating invariants are enough for an orbit-separating embedding, and by [9, 19] where the *max filter* is introduced. We work with a generalization of the *max filter*: instead of choosing the maximum element of the orbit we choose other subsets of orbit. The problem of finding permutation invariant embeddings seems to be closely connected to the phase retrieval problem where there already are a lot of important results [5, 6, 2, 1, 4, 13].

In the second chapter we construct an injective embedding for the case of a finite subset of a vector space \mathcal{V} . In the third chapter we propose two different embedding schemes for a d -dimensional vector space \mathcal{V} . In forth chapter we prove that for the proposed embeddings, injectivity implies stability and that a linear projection into a target space of dimension $2d$ preserves injectivity. Finally in fifth we establish some results for specific group actions.

1.1 Notation

Let $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ be a d -dimensional real vector space, where $d \geq 2$. Assume (G, \cdot) is a finite group of order $|G| = N$ acting unitarily on \mathcal{V} . For every $g \in G$, we denote by $U_g x$ the group action on vector $x \in \mathcal{V}$. Let $\hat{\mathcal{V}} = \mathcal{V} / \sim$ denote the quotient space with respect to the action of group G . We denote by $[x]$ the orbit of vector x , i.e. $[x] = \{U_g x : g \in G\}$. Consider now the natural metric, $\mathbf{d} : \hat{\mathcal{V}} \times \hat{\mathcal{V}} \rightarrow \mathbb{R}$,

$$\mathbf{d}([x], [y]) = \min_{h_1, h_2 \in G} \|U_{h_1} x - U_{h_2} y\| = \min_{g \in G} \|x - U_g y\|. \quad (1)$$

Our goal is to construct a bi-Lipschitz Euclidean embedding of the metric space $(\mathcal{V}, \mathbf{d})$ into an Euclidean space \mathbb{R}^m .

Specifically, we want to construct a function $\Psi : \mathcal{V} \rightarrow \mathbb{R}^m$ such that

1. $\Psi(U_g x) = \Psi(x)$, $\forall x \in \mathcal{V}$, $\forall g \in G$,
2. If $x, y \in \mathcal{V}$ are such that $\Psi(x) = \Psi(y)$, then there exist $g \in G$ such that $y = U_g x$,
3. There are $0 < a < b < \infty$ such that for any $x, y \in \mathcal{V}$

$$a \mathbf{d}([x], [y])^2 \leq \|\Psi(x) - \Psi(y)\|^2 \leq b(\mathbf{d}([x], [y]))^2.$$

The invariance property (1) lifts Ψ to a map $\hat{\Psi}$ acting on the quotient space $\hat{\mathcal{V}} = \mathcal{V} / \sim$, where $x \sim y$ if and only if $y = U_g x$ for some $g \in G$:

$$\hat{\Psi} : \hat{\mathcal{V}} \rightarrow \mathbb{R}^m, \quad \hat{\Psi}([x]) = \Psi(x), \quad \forall [x] \in \hat{\mathcal{V}}.$$

If a G -invariant map Ψ satisfies property (2) we say that Ψ separates the G -orbits in \mathbb{R}^d .

Our construction for the embedding Ψ is based on a non-linear sorting map.

Notation 1.1. Let $\downarrow : \mathbb{R}^r \rightarrow \mathbb{R}^r$ be the operator that takes as input a vector in \mathbb{R}^r and returns a monotonically decreasing sorted vector of same length r that has same entries as the input vector.

For a number $p \in \mathbb{N}$, fix a p -tuple of vectors $\mathbf{w} = (w_1, \dots, w_p) \in \mathcal{V}^p$. For any $i \in [p]$ and $j \in [N]$ we define the operator $\Phi_{w_i, j} : \mathcal{V} \rightarrow \mathbb{R}$ so that $\Phi_{w_i, j}(x)$ is the j -th coordinate of vector $\downarrow \langle U_g w_i, x \rangle_{g \in G}$. Now fix a set $S \subset [N] \times [p]$ such that $|S| = m$, and for $i \in [p]$, set $S_i = \{k \in [N] : (k, i) \in S\}$ (the i^{th} column of S). Denote by m_i the cardinal of the set S_i , $m_i = |S_i|$. Thus $m = \sum_{i=1}^p m_i$.

Notation 1.2. The coorbit embedding $\Phi_{\mathbf{w}, S}$ associated to windows $\mathbf{w} \in \mathcal{V}^p$ and index set $S \subset [N] \times [p]$ is given by the map

$$\Phi_{\mathbf{w}, S} : \mathcal{V} \rightarrow \mathbb{R}^m, \quad \Phi_{\mathbf{w}, S}(x) = [\{\Phi_{w_1, j}(x)\}_{j \in S_1}, \dots, \{\Phi_{w_p, j}(x)\}_{j \in S_p}] \in \mathbb{R}^m. \quad (2)$$

Let $\ell : \mathbb{R}^m \rightarrow \mathbb{R}^q$ be a linear transformation.

Notation 1.3. The embedding $\Psi_{\mathbf{w},S,\ell}$ associated to windows $\mathbf{w} \in \mathcal{V}^p$, index set $S \subset [N] \times [p]$ and linear map $\ell : \mathbb{R}^m \rightarrow \mathbb{R}^q$ is given by the map

$$\Psi_{\mathbf{w},S,\ell} = \ell \circ \Phi_{\mathbf{w},S} : \mathcal{V} \rightarrow \mathbb{R}^q, \quad \Psi_{\mathbf{w},S,\ell}(x) = \ell(\Phi_{\mathbf{w},S}(x)) \quad (3)$$

obtained by composition of ℓ with the coorbit embedding $\Phi_{\mathbf{w},S}$.

In this paper we focus on stability properties of maps $\Phi_{\mathbf{w},S}$ and $\Psi_{\mathbf{w},S,\ell}$.

Informally, our main two results Theorem 2.1 and Theorem 3.1 state that: (1) "injectivity" implies "(bi-Lipschitz) stability", and (2) stable bi-Lipschitz embedding can be achieved in an Euclidean space twice the dimension of the input data space.

For the rest of the paper we shall use interchangeably $\Phi_{i,j}$ instead of $\Phi_{w_i,j}$. We also overload the notation and use the same letter to denote maps $\Phi_{\mathbf{w},S}$ and $\Psi_{\mathbf{w},S,\ell}$ either defined on \mathcal{V} or $\hat{\mathcal{V}}$.

2 Stability of Embedding

Suppose that for $\mathbf{w} = (w_1, \dots, w_p) \in \mathcal{V}^p$ and $S \subset [N] \times [p]$ the map $\Phi_{\mathbf{w},S}$ is injective. In this case, we claim the map $\Phi_{\mathbf{w},S}$ is also bi-Lipschitz. We state this claim in the next Theorem 2.1.

Theorem 2.1. *Let G be a finite subgroup of $O(d)$. For fixed $\mathbf{w} \in \mathcal{V}^p$ and $S \subset [N] \times [p]$, where $|S| = m$, suppose that the map $\Phi_{\mathbf{w},S} : \mathcal{V} \rightarrow \mathbb{R}^m$, is injective on the quotient space $\hat{\mathcal{V}} = \mathcal{V}/G$. Then, there exist $0 < a \leq b < \infty$ such that for all $(x, y) \in \hat{\mathcal{V}}$, where $x \approx y$*

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

A corollary of Theorem 2.1 is that for *max filter* bank introduced in [9], injectivity implies stability:

Corollary 2.2. *Let a finite group G acting unitarily on a vector space \mathcal{V} . If the max filter bank $\Phi_{\mathbf{w},S_{max}} : \hat{\mathcal{V}} \rightarrow \mathbb{R}^m$ is injective then $\Phi_{\mathbf{w},S_{max}}$ is also bi-Lipschitz, where $\Phi_{\mathbf{w},S_{max}}$ is the embedding associated to $\mathbf{w} \in \mathcal{V}^p$ and set $S_{max} = \{(1, k), k \in [p]\}$.*

In the next two subsections, we prove Theorem 2.1.

2.1 Upper Lipschitz bound

The upper bound can be obtained easily.

Lemma 2.3. *Consider G be a finite group of size N acting unitarily on \mathcal{V} . Let $\mathbf{w} \in \mathcal{V}^p$ and $S \subset [N] \times [p]$. Let also,*

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

where $S_i = \{j \in [N], (i, j) \in S\}$ and $m_i = |S_i|$. Then $\Phi_{\mathbf{w}, S} : (\widehat{V}, \mathbf{d}) \rightarrow \mathbb{R}^m$ is Lipschitz with constant upper bounded by \sqrt{B} .

Proof. For fixed $x, y \in \mathcal{V}$, $i \in [p]$, $j \in S_i$, let $\psi_{i,j} : \mathbb{R} \rightarrow \mathbb{R}$, where

$$\psi_{i,j}(t) = \Phi_{i,j}((1-t)x + ty) = \langle (1-t)x + ty, U_{g_t} w_i \rangle.$$

By Lebesgue differentiation theorem we have that $\psi_{i,j}$ is differentiable almost everywhere, consequently $\Phi_{i,j}$ is also differentiable almost everywhere. Notice that

$$\psi'_{i,j}(t) = \langle y - x, U_{g_t} w_i \rangle$$

for almost every $t \in \mathbb{R}$. Specifically for almost every $t \in \mathbb{R}$ such that exists $\epsilon > 0$ such that $g_{t-\epsilon, t+\epsilon}$ is the same group element.

From fundamental theorem of calculus we get

$$\Phi_{i,j}(x) - \Phi_{i,j}(y) = \int_0^1 \frac{d}{dt} \Phi_{i,j}((1-t)x + ty) dt.$$

Therefore,

$$\Phi_{i,j}(y) - \Phi_{i,j}(x) = \int_0^1 \langle y - x, U_{g_{jt}} w_i \rangle dt$$

so

$$\|\downarrow\{\Phi_{i,j}(x)\}_{j \in S_i} - \downarrow\{\Phi_{i,j}(y)\}_{j \in S_i}\| \leq \int_0^1 \left(\sum_{j \in S_i} \langle y - x, U_{g_{jt}} w_i \rangle^2 \right)^{1/2} dt \leq \sqrt{B_i} \|x - y\|$$

where $B_i = \max_{\substack{\sigma \subset G, \\ |\sigma| = |S_i|}} \lambda_{max} \left(\sum_{g \in \sigma} U_g w_i w_i^T U_g^T \right)$.

Hence,

$$\|\Phi_{\mathbf{w}, S}(x) - \Phi_{\mathbf{w}, S}(y)\|^2 = \sum_{i=1}^p \|\Phi_{i,j}(y) - \Phi_{i,j}(x)\|^2 \leq \sum_{i=1}^p B_i \|x - y\|^2 \leq B \mathbf{d}(x, y)^2.$$

□

2.2 Lower Lipschitz bound

Before we show that a strictly positive lower Lipschitz bound exists, we will prove some helpful geometric results.

2.2.1 Geometric Analysis of Coorbits

First let us introduce some additional notation. For fixed $i \in [p]$, $j \in [N]$ and $x \in \mathcal{V}$ we define the following non-empty subset of the group G :

$$L^{i,j}(x) = \{g \in G : \langle U_g w_i, x \rangle = \Phi_{i,j}(x)\}. \quad (4)$$

Let also the map

$$\Delta^{i,j}(x) = \begin{cases} \min_{g \notin L^{i,j}(x)} (|\langle U_g w_i, x \rangle - \Phi_{i,j}(x)|) \frac{1}{\|w_i\|}, & \text{if } L^{i,j}(x) \neq G \\ \frac{\|x\|}{\|w_i\|}, & \text{if } L^{i,j}(x) = G. \end{cases} \quad (5)$$

Lemma 2.4.

a. For any $x \in \mathcal{V}$, $i \in [p]$, and $j \in [N]$,

$$|\{g \in G, \langle U_g w_i, x \rangle > \Phi_{i,j}(x)\}| \leq j - 1 \quad (6)$$

$$|\{g \in G, \langle U_g w_i, x \rangle < \Phi_{i,j}(x)\}| \leq N - j. \quad (7)$$

b. For any $x \in \mathcal{V}$, $i \in [p]$, and $j \in [N - 1]$,

(i) either $\Phi_{i,j}(x) = \Phi_{i,j+1}(x)$, in which case $L^{i,j}(x) = L^{i,j+1}(x)$,

or

(ii) $\Phi_{i,j}(x) > \Phi_{i,j+1}(x)$, in which case $L^{i,j}(x) \neq L^{i,j+1}(x)$ and

$$\{g \in G, \langle U_g w_i, x \rangle > \Phi_{i,j+1}(x)\} = \cup_{k \leq j} L^{i,k}(x) \quad , \quad \left| \cup_{k \leq j} L^{i,k}(x) \right| = j$$

and

$$\{g \in G, \langle U_g w_i, x \rangle < \Phi_{i,j}(x)\} = \cup_{k \geq j+1} L^{i,k}(x) \quad , \quad \left| \cup_{k \geq j+1} L^{i,k}(x) \right| = N - j.$$

Proof. Recall that $\Phi_{i,j}(x)$, is the j -th coordinate of the sorted in decreasing order vector $\downarrow \{\langle U_g w_i, x \rangle\}_{g \in G}$. Suppose that

$$|\{g \in G, \langle U_g w_i, x \rangle > \Phi_{i,j}(x)\}| > j - 1.$$

Then there are at least j , distinct elements of group G , (h_1, \dots, h_j) , such that $\langle U_{h_k} w_i, x \rangle > \langle U_g w_i, x \rangle$, $\forall k \in [j]$. But this is a contradiction. Similarly, if

$$|\{g \in G, \langle U_g w_i, x \rangle < \Phi_{i,j}(x)\}| > N - j$$

there exist at least $N - j + 1$, distinct elements of group G , (h_1, \dots, h_j) such that $\langle U_{h_k} w_i, x \rangle < \langle U_g w_i, x \rangle \forall k \in [N - j + 1]$, which is also a contradiction.

Moreover, if $\Phi_{i,j}(x) = \Phi_{i,j+1}(x)$, a group element g achieves $\Phi_{i,j}(x)$ if and only if, also achieves $\Phi_{i,j+1}(x)$, therefore $L^{i,j}(x) = L^{i,j+1}(x)$. On the other hand, if $\Phi_{i,j}(x) > \Phi_{i,j+1}(x)$ then $L^{i,j}(x)$ and $L^{i,j+1}(x)$ are disjoint sets. Assuming otherwise, there is $g \in L^{i,j}(x) \cap L^{i,j+1}(x)$, but then $\Phi_{i,j}(x) = \langle U_g w_i, x \rangle = \Phi_{i,j+1}(x)$. Now assume that

$$\{g \in G, \langle U_g w_i, x \rangle > \Phi_{i,j+1}(x)\} \neq \emptyset.$$

Without loss of generality

$$\{g \in G, \langle U_g w_i, x \rangle > \Phi_{i,j+1}(x)\} > j$$

so there exists at least $j + 1$ group elements (h_1, \dots, h_{j+1}) , such that

$$\langle U_{h_k} w_i, x \rangle > \Phi_{i,j+1}, \forall k \in [j + 1]$$

but this is a contradiction. Similarly,

$$\{g \in G, \langle U_g w_i, x \rangle < \Phi_{i,j}(x)\} = \cup_{k \geq j+1} L^{i,k}(x), \quad \left| \cup_{k \geq j+1} L^{i,k}(x) \right| = N - j.$$

□

Note that for any $w_1, \dots, w_p \in \mathcal{V} \setminus \{0\}$ the subset $L^{i,j}(x) \subset G$ has the following “nesting” property.

Lemma 2.5. *For any $x, y \in \mathcal{V}$ such that $\|y\| < \frac{1}{2}\Delta^{i,j}(x)$, we have that $L^{i,j}(x + y) \subset L^{i,j}(x)$. Furthermore,*

$$\{g \in G, \langle U_g w_i, x \rangle > \Phi_{i,j}(x)\} \subset \{g \in G, \langle U_g w_i, x + y \rangle > \Phi_{i,j}(x + y)\},$$

$$\{g \in G, \langle U_g w_i, x \rangle < \Phi_{i,j}(x)\} \subset \{g \in G, \langle U_g w_i, x + y \rangle < \Phi_{i,j}(x + y)\},$$

$$\{g \in G, \langle U_g w_i, x + y \rangle \geq \Phi_{i,j}(x + y)\} \subset \{g \in G, \langle U_g w_i, x \rangle \geq \Phi_{i,j}(x)\}$$

and

$$\{g \in G, \langle U_g w_i, x + y \rangle \leq \Phi_{i,j}(x + y)\} \subset \{g \in G, \langle U_g w_i, x \rangle \leq \Phi_{i,j}(x)\}.$$

Proof. Suppose that exists $g \in G$ such that $g \in L^{i,j}(x+y)$ but $g \notin L^{i,j}(x)$. Without loss of generality assume that $\langle U_g w_i, x \rangle < \Phi_{i,j}(x)$. Then for every $h \in \cup_{k \leq j} L^{i,k}(x)$

$$\langle U_h w_i, x+y \rangle - \langle U_g w_i, x+y \rangle \geq \langle U_h w_i, x \rangle - \langle U_g w_i, x \rangle - 2\|y\|\|w_i\| > 0.$$

On the other hand, $\langle U_g w_i, x+y \rangle = \Phi_{i,j}(x+y)$. Thus

$$\cup_{k \leq j} L^{i,k}(x) \subset \{h \in G, \langle U_h w_i, x+y \rangle > \Phi_{i,j}(x+y)\}.$$

But the set $\cup_{k \leq j} L^{i,k}(x)$ contains at least j elements (since each $L^{i,j}(x)$ is non-empty) and so we derived a contradiction with Lemma 2.4(a) Equation (6). \square

Lemma 2.6. For $i \in [p]$ and $j \in [N]$, fix vectors $x, y \in \mathcal{V}$ and positive numbers $c_1, c_2 > 0$ such that $\max(c_1, c_2)\|y\| < \frac{1}{4}\Delta^{i,j}(x)$. Then $L^{i,j}(x+c_1y) = L^{i,j}(x+c_2y)$.

Proof.

Assume that exist $g_1 \in L^{i,j}(x+c_2y)$ with $g_1 \notin L^{i,j}(x+c_1y)$. Without loss of generality assume that $\langle U_{g_1} w_i, x+c_1y \rangle < \Phi_{i,j}(x+c_1y)$. Let $q > j$ be the smallest integer such that $g_1 \in L^{i,q}(x+c_1y)$. Then $\Phi_{i,q}(x+c_1y) = \langle U_{g_1} w_i, x+c_1y \rangle < \Phi_{i,j}(x+c_1y)$. By Lemma 2.4 (b)(ii),

$$\left| \cup_{r \leq j} L^{i,r}(x+c_1y) \right| = q-1 \geq j,$$

and $g_1 \notin \cup_{r \leq j} L^{i,r}(x+c_1y)$. On the other hand, from Lemma 2.4 (a), Equation (6),

$$|\{h \in G, \langle U_h w_i, x+c_2y \rangle > \Phi_{i,j}(x+c_2y)\}| \leq j-1$$

Hence

$$\cup_{r \leq j} L^{i,r}(x+c_1y) \setminus \{h \in G, \langle U_h w_i, x+c_2y \rangle > \Phi_{i,j}(x+c_2y)\} \neq \emptyset$$

Therefore there exists $h \in \cup_{r \leq j} L^{i,r}(x+c_1y)$ such that

$$\langle U_h w_i, x+c_2y \rangle \leq \Phi_{i,j}(x+c_2y) = \langle U_{g_1} w_i, x+c_2y \rangle. \quad (8)$$

On the other hand, by Lemma 2.5, $g_1 \in L^{i,j}(x)$.

But if $\langle U_h w_i, x \rangle - \langle U_{g_1} w_i, x \rangle > 0$ then

$$\langle U_h w_i, x + c_2 y \rangle - \langle U_{g_1} w_i, x + c_2 y \rangle \geq \|w_i\|(\Delta^{i,j}(x) - 2c_2\|y\|) > 0$$

which is a contradiction with (8). If $\langle U_h w_i, x \rangle - \langle U_{g_1} w_i, x \rangle < 0$ then

$$\langle U_{g_1} w_i, x + c_1 y \rangle - \langle U_h w_i, x + c_1 y \rangle \geq \|w_i\|(\Delta^{i,j}(x) - 2c_1\|y\|) > 0$$

which is a contradiction with $h \in \cup_{r \leq j} L^{i,r}(x + c_1 y)$. Therefore $\langle U_h w_i, x \rangle = \langle U_{g_1} w_i, x \rangle$ and thus $h \in L^{i,j}(x)$. But then

$$\begin{aligned} 0 &\geq \langle U_h w_i, x + c_2 y \rangle - \langle U_{g_1} w_i, x + c_2 y \rangle = \langle U_h w_i, c_2 y \rangle - \langle U_{g_1} w_i, c_2 y \rangle \\ &= c_2(\langle U_h w_i, y \rangle - \langle U_{g_1} w_i, y \rangle) \end{aligned}$$

and

$$\begin{aligned} 0 &< \langle U_h w_i, x + c_1 y \rangle - \langle U_{g_1} w_i, x + c_1 y \rangle = \langle U_h w_i, c_1 y \rangle - \langle U_{g_1} w_i, c_1 y \rangle \\ &= c_1(\langle U_h w_i, y \rangle - \langle U_{g_1} w_i, y \rangle). \end{aligned}$$

Which contradict one another since $c_1, c_2 > 0$.

□

Lemma 2.7. *For any $w_1, \dots, w_p \in \mathcal{V} \setminus \{0\}$ and $x \in \mathcal{V}$, the sets $L^{i,j}(x)$ and perturbation bounds $\Delta^{i,j}(x)$ have the following properties:*

1. For any $t > 0$, $L^{i,j}(tx) = L^{i,j}(x)$.
2. For any $i \in [p]$, $j \in [N]$, and $t > 0$, $\Delta^{i,j}(tx) = t\Delta^{i,j}(x)$.
3. For any $i \in [p]$, $j \in [N]$, and $x \in \mathcal{V} \setminus \{0\}$, $\Delta^{i,j}(x) > 0$.

Proof. 1.,2. For $t > 0$, $\Phi_{i,j}(tx) = t\Phi_{i,j}(x)$, from where the claims follow from the definitions of $L^{i,j}(x)$ and $\Delta^{i,j}(x)$.

3. This claim follows from definitions of $\Delta_{i,j}$ which is the minimum of a finite set of positive numbers.

□

Lemma 2.8. Fix $w_i \in \mathcal{V} \setminus \{0\}$ and $j \in [N]$. For any $k > 1$, fix $z_1 \in \mathcal{V}$ of unit norm, $\|z_1\| = 1$, and choose z_2, \dots, z_k inductively such that

$$\|z_{l+1}\| \leq \min\left(\frac{1}{4}\Delta^{i,j}\left(\sum_{r=1}^l z_r\right), \frac{1}{4}\|z_l\|\right) \quad \forall l \in [k-1].$$

For any positive numbers $a_1, \dots, a_k \in \left(1 - \frac{1}{16k}\Delta^{i,j}\left(\sum_{r=1}^k z_r\right), 1 + \frac{1}{16k}\Delta^{i,j}\left(\sum_{r=1}^k z_r\right)\right)$ the following hold true:

1.

$$\frac{1}{4}\Delta^{i,j}\left(\sum_{r=1}^k a_r z_r\right) < \Delta^{i,j}\left(\sum_{r=1}^k z_r\right) < 4\Delta^{i,j}\left(\sum_{r=1}^k a_r z_r\right). \quad (9)$$

2.

$$L^{i,j}\left(\sum_{r=1}^k a_r z_r\right) = L^{i,j}\left(\sum_{r=1}^k z_r\right), \quad (10)$$

$$\bigcup_{l \leq j} L^{i,l}\left(\sum_{r=1}^k a_r z_r\right) = \bigcup_{l \leq j} L^{i,l}\left(\sum_{r=1}^k z_r\right), \quad (11)$$

$$\bigcup_{l \geq j} L^{i,l}\left(\sum_{r=1}^k a_r z_r\right) = \bigcup_{l \geq j} L^{i,l}\left(\sum_{r=1}^k z_r\right). \quad (12)$$

3. For every $e \in \mathcal{V}$ where $\|e\| < \frac{1}{16}\Delta^{i,j}\left(\sum_{r=1}^k z_r\right)$.

$$L^{i,j}\left(\sum_{r=1}^k a_r z_r + e\right) = L^{i,j}\left(\sum_{r=1}^k z_r + e\right) \quad (13)$$

Proof.

1. Case 1. Suppose $L^{i,j}\left(\sum_{r=1}^k z_r\right) = G$. From lemma 2.5 we know that

$$G = L^{i,j}\left(\sum_{r=1}^k z_r\right) \subset L^{i,j}\left(\sum_{r=1}^{k-1} z_r\right) \subset \dots \subset L^{i,j}(z_1).$$

Therefore, $L^{i,j}(z_r) = G$, $\forall r \in [k]$ and consequently $L^{i,j}(a_r z_r) = G$, $\forall r \in [k]$.

Moreover, $a_1, \dots, a_k \in (7/8, 9/8)$. Therefore,

$$\Delta^{i,j}\left(\sum_{r=1}^k a_r z_r\right) = \frac{1}{\|w_i\|} \left\| \sum_{r=1}^k a_r z_r \right\| \leq \frac{9}{8\|w_i\|} \left(\sum_{r=1}^k \|z_r\| \right) < \frac{3}{2\|w_i\|} \|z_1\|$$

and

$$\Delta^{i,j}\left(\sum_{r=1}^k a_r z_r\right) = \frac{1}{\|w_i\|} \left\| \sum_{r=1}^k a_r z_r \right\| \leq \frac{1}{\|w_i\|} \left(\frac{7}{8} \|z_1\| - \frac{9}{8} \sum_{r=2}^k \|z_r\| \right) > \frac{1}{2\|w_i\|} \|z_1\|.$$

Similarly,

$$\Delta^{i,j}\left(\sum_{r=1}^k z_r\right) = \frac{1}{\|w_i\|} \left\| \sum_{r=1}^k z_r \right\| \leq \frac{1}{\|w_i\|} \left(\sum_{r=1}^k \|z_r\| \right) < \frac{4}{3\|w_i\|} \|z_1\|$$

and

$$\Delta^{i,j}\left(\sum_{r=1}^k z_r\right) = \frac{1}{\|w_i\|} \left\| \sum_{r=1}^k z_r \right\| \leq \frac{1}{\|w_i\|} \left(\|z_1\| - \sum_{r=2}^k \|z_r\| \right) > \frac{2}{3\|w_i\|} \|z_1\|.$$

So,

$$\frac{1}{4} \Delta^{i,j}\left(\sum_{r=1}^k z_r\right) \leq \Delta^{i,j}\left(\sum_{r=1}^k a_r z_r\right) \leq 4 \Delta^{i,j}\left(\sum_{r=1}^k z_r\right).$$

Case 2. Now assume that $L^{i,j}(\sum_{r=1}^k z_r) \neq G$. Fix $g_1 \in G$ that achieves $\Delta^{i,j}(\sum_{r=1}^k a_r z_r)$, i.e.

$$\frac{1}{\|w_i\|} \left| \langle U_{g_1} w_i, \sum_{r=1}^k a_r z_r \rangle - \Phi_{i,j}\left(\sum_{r=1}^k a_r z_r\right) \right| = \Delta^{i,j}\left(\sum_{r=1}^k a_r z_r\right)$$

and $g_2 \in L^{i,j}(\sum_{r=1}^k a_r z_r)$. Then,

$$\begin{aligned} \Delta^{i,j}\left(\sum_{r=1}^k a_r z_r\right) &= \frac{1}{\|w_i\|} \left| \langle U_{g_1} w_i, \sum_{r=1}^k a_r z_r \rangle - \langle U_{g_2} w_i, \sum_{r=1}^k a_r z_r \rangle \right| \\ &\geq \frac{1}{\|w_i\|} \left| \langle U_{g_1} w_i, \sum_{r=1}^k z_r \rangle - \langle U_{g_2} w_i, \sum_{r=1}^k z_r \rangle \right| - \sum_{r=1}^k 2|1 - a_r| \|z_r\| \\ &\geq \Delta^{i,j}\left(\sum_{r=1}^k z_r\right) - 2 \sum_{r=1}^k |1 - a_r| \|z_r\| > \frac{1}{2} \Delta^{i,j}\left(\sum_{r=1}^k z_r\right). \end{aligned}$$

and,

$$\begin{aligned}
\Delta^{i,j}(\sum_{r=1}^k a_r z_r) &= \frac{1}{\|w_i\|} |\langle U_{g_1} w_i, \sum_{r=1}^k a_r z_r \rangle - \langle U_{g_2} w_i, \sum_{r=1}^k a_r z_r \rangle| \\
&\leq \frac{1}{\|w_i\|} |\langle U_{g_1} w_i, \sum_{r=1}^k z_r \rangle - \langle U_{g_2} w_i, \sum_{r=1}^k z_r \rangle| + \sum_{r=1}^k 2|1 - a_r| \|z_r\| \\
&\geq \Delta^{i,j}(\sum_{r=1}^k z_r) + 2 \sum_{r=1}^k |1 - a_r| \|z_r\| \leq 2\Delta^{i,j}(\sum_{r=1}^k z_r).
\end{aligned}$$

Therefore,

$$\frac{1}{2}\Delta^{i,j}(\sum_{r=1}^k a_r z_r) \leq \Delta^{i,j}(\sum_{r=1}^k z_r) \leq 2\Delta^{i,j}(\sum_{r=1}^k a_r z_r).$$

2. From definition of vectors z_j and positive constants a_j we have that

$$\|\sum_{r=1}^k (1-a_r)z_r\| \leq \sum_{r=1}^k |a_r-1| \leq \frac{1}{16}\Delta^{i,j}(\sum_{r=1}^k z_r) \leq \frac{1}{4} \min(\Delta^{i,j}(\sum_{r=1}^k z_r), \Delta^{i,j}(\sum_{r=1}^k a_r z_r))$$

Apply Lemma 2.5 twice: first with $x = \sum_{r=1}^k z_r$ and $y = \sum_{r=1}^k (a_r - 1)z_r$ to obtain

$$L^{i,j}(\sum_{r=1}^k z_r) \subset L^{i,j}(\sum_{r=1}^k a_r z_r)$$

and a second time with $x = \sum_{r=1}^k a_r z_r$ and $y = \sum_{r=1}^k (1 - a_r)z_r$ to obtain

$$L^{i,j}(\sum_{r=1}^k a_r z_r) \subset L^{i,j}(\sum_{r=1}^k z_r).$$

So we conclude that

$$L^{i,j}(\sum_{r=1}^k a_r z_r) = L^{i,j}(\sum_{r=1}^k z_r).$$

Let $g_1 \in \cup_{l=1}^j L^{i,l}(\sum_{r=1}^k z_r)$.

If $g_1 \in L^{i,j}(\sum_{r=1}^k z_r)$ then we just showed that $g_1 \in L^{i,j}(\sum_{r=1}^k a_r z_r)$.

If $g_1 \notin L^{i,j}(\sum_{r=1}^k z_r)$ then for every $h \in \cup_{l \geq j} L^{i,l} L^{i,l}(\sum_{r=1}^k z_r)$

$$\begin{aligned} \langle U_{g_1} w_i, \sum_{r=1}^k a_r z_r \rangle - \langle U_h w_i, \sum_{r=1}^k a_r z_r \rangle &\geq \langle U_{g_1} w_i, \sum_{r=1}^k z_r \rangle - \|w_i\| \left\| \sum_{r=1}^k |1 - a_r| z_r \right\| \\ &> \|w_i\| \left(\Delta^{i,j}(\sum_{r=1}^k z_r) - \left\| \sum_{r=1}^k |1 - a_r| z_r \right\| \right) > 0 \end{aligned}$$

But from Lemma 2.4 $|\cup_{l \geq j} L^{i,l} L^{i,l}(\sum_{r=1}^k z_r)| \geq N - j + 1$. So $g_1 \in \cup_{l=1}^j L^{i,l}(\sum_{r=1}^k z_r)$. Therefore,

$$\cup_{l \leq j} L^{i,l}(\sum_{r=1}^k z_r) \subset \cup_{l \leq j} L^{i,l}(\sum_{r=1}^k a_r z_r).$$

The other inclusions are obtained similarly.

3.

We prove the equality between complements: $(L^{i,j}(\sum_{r=1}^k z_r + e))^c = (L^{i,j}(\sum_{r=1}^k a_r z_r + e))^c$.

First notice that by Lemma 2.5 we have that $L^{i,j}(\sum_{r=1}^k z_r + e) \subset L^{i,j}(\sum_{r=1}^k z_r)$ and $L^{i,j}(\sum_{r=1}^k a_r z_r + e) \subset L^{i,j}(\sum_{r=1}^k a_r z_r)$. From Lemma 2.8 part (2) we have that $L^{i,j}(\sum_{r=1}^k z_r) = L^{i,j}(\sum_{r=1}^k a_r z_r)$.

Take $g \in L^{i,j}(\sum_{r=1}^k z_r + e)$ and $h \in (L^{i,j}(\sum_{r=1}^k z_r + e))^c$. Hence $\langle U_g w_i, \sum_{r=1}^k z_r + e \rangle \neq \langle U_h w_i, \sum_{r=1}^k z_r + e \rangle$.

There are two cases:

Case 1. $h \in L^{i,j}(\sum_{r=1}^k z_r) \setminus L^{i,j}(\sum_{r=1}^k z_r + e)$. Thus $\langle U_g w_i, \sum_{r=1}^k z_r \rangle = \langle U_h w_i, \sum_{r=1}^k z_r \rangle$. Therefore $\langle U_g w_i, e \rangle \neq \langle U_h w_i, e \rangle$. On the other hand $h \in L^{i,j}(\sum_{r=1}^k a_r z_r)$ since $L^{i,j}(\sum_{r=1}^k z_r) = L^{i,j}(\sum_{r=1}^k a_r z_r)$. Hence $\langle U_g w_i, \sum_{r=1}^k a_r z_r \rangle = \langle U_h w_i, \sum_{r=1}^k a_r z_r \rangle$, which implies $\langle U_g w_i, \sum_{r=1}^k a_r z_r + e \rangle \neq \langle U_h w_i, \sum_{r=1}^k a_r z_r + e \rangle$. Thus $h \in (L^{i,j}(\sum_{r=1}^k a_r z_r + e))^c$.

Case 2. $h \in G \setminus L^{i,j}(\sum_{r=1}^k z_r)$. Thus $\langle U_g w_i, \sum_{r=1}^k z_r \rangle \neq \langle U_h w_i, \sum_{r=1}^k z_r \rangle$. In this case $|\langle U_h w_i, \sum_{r=1}^k z_r \rangle - \langle U_g w_i, \sum_{r=1}^k z_r \rangle| \geq \|w_i\| \Delta^{i,j}(\sum_{r=1}^k z_r)$ and

$$\begin{aligned} &\left| \langle U_h w_i, \sum_{r=1}^k a_r z_r + e \rangle - \langle U_g w_i, \sum_{r=1}^k a_r z_r + e \rangle \right| \geq \\ &\geq \left| \langle U_h w_i, \sum_{r=1}^k z_r \rangle - \langle U_g w_i, \sum_{r=1}^k z_r \rangle \right| - \left| \langle U_h w_i, \sum_{r=1}^k (a_r - 1) z_r + e \rangle \right| - \left| \langle U_g w_i, \sum_{r=1}^k (a_r - 1) z_r + e \rangle \right| \geq \\ &\geq \|w_i\| \left[\Delta^{i,j}(\sum_{r=1}^k z_r) - 2 \left(\sum_{r=1}^k |a_r - 1| \|z_r\| + \|e\| \right) \right] > 0 \quad (*) \end{aligned}$$

Hence again $h \in \left(L^{i,j}(\sum_{r=1}^k a_r z_r + e)\right)^c$.

This proves that $\left(L^{i,j}(\sum_{r=1}^k z_r + e)\right)^c \subset \left(L^{i,j}(\sum_{r=1}^k a_r z_r + e)\right)^c$.

The reverse inclusion is shown similarly, with $\Delta^{i,j}(\sum_{r=1}^k z_r)$ replaced by $\Delta^{i,j}(\sum_{r=1}^k a_r z_r)$ in (*). \square

2.2.2 Positivity of the Lower Lipschitz Constant

Now we prove that the lower Lipschitz bound must be positive if the embedding map $\Phi_{\mathbf{w},S}$ is injective. We do so by contradiction.

The strategy is the following: Assume the lower Lipschitz constant is zero.

- First we find a unit norm vector z_1 where the local lower Lipschitz constant vanishes.
- Next we construct inductively a sequence of non-zero vectors z_2, z_3, \dots, z_k so that the local lower Lipschitz constant vanishes in a convex set of the form $\{\sum_{r=1}^k a_r z_r, |a_r - 1| < \delta\}$ for some $\delta > 0$ small enough, and where sets $L^{i,j}$ remain constant.
- For $k = d$ this construction defines a non-empty open set where the local lower Lipschitz constant vanishes and $L^{i,j}$ remain constants. This allows us to 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 $x \not\sim y$ and yet $\Phi_{\mathbf{w},S}(x) = \Phi_{\mathbf{w},S}(y)$. This contradicts the injectivity hypothesis.

First, we show that if the lower bound is zero then it can be achieved locally.

Lemma 2.9. *Fix $\mathbf{w} = (w_1, \dots, w_p) \in \mathcal{V}^p$ and $S \subset [N] \times [p]$. If the lower Lipschitz constant of map $\Phi_{\mathbf{w},S}$ is zero, then there exist sequences $(x_n)_n, (y_n)_n$ in \mathcal{V} such that*

$$\lim_{n \rightarrow \infty} \frac{\|\Phi_{\mathbf{w},S}(x_n) - \Phi_{\mathbf{w},S}(y_n)\|^2}{d(x_n, y_n)^2} = 0$$

and, additionally, satisfy the following relations:

1. (convergence) They share a common limit z_1 ,

$$\lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} y_n = z_1, \tag{14}$$

with $\|z_1\| = 1$;

2. (boundedness) For all k :

$$\|x_n\| = 1 \quad (15)$$

$$\|y_n\| \leq 1 \quad (16)$$

3. (alignment) For all k :

$$\|x_n - y_n\| = \min_{g \in G} \|x_n - U_g y_n\| \quad (17)$$

$$\|x_n - z_1\| = \min_{g \in G} \|x_n - U_g z_1\| \quad (18)$$

$$\|y_n - z_1\| = \min_{g \in G} \|y_n - U_g z_1\| \quad (19)$$

Proof. Because the lower Lipschitz bound of map $\Phi_{\mathbf{w},S}$ is zero we have that

$$\inf_{\substack{x, y \in \mathcal{V} \\ x \neq y}} \frac{\|\Phi_{\mathbf{w},S}(x) - \Phi_{\mathbf{w},S}(y)\|^2}{\mathbf{d}(x, y)^2} = 0.$$

Thus, we can find sequences $(x_n)_n, (y_n)_n \in \mathcal{V}$ such that

$$\lim_{n \rightarrow \infty} \frac{\|\Phi_{\mathbf{w},S}(x_n) - \Phi_{\mathbf{w},S}(y_n)\|^2}{\mathbf{d}(x_n, y_n)^2} = 0$$

Now, notice that for all $t > 0$ we have $\Phi_{\mathbf{w},S}(tx) = t\Phi_{\mathbf{w},S}(x)$ and $\mathbf{d}(tx, ty) = t\mathbf{d}(x, y)$. So, for every $t > 0$

$$\frac{\|\Phi_{\mathbf{w},S}(x_n) - \Phi_{\mathbf{w},S}(y_n)\|^2}{\mathbf{d}(x_n, y_n)^2} = \frac{\|t\Phi_{\mathbf{w},S}(x_n) - t\Phi_{\mathbf{w},S}(y_n)\|^2}{\mathbf{d}(tx_n, ty_n)^2} = \frac{\|\Phi_{\mathbf{w},S}(tx_n) - \Phi_{\mathbf{w},S}(ty_n)\|^2}{\mathbf{d}(tx_n, ty_n)^2}.$$

By setting $t = \frac{1}{\max(\|x_n\|, \|y_n\|)}$ we can always assume that both x_n and y_n , lie in the unit ball, and what is more thanks to the symmetry of the formulas we can additionally assume that one of the sequences, say x_n , lies on unit sphere. In other words, $\|x_n\| = 1$ and $\|y_n\| \leq 1$ for all $n \in \mathbb{N}$.

Because of this, we can find a convergent subsequence $(x_{n_k})_k$ of $(x_n)_n$ with $x_{n_k} \rightarrow x_\infty$. Similarly, we can find a convergent subsequence $(y_{n_{k_l}})_l$ of $(y_{n_k})_k$ with $y_{n_{k_l}} \rightarrow y_\infty$. Clearly, $x_{n_{k_l}} \rightarrow x_\infty$. For easiness of notation, we denote the sequences $(x_{n_{k_l}})_l$ and $(y_{n_{k_l}})_l$ by $(x_n)_n$ and $(y_n)_n$, respectively.

Next, suppose that $x_\infty \approx y_\infty$. Then,

$$\frac{\|\Phi_{\mathbf{w},S}(x_\infty) - \Phi_{\mathbf{w},S}(y_\infty)\|^2}{\mathbf{d}(x_\infty, y_\infty)^2} = \lim_{k \rightarrow \infty} \frac{\|\Phi_{\mathbf{w},S}(x_n) - \Phi_{\mathbf{w},S}(y_n)\|^2}{\mathbf{d}(x_n, y_n)^2} = 0,$$

and thus, $\Phi_{\mathbf{w},S}(x_\infty) = \Phi_{\mathbf{w},S}(y_\infty)$, which contradict the injectivity assumption. Hence, $x_\infty \sim y_\infty$.

Now, let us denote by g_∞ a group element such that $x_\infty = U_{g_\infty}y_\infty$. Observe that $\lim_{n \rightarrow \infty} \|x_n - U_{g_\infty}y_n\| = 0$. For each $n \in \mathbb{N}$ there exists at least one element $g_n \in G$, which achieves the Euclidean distance between x_n and $U_{g_\infty}y_n$, i.e. satisfying $\mathbf{d}(x_n, U_{g_\infty}y_n) = \|x_n - U_{g_n g_\infty}y_n\|$. But G is a finite group, meaning that, as n goes to infinity, there must exist an element $g_0 \in G$ for which $g_n = g_0$ for infinitely many n . Let $(n_m)_m$ be the sequence of all such indices. We see that $\mathbf{d}(x_{n_m}, U_{g_\infty}y_{n_m}) = \|x_{n_m} - U_{g_0 g_\infty}y_{n_m}\|$ for all $m \in \mathbb{N}$. Finally, for every $m \in \mathbb{N}$, let $g_m \in G$ be a group element that achieves the Euclidean distance between x_{n_m} and x_∞ , that is

$$\mathbf{d}(x_{n_m}, x_\infty) = \|U_{g_m}x_{n_m} - x_\infty\|.$$

Denote $U_{g_m}x_{n_m}$ by x_n and $U_{g_m g_0 g_\infty}y_{n_m}$ by y_n . So far we obtained two sequences $(x_n)_n$ and $(y_n)_n$ that satisfy (14-18). Now let $h_n \in G$ denote a group element so that $d(y_n, z_1) = \|y_n - U_{h_n}z_1\|$. Since G is finite, pass to a subsequence (again indexed by n) so that $h_n = h_0$. Therefore $d(y_n, z_1) = \|y_n - U_{h_0}z_1\| \leq \|y_n - z_1\|$. But $\lim_{y_n} = z_1$. Thus $U_{h_0}z_1 = z_1$. This shows (19) and the lemma is now proved. \square

In what follows, we will denote by $H(z)$ the stabilizer group of z ; recall that

$$H(z) = \{g \in G : U_g z = z\}.$$

For a fixed vector z we define the strictly positive number

$$\rho_0(z) = \begin{cases} \min_{g \in G \setminus H(z)} \|z - U_g z\|, & \text{if } H(z) \neq G \\ \|z\|, & \text{if } H(z) = G. \end{cases}$$

Assume N_0 is large enough so that $d(x_{1,k}, z_1) < \frac{1}{8}\rho_0(z_1)$ and $d(x_{1,k}, y_{1,k}) < \frac{1}{8}\rho_0(z_1)$ for all $k > N_0$. Then

$$\|y_{1,k} - z_1\| \leq \|y_{1,k} - x_{1,k}\| + \|x_{1,k} - z_1\| = d(y_{1,k}, x_{1,k}) + d(x_{1,k}, z_1) < \frac{\rho_0(z_1)}{4}.$$

Lemma 2.10. *Assume that $\|u\|, \|v\| < \frac{1}{4}\Delta_0(z_1)$ and let $x = z_1 + u$ and $y = z_1 + v$. Then, the following properties hold:*

1. $\mathbf{d}(x, z_1) = \|u\|$ and $\mathbf{d}(y, z_1) = \|v\|$,

2. $\mathbf{d}(x, y) = \min_{g \in H(z_1)} \|u - U_g v\| = \min_{g \in H(z_1)} \|U_g u - v\|$, and

3. the following are equivalent:

(a) $\mathbf{d}(x, y) = \|u - v\|$,

(b) $\|u - v\| \leq \|U_g u - v\|$, for all $g \in H(z_1)$,

(c) $\langle u, v \rangle \geq \langle U_g u, v \rangle$, for all $g \in H(z_1)$.

Proof. 1. If $u = 0$ then the claim follows. If $u \neq 0$, then $\mathbf{d}(x, z_1) = \min_{g \in G} \|x - U_g z_1\| = \min_{g \in G} \|z_1 - U_g z_1 + u\| \leq \|u\|$. From the other hand, suppose that minimum is achieved for a permutation $g \in G$. If $g \in H(z_1)$, then $\mathbf{d}(x, z_1) = \|u\|$. If $g \notin H(z_1)$, then $\mathbf{d}(x, z_1) > \|u\| \leq \mathbf{d}(x, z_1)$, which is a contradiction.

2. Obviously $\mathbf{d}(x, z_1) \leq \min_{g \in H(z_1)} \|U_g u - v\|$. On the other hand, for $g \in G \setminus K$ and $h \in G$,

$$\begin{aligned} \|U_g x - y\| &= \|U_g z_1 - z_1 + U_g u - v\| \\ &\geq \|U_g z_1 - z_1\| - \|u\| - \|v\| \\ &\geq \rho_0(z_1) - 2\|u\| - 2\|v\| + \|U_h u - v\| \\ &\geq \mathbf{d}(x, y). \end{aligned}$$

3. • (a) \Rightarrow (b). If $\mathbf{d}(x, y) = \|u - v\|$, then $\|u - v\| \leq \|U_g x - y\| = \|U_g z_1 - z_1 + U_g u - v\|$, $\forall g \in G$. For $g \in H(z_1)$ this reduces to (b)
- (b) \Rightarrow (a). Assume that $\forall g \in H(z_1)$, $\|u - v\| \leq \|U_g u - v\|$. Then $\|u - v\| = \|x - y\| \leq \|U_g u - v\| = \|U_g x - y\|$. For, $g \in G \setminus H(z_1)$ $\|U_g x - y\| = \|U_g z_1 - z_1 + U_g u - v\| \geq \|U_g z_1 - z_1\| - \|u\| - \|v\| \geq \rho_0(z_1) - 2\|u\| - 2\|v\| + \|u - v\| \geq \|u - v\| = \|x - y\|$. Thus, $\mathbf{d}(x, y) = \|x - y\| = \|u - v\|$
- (b) \Leftrightarrow (c) is immediate from definition of inner product

□

Remark 2.11. Applying Lemma 2.10 to two sequences $(x_k)_k$ and $(y_k)_k$ that satisfy (14-17) in Lemma 2.9, it follows that $d(x_k, z_1) = \|x_k - z_1\|$ and $d(y_k, z_1) = \|y_k - z_1\|$ for k large enough. Hence alignment must occur from some rank on.

Lemma 2.12. For fixed $i \in [p]$, $j \in S_i$ and two sequences $(x_n)_n, (y_n)_n$ produced by Lemma 2.9, we denote by $g_{1,n,i,j}$ the group elements that achieves $\Phi_{i,j}(x_n)$ and by $g_{2,n,i,j}$ the group element that achieves the $\Phi_{i,j}^j(y_n)$. That is $\Phi_{i,j}(x_n) = \langle U_{g_{1,n,i,j}} w_i, x_n \rangle$ and $\Phi_{i,j}(y_n) = \langle U_{g_{2,n,i,j}} w_i, y_n \rangle$.

We can find a sequence of natural numbers $(n_r)_r$, such that, $g_{1,n_r,i,j} = g_{1,i,j}$ and $g_{2,n_r,i,j} = g_{2,i,j} \forall r \in \mathbb{N}, i \in [p], j \in S_i$.

Proof. For $i = 1, j = 1$ there is a subsequence $(x_{n_m})_m$ such that $g_{1,1,1,n_m} = g_{1,1,1}$ for every $m \in \mathbb{N}$. Similarly, for $i = 1, j = 2$ we can find a subsequence of $(x_{n_m})_m$, lets call it $(x_{n_l})_l$, such that $g_{1,1,2,n_l} = g_{1,1,2}, \forall l \in \mathbb{N}$. So by induction after $\sum_{i \in [p]} m_i = m$ steps we construct a subsequence of $(x_n)_n$ lets call it $(x_{n_m})_m$ such that $g_{1,i,j,n_m} = g_{1,i,j}$ for every $i \in [p], j \in S_i$. Starting from sequence $(y_{1,n_m})_m$ we repeat the same procedure concluding in a subsequence $(y_{1,n_r})_r$ such that $g_{2,i,j,n_r} = g_{2,i,j}$ for every $r \in \mathbb{N}, i \in [p], j \in S_i$. Notice that sequences $(x_{n_r})_r$ and $(y_{n_r})_r$ that from now on we will call them $(x_n)_n$ and $(y_n)_n$ for easiness of notation, satisfy the assumptions of lemma. \square

For sequences $(x_n)_n, (y_n)_n$ and z_1 defined before, let $u_n = x_n - z_1$ and $v_n = y_n - z_1$. Notice that

$$\begin{aligned} \|\Phi_{\mathbf{w},S}(x_n) - \Phi_{\mathbf{w},S}(y_n)\|^2 &= \sum_{i=1}^p \sum_{j \in S_i} |\langle U_{g_{1,i,j}} w_i, x_n \rangle - \langle U_{g_{2,i,j}} w_i, y_n \rangle|^2 \\ &= \sum_{i=1}^p \sum_{j \in S_i} |\langle U_{g_{1,i,j}} w_i - U_{g_{2,i,j}} w_i, z_1 \rangle \\ &\quad + \langle w_i, U_{g_{1,i,j}^{-1}} u_n - U_{g_{2,i,j}^{-1}} v_n \rangle|^2. \end{aligned}$$

This sequence converge to 0, as $k \rightarrow \infty$ while also $u_n, v_n \rightarrow 0$. So we conclude that for each $i \in [p]$ and $j \in S_i$, $\langle U_{g_{1,i,j}} w_i - U_{g_{2,i,j}} w_i, z_1 \rangle = 0$. So

$$\|\Phi_{\mathbf{w},S}(x_n) - \Phi_{\mathbf{w},S}(y_n)\|^2 = \sum_{i=1}^p \sum_{j \in S_i} |\langle w_i, U_{g_{1,i,j}^{-1}} u_n - U_{g_{2,i,j}^{-1}} v_n \rangle|^2.$$

Thus we have

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^p \sum_{j \in S_i} |\langle w_i, U_{g_{1,i,j}^{-1}} u_n - U_{g_{2,i,j}^{-1}} v_n \rangle|^2}{\|u_n - v_n\|^2} = 0 \quad (20)$$

where $\|u_n\|, \|v_n\| \rightarrow 0$, so for large enough n we have that $\|u_n\|, \|v_n\| \leq \frac{1}{4}\rho_0(z_1)$. Recall that from Lemma 2.10, we conclude that exists $N_0 \in \mathbb{N}$, such that $\|u_n - v_n\| \leq \|U_g u_n - v_n\|$ for all $g \in H(z_1)$ and $k \geq N_0$.

Lemma 2.13. Fix $p \in \mathbb{N}$, $\mathbf{w} \in \mathcal{V}^p$ and $S \subset [N] \times [p]$. Let $\Delta : \mathcal{V} \rightarrow \mathbb{R}$, where $\Delta(x) = \min_{(i,j) \in [p] \times [N]} \Delta^{i,j}(x)$, where the map $\Delta^{i,j}$ is defined in (5). Fix nonzero vectors $z_1, \dots, z_k \in \mathcal{V}$, such that

$$\|z_1\| = 1, \quad \langle z_i, z_j \rangle = 0, \quad \forall i, j \in [k], \quad i \neq j$$

and

$$\|z_{l+1}\| \leq \min \left(\frac{1}{4} \Delta \left(\sum_{r=1}^l z_r \right), \frac{1}{4} \|z_l\| \right), \quad \forall l \in [k-1].$$

Assume that the local lower Lipschitz constant of $\Phi_{\mathbf{w},S}$ vanishes at $z_1 + z_2 + \dots + z_k$.

1. The local lower Lipschitz constant vanishes on the non-empty convex box $\{\sum_{r=1}^k a_r z_r, |a_r - 1| < \frac{1}{16k} \Delta(\sum_{l=1}^k z_l)\}$ centered at $z_1 + z_2 + \dots + z_k$.
2. Assume $\Phi_{\mathbf{w},S}$ is injective. If $k < d$ then there exists a nonzero vector z_{k+1} such that:
 - (i) $\langle z_{k+1}, z_j \rangle = 0, \quad \forall j \in [k]$;
 - (ii) $\|z_{k+1}\| \leq \min \left(\frac{1}{4} \Delta(\sum_{r=1}^k z_r), \frac{1}{4} \|z_k\| \right)$; and
 - (iii) The local lower Lipschitz constant vanishes at $z_1 + z_2 + \dots + z_{k+1}$, i.e. there are sequences of vectors $(x_n)_n, (y_n)_n$ such that

$$\lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} y_n = \sum_{r=1}^{k+1} z_r$$

and

$$\lim_{n \rightarrow \infty} \frac{\|\Phi_{\mathbf{w},S}(x_n) - \Phi_{\mathbf{w},S}(y_n)\|^2}{\mathbf{d}(x_n, y_n)^2} = 0.$$

Proof. 1. Let $(x_n)_n, (y_n)_n$ be sequences in \mathcal{V} such that

$$\lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} y_n = \sum_{r=1}^k z_r$$

and

$$\lim_{n \rightarrow \infty} \frac{\|\Phi_{\mathbf{w},S}(x_n) - \Phi_{\mathbf{w},S}(y_n)\|^2}{\mathbf{d}(x_n, y_n)^2} = 0.$$

Claim: For any $a_1, \dots, a_k \in \left(1 - \frac{1}{16k} \Delta(\sum_{r=1}^k z_r), 1 + \frac{1}{16k} \Delta(\sum_{r=1}^k z_r)\right)$ the sequences

$$\tilde{x}_n = x_n + \sum_{r=1}^k (a_r - 1) z_r$$

and

$$\tilde{y}_n = y_n + \sum_{r=1}^k (a_r - 1) z_r$$

also achieve a zero lower Lipschitz constant, i.e.

$$\lim_{n \rightarrow \infty} \frac{\|\Phi_{\mathbf{w}, S}(\tilde{x}_n) - \Phi_{\mathbf{w}, S}(\tilde{y}_n)\|^2}{\mathbf{d}(\tilde{x}_n, \tilde{y}_n)^2} = 0.$$

First we denote by u_n and v_n the difference sequences x_n and y_n to their common limit $\sum_{r=1}^k z_r$,

$$u_n = x_n - \sum_{r=1}^k z_r = \tilde{x}_n - \sum_{r=1}^k a_r z_r$$

and

$$v_n = y_n - \sum_{r=1}^k z_r = \tilde{y}_n - \sum_{r=1}^k a_r z_r.$$

Sequences $(u_n)_n$ and $(v_n)_n$ converge to zero. Therefore there exists $M_0 \in \mathbb{N}$ such that $\forall n \geq M_0$

- (a) $\|u_n\| = \|x_n - \sum_{r=1}^k z_r\| < \frac{1}{16} \Delta(\sum_{r=1}^k z_r)$
- (b) $\|u_n\| = \|\tilde{x}_n - \sum_{r=1}^k a_r z_r\| < \frac{1}{16} \Delta(\sum_{r=1}^k a_r z_r)$
- (c) $\|v_n\| = \|y_n - \sum_{r=1}^k z_r\| < \frac{1}{16} \Delta(\sum_{r=1}^k z_r)$
- (d) $\|v_n\| = \|\tilde{y}_n - \sum_{r=1}^k a_r z_r\| < \frac{1}{16} \Delta(\sum_{r=1}^k a_r z_r)$.

Thus from part (3) of Lemma 2.8, Lemma 2.5 and part (2) of Lemma 2.8 we have that for any $n \geq M_0$ and $(i, j) \in S$

$$L^{i,j}(\tilde{x}_n) = L^{i,j}(x_n) \subset L^{i,j}\left(\sum_{r=1}^k z_r\right) = L^{i,j}\left(\sum_{r=1}^k a_r z_r\right)$$

and

$$L^{i,j}(\tilde{y}_n) = L^{i,j}(y_n) \subset L^{i,j}\left(\sum_{r=1}^k z_r\right) = L^{i,j}\left(\sum_{r=1}^k a_r z_r\right).$$

Therefore,

$$\begin{aligned}
0 &= \lim_{n \rightarrow \infty} \frac{\|\Phi_{\mathbf{w},S}(x_n) - \Phi_{\mathbf{w},S}(y_n)\|^2}{\mathbf{d}(x_n, y_n)^2} = \\
&= \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^p \sum_{j \in S_i} |\langle w_i, U_{g_{1,i,j}}^{-1} u_n - U_{g_{2,i,j}}^{-1} v_n \rangle|^2}{\|u_n - v_n\|^2} = \\
&= \lim_{n \rightarrow \infty} \frac{\|\Phi_{\mathbf{w},S}(\tilde{x}_n) - \Phi_{\mathbf{w},S}(\tilde{y}_n)\|^2}{\mathbf{d}(\tilde{x}_n, \tilde{y}_n)^2},
\end{aligned}$$

where

$$g_{1,i,j} \in L^{i,j}(x_n) = L^{i,j}(\tilde{x}_n)$$

and

$$g_{2,i,j} \in L^{i,j}(y_n) = L^{i,j}(\tilde{y}_n).$$

This proves the lower Lipschitz constant of $\Phi_{\mathbf{w},S}$ vanishes at $\sum_{r=1}^k a_r z_r$.

2. Let two sequences $(x_n)_n, (y_n)_n$ that both converge to $\sum_{r=1}^k z_r$, and achieve lower Lipschitz bound zero for map $\Phi_{\mathbf{w},S}$. We align sequences $(x_n)_n$ and $(y_n)_n$ to satisfy the properties of Lemma 2.9. We denote by $a_n = P_{E_k} x_n$ and $b_n = P_{E_k} y_n$ the orthogonal projections of the sequences $(x_n)_n$ and $(y_n)_n$ respectively, on the linear subspace $E_k = \text{span}\{z_1, \dots, z_k\}^\perp$.

Claim 1: First we will show that $\exists M_0$ such that $\forall n \geq M_0$, $a_n \neq 0$ or $b_n \neq 0$. Assuming otherwise, there are two sequences of vectors $x_n = \sum_{r=1}^k c_{r,n} z_r$ and $y_n = \sum_{r=1}^k d_{r,n} z_r$, where $\lim_{n \rightarrow \infty} c_{r,n} = \lim_{n \rightarrow \infty} d_{r,n} = 1$, $\forall r \in [k]$ that achieve lower Lipschitz bound zero. Recall that from part (2) of Lemma 2.8 we have that $\exists M_0 \in \mathbb{N}$ such that $\forall n \geq M_0$ and $(i, j) \in S$

$$L^{i,j}\left(\sum_{r=1}^k c_{r,n} z_r\right) = L^{i,j}\left(\sum_{r=1}^k d_{r,n} z_r\right) = L^{i,j}\left(\sum_{r=1}^k z_r\right).$$

Then, for $g_{i,j} \in L^{i,j}(\sum_{r=1}^k z_r)$,

$$\begin{aligned}
0 &= \lim_{n \rightarrow \infty} \frac{\|\Phi_{\mathbf{w},S}(x_n) - \Phi_{\mathbf{w},S}(y_n)\|^2}{\mathbf{d}(x_n, y_n)^2} \\
&= \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^p \sum_{j \in S_i} |\langle U_{g_{i,j}} w_i, x_n - y_n \rangle|^2}{\mathbf{d}(x_n, y_n)^2} \\
&\geq \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^p \sum_{j \in S_i} |\langle U_{g_{i,j}} w_i, \sum_{r=1}^k (c_{r,n} - d_{r,n}) z_r \rangle|^2}{\|\sum_{r=1}^k (c_{r,n} - d_{r,n}) z_r\|^2} \\
&= \sum_{i=1}^p \sum_{j \in S_i} |\langle U_{g_{i,j}} w_i, \tilde{z} \rangle|^2,
\end{aligned}$$

where

$$\tilde{z} = \lim_{m \rightarrow \infty} \frac{\sum_{r=1}^k (c_{r,n_m} - d_{r,n_m}) z_r}{\|\sum_{r=1}^k (c_{r,n_m} - d_{r,n_m}) z_r\|}$$

is a unit vector obtained as the limit of a convergent subsequence of the sequence of unit vectors $\frac{\sum_{r=1}^k (c_{r,n} - d_{r,n}) z_r}{\|\sum_{r=1}^k (c_{r,n} - d_{r,n}) z_r\|}$. Since the group G is finite, we can find a positive number $\epsilon > 0$ such that $\epsilon \|\tilde{z}\| < \frac{1}{4} \Delta(\sum_{r=1}^k z_r)$ and $\sum_{r=1}^k z_r \approx \sum_{r=1}^k z_r + \epsilon \tilde{z}$. In this case

$$\Phi_{\mathbf{w},S}\left(\sum_{r=1}^k z_r\right) = \Phi_{\mathbf{w},S}\left(\sum_{r=1}^k z_r + \epsilon \tilde{z}\right)$$

which contradict the injectivity property. This establishes Claim 1.

Now we can assume for all $n \geq M_0$, $a_n = P_{E_k} x_n \neq 0$ or $b_n = P_{E_k} y_n \neq 0$. If need be, pass to a subsequence and/or switch the definitions of x_n and y_n , so that $\|b_n\| \geq \|a_n\|$ for all n . In doing so we no longer claim the normalization (15). Nevertheless, both $\|x_n\|, \|y_n\| \leq 1$.

Let $c_{r,n}, d_{r,n}$ be the unique coefficients determined by $x_n = \sum_{r=1}^k c_{r,n} z_r + a_n$, $y_n = \sum_{r=1}^k d_{r,n} z_r + b_n$. Note $\lim_{n \rightarrow \infty} c_{r,n} = \lim_{n \rightarrow \infty} d_{r,n} = 1$.

Let $e_n = \sum_{r=1}^k (d_{r,n} - c_{r,n}) z_r + b_n$ and

$$s_n = \frac{\min\left(\|z_k\|, \Delta(\sum_{r=1}^k z_r), \rho_0(\sum_{r=1}^k z_r)\right)}{16\|e_n\|}.$$

Note $\|e_n\| \geq \|b_n\| \geq \|a_n\|$ for all n .

Claim 2: Sequences $\tilde{x}_n = \sum_{r=1}^k z_r + s_n a_n$ and $\tilde{y}_n = \sum_{r=1}^k z_r + s_n e_n$ achieve also the lower Lipschitz constant zero at $\sum_{r=1}^k z_r$.

Note that $\max(\|s_n a_n\|, \|s_n e_n\|) \leq \frac{1}{16}$. Pass to subsequences of $(a_n)_n$ and $(e_n)_n$ so that both $\lim_{n \rightarrow \infty} s_n a_n$ and $\lim_{n \rightarrow \infty} s_n e_n$ converge. Let $\alpha = \lim_{n \rightarrow \infty} s_n a_n$ and $\delta = \lim_{n \rightarrow \infty} s_n e_n$. Notice $\delta \neq 0$.

The limits

$$\lim_{n \rightarrow \infty} c_{r,n} = \lim_{n \rightarrow \infty} d_{r,n} = 1, \quad \forall r \in [k] \text{ and } \lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} e_n = 0$$

imply that $\exists m_0 \in \mathbb{N}$ such that $\forall n \geq m_0$, and $\forall r \in [k]$

- (a) $|1 - c_{r,n}| < \frac{1}{16k} \Delta(\sum_{r=1}^k z_r)$
- (b) $|1 - d_{r,n}| < \frac{1}{16k} \Delta(\sum_{r=1}^k z_r)$
- (c) $|c_{r,n} - d_{r,n}| < \frac{1}{16k} \Delta(\sum_{r=1}^k z_r)$
- (d) $\|a_n\| < \frac{1}{16k} \Delta(\sum_{r=1}^k z_r)$
- (e) $\|e_n\| < \frac{1}{16k} \Delta(\sum_{r=1}^k z_r)$

From Lemma 2.8 part (1),

$$\Delta\left(\sum_{r=1}^k c_{r,n} z_r\right) \geq \frac{1}{4} \Delta\left(\sum_{r=1}^k z_r\right).$$

Also

$$\max(\|a_n\|, \|s_n a_n\|) < \frac{1}{16} \Delta\left(\sum_{r=1}^k z_r\right) \leq \frac{1}{4} \Delta\left(\sum_{r=1}^k c_{r,n} z_r\right)$$

and

$$\max(\|e_n\|, \|s_n e_n\|) < \frac{1}{16} \Delta\left(\sum_{r=1}^k z_r\right) \leq \frac{1}{4} \Delta\left(\sum_{r=1}^k d_{r,n} z_r\right)$$

So, for any $(i, j) \in S$

$$\begin{aligned} L^{i,j}(x_n) &= L^{i,j}\left(\sum_{r=1}^k c_{r,n} z_r + a_n\right) = L^{i,j}\left(\sum_{r=1}^k z_r + a_n\right) \\ &= L^{i,j}\left(\sum_{r=1}^k z_r + s_n a_n\right) = L^{i,j}(\tilde{x}_n) \subset L^{i,j}\left(\sum_{r=1}^k z_r\right) = L^{i,j}\left(\sum_{r=1}^k c_{r,n} z_r\right). \end{aligned}$$

Where the second equality comes from Lemma 2.8 part 3, third equality from Lemma 2.6 the fifth inclusion from Lemma 2.5, and the last equality from Lemma 2.8 part 2.

Similarly,

$$\begin{aligned} L^{i,j}(y_n) &= L^{i,j}\left(\sum_{r=1}^k d_{r,n}z_r + b_n\right) = L^{i,j}\left(\sum_{r=1}^k (1 + d_{r,n} - c_{r,n})z_r + b_n\right) \\ &= L^{i,j}\left(\sum_{r=1}^k z_r + e_n\right) = L^{i,j}\left(\sum_{r=1}^k z_r + s_n e_n\right) = L^{i,j}(\tilde{y}_n) \subset L^{i,j}\left(\sum_{r=1}^k z_r\right) = L^{i,j}\left(\sum_{r=1}^k c_{r,n}z_r\right). \end{aligned}$$

Therefore,

$$\begin{aligned} 0 &= \lim_{n \rightarrow \infty} \frac{\|\Phi_{\mathbf{w},S}(x_n) - \Phi_{\mathbf{w},S}(y_n)\|^2}{\mathbf{d}(x_n, y_n)^2} = \\ &= \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^p \sum_{j \in S_i} \langle w_i, U_{g_{1,i,j}^{-1}} a_n - U_{g_{2,i,j}^{-1}} e_n \rangle^2}{\|a_n - e_n\|^2} \\ &= \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^p \sum_{j \in S_i} \langle w_i, U_{g_{1,i,j}^{-1}} s_n a_n - U_{g_{2,i,j}^{-1}} s_n e_n \rangle^2}{\|s_n a_n - s_n e_n\|^2} \\ &= \lim_{n \rightarrow \infty} \frac{\|\Phi_{\mathbf{w},S}(\tilde{x}_n) - \Phi_{\mathbf{w},S}(\tilde{y}_n)\|^2}{\mathbf{d}(\tilde{x}_n, \tilde{y}_n)^2}. \end{aligned}$$

where

$$g_{1,i,j} \in L^{i,j}(x_n) \text{ and } g_{2,i,j} \in L^{i,j}(y_n)$$

are chosen independent of n by possibly passing to subsequences since G is finite. So,

$$\Phi_{\mathbf{w},S}\left(\sum_{r=1}^k z_r + \alpha\right) - \Phi_{\mathbf{w},S}\left(\sum_{r=1}^k z_r + \delta\right) = 0.$$

Since $\Phi_{\mathbf{w},S}$ is injective,

$$\sum_{r=1}^k z_r + \alpha \sim \sum_{r=1}^k z_r + \delta$$

Let $g_1 \in G$ denote a group element that achieves this equivalence, i.e.

$$\sum_{r=1}^k z_r + \alpha = U_{g_1}\left(\sum_{r=1}^k z_r + \delta\right)$$

Note that $g_1 \in H(\sum_{r=1}^k z_r)$ because otherwise

$$\begin{aligned} 0 &= \left\| \sum_{r=1}^k z_r + \alpha - U_{g_1} \left(\sum_{r=1}^k z_r \right) + U_{g_1} \delta \right\| = \left\| \sum_{r=1}^k z_r + \alpha - U_{g_1} \left(\sum_{r=1}^k z_r \right) + U_{g_1} \delta \right\| \\ &\geq \left\| \sum_{r=1}^k z_r - U_{g_1} \left(\sum_{r=1}^k z_r \right) \right\| - \|\alpha - U_{g_1} \delta\| \geq \rho_0 \left(\sum_{r=1}^k z_r \right) - \|\alpha\| - \|\delta\| > 0 \end{aligned}$$

The last inequality comes from the fact that $\|\alpha\| < \frac{1}{4} \rho_0(\sum_{r=1}^k z_r)$, and $\|\delta\| < \frac{1}{4} \rho_0(\sum_{r=1}^k z_r)$.

Additionally, $\alpha = U_{g_1} \delta$ because

$$0 = \left\| \sum_{r=1}^k z_r + a - U_{g_1} \left(\sum_{r=1}^k z_r \right) + U_{g_1} \delta \right\| = \|a - U_{g_1} \delta\|.$$

Claim 3: The two vectors α and δ are equal, $\alpha = \delta$.

We prove this claim by contradiction. Assume that $\alpha \neq \delta$. From Lemma 2.10, $\exists M_0 \in \mathbb{N}$ such that $\forall n \geq M_0$

$$\|s_n a_n - s_n e_n\| \leq \|s_n a_n - s_n U_{g_1} e_n\|.$$

Therefore,

$$0 < \|a - \delta\| = \lim_{n \rightarrow \infty} \|s_n a_n - s_n e_n\| \leq \lim_{n \rightarrow \infty} \|s_n a_n - s_n U_{g_1} e_n\| = 0.$$

We conclude that $\alpha = \delta \neq 0$.

Set $z_{k+1} = \alpha = \delta$. Together with sequences \tilde{x}_n and \tilde{y}_n , they satisfy the assertions of part 2 of this Lemma. □

Remark 2.14. *Our construction produces z_{k+1} that has norm equal to $\frac{1}{16} \min \left(\|z_k\|, \Delta(\sum_{r=1}^k z_r), \rho_0(\sum_{r=1}^k z_r) \right)$.*

Now we can complete the proof of Theorem 2.1.

Proof of Theorem 2.1. Starting from vector z_1 and the sequences $(x_n)_n$, $(y_n)_n$ observed in Lemma 2.3 after $d - 1$ steps of algorithmic construction of

part (2) of Lemma 2.13 we get d non-zero vectors $\{z_1, \dots, z_d\}$ and a pair of sequences $(\tilde{x}_n)_n, (\tilde{y}_n)_n$ such that

- (i) $\langle z_i, z_j \rangle = 0, \forall i, j \in [d], i \neq j$;
- (ii) $\|z_{k+1}\| \leq \min(\frac{1}{4}\Delta(\sum_{r=1}^k z_r), \frac{1}{4}\|z_k\|), \forall k \in [d-1]$; and
- (iii) $\lim_{n \rightarrow \infty} \tilde{x}_n = \lim_{n \rightarrow \infty} \tilde{y}_n = \sum_{r=1}^d z_r$ and

$$\lim_{n \rightarrow \infty} \frac{\|\Phi_{\mathbf{w}, S}(\tilde{x}_n) - \Phi_{\mathbf{w}, S}(\tilde{y}_n)\|^2}{\mathbf{d}(\tilde{x}_n, \tilde{y}_n)^2} = 0.$$

Let $\tilde{x}_n = \sum_{r=1}^d l_{r,n} z_r$ and $\tilde{y}_n = \sum_{r=1}^d t_{r,n} z_r$. Notice that $\lim_{n \rightarrow \infty} l_{r,n} = \lim_{n \rightarrow \infty} t_{r,n} = 1, \forall r \in [d]$.

Recall that from part (2) of Lemma 2.8 we have that $\exists M_0 \in \mathbb{N}$ such that $\forall n \geq M_0$ and $(i, j) \in S$

$$L^{i,j}(\sum_{r=1}^d l_{r,n} z_r) = L^{i,j}(\sum_{r=1}^d t_{r,n} z_r) = L^{i,j}(\sum_{r=1}^d z_r).$$

Then, for $g_{i,j} \in L^{i,j}(\sum_{r=1}^d z_r)$,

$$\begin{aligned} 0 &= \lim_{n \rightarrow \infty} \frac{\|\Phi_{\mathbf{w}, S}(x_n) - \Phi_{\mathbf{w}, S}(y_n)\|^2}{\mathbf{d}(x_n, y_n)^2} \\ &= \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^p \sum_{j \in S_i} |\langle U_{g_{i,j}} w_i, x_n - y_n \rangle|^2}{\mathbf{d}(x_n, y_n)^2} \\ &\geq \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^p \sum_{j \in S_i} |\langle U_{g_{i,j}} w_i, \sum_{r=1}^d (l_{r,n} - t_{r,n}) z_r \rangle|^2}{\|\sum_{r=1}^d (l_{r,n} - t_{r,n}) z_r\|^2} \\ &= \sum_{i=1}^p \sum_{j \in S_i} |\langle U_{g_{i,j}} w_i, \tilde{z} \rangle|^2, \end{aligned}$$

where

$$\tilde{z} = \lim_{m \rightarrow \infty} \frac{\sum_{r=1}^d (l_{r,n_m} - t_{r,n_m}) z_r}{\|\sum_{r=1}^d (l_{r,n_m} - t_{r,n_m}) z_r\|}$$

is a unit vector obtained as the limit of a convergent subsequence of the sequence of unit vectors $\frac{\sum_{r=1}^d (l_{r,n} - t_{r,n}) z_r}{\|\sum_{r=1}^d (l_{r,n} - t_{r,n}) z_r\|}$. Since the group G is finite, we can find a positive number $\epsilon > 0$ such that $\epsilon \|\tilde{z}\| < \frac{1}{4}\Delta(\sum_{r=1}^d z_r)$ and $\sum_{r=1}^d z_r \approx \sum_{r=1}^d z_r + \epsilon \tilde{z}$. In this case

$$\Phi_{\mathbf{w}, S}(\sum_{r=1}^d z_r) = \Phi_{\mathbf{w}, S}(\sum_{r=1}^d z_r + \epsilon \tilde{z})$$

which contradict the injectivity property. Theorem 2.1 is now proved. \square

3 Dimension reduction using linear projections

Note that the vector $\Phi_{\mathbf{w},S}$ can have more than $2d$ entries. In this section we show that, no matter how big the dimension m of the target space of an injective embedding $\Phi_{\mathbf{w},S}$ is, a generic linear projection of $\Phi_{\mathbf{w},S}$ onto a linear space of dimension twice the dimension of the data space, preserves injectivity (similar to [7]).

Theorem 3.1. *Let G be a finite group of order N acting unitarily on \mathcal{V} and $\mathbf{w} \in \mathcal{V}^p$, $S \subset [N] \times [p]$ so that $\Phi_{\mathbf{w},S} : \mathcal{V} \rightarrow \mathbb{R}^m$ is injective on the quotient space $\hat{\mathcal{V}}$. Then for a generic linear transformation $\ell : \mathbb{R}^m \rightarrow \mathbb{R}^{2d}$, the map $\Psi_{\mathbf{w},S,\ell} = \ell \circ \Phi_{\mathbf{w},S}$ is also injective. Here generic means open dense with respect to Zariski topology over the set of matrices.*

Let $D : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}^m$ be the map $D(x, y) = \Phi_{\mathbf{w},S}(x) - \Phi_{\mathbf{w},S}(y)$. Its range E is defined by $E = \text{Ran}(D) = \{\Phi_{\mathbf{w},S}(x) - \Phi_{\mathbf{w},S}(y) : x, y \in \mathcal{V}\} = \text{Ran}(\Phi_{\mathbf{w},S}) - \text{Ran}(\Phi_{\mathbf{w},S})$.

Let g_1, \dots, g_N be an enumeration of the elements of the group G and define $\lambda_{i,j}(x) : \mathcal{V} \rightarrow \mathbb{R}$ where, $\lambda_{i,j}(x) = \langle U_{g_i} w_j, x \rangle$. Notice that $\lambda_{i,j}$ is a linear map (unlike $\Phi_{i,j}$) and also that

$$\begin{aligned} \Phi_{\mathbf{w},S}(x) - \Phi_{\mathbf{w},S}(y) = & [\lambda_{1,\pi_1(1)}(x) - \lambda_{1,\pi_{p+1}(1)}(y), \dots, \lambda_{1,\pi_1(m_1)}(x) - \lambda_{1,\pi_{p+1}(m_1)}(y), \\ & \dots, \lambda_{p,\pi_p(1)}(x) - \lambda_{p,\pi_{2p}(1)}(y), \dots, \lambda_{p,\pi_p(m_p)}(x) - \lambda_{p,\pi_{2p}(m_p)}(y)] \end{aligned}$$

for some $\pi_1, \dots, \pi_{2p} \in S_N$ that depend on x and y . Let $m_j = |S_j| = \{i \in [N], (i, j) \in S\}$ so that $m_1 + \dots + m_p = m$.

Now, fix permutations $\pi_1, \dots, \pi_{2p} \in S_N$ and let $L_{\pi_1, \dots, \pi_{2p}} : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}^m$ denote the *linear* maps

$$\begin{aligned} L_{\pi_1, \dots, \pi_{2p}}(x, y) = & [\lambda_{1,\pi_1(1)}(x) - \lambda_{1,\pi_{p+1}(1)}(y), \dots, \lambda_{1,\pi_1(m_1)}(x) - \lambda_{1,\pi_{p+1}(m_1)}(y), \\ & \dots, \lambda_{p,\pi_p(1)}(x) - \lambda_{p,\pi_{2p}(1)}(y), \dots, \lambda_{p,\pi_p(m_p)}(x) - \lambda_{p,\pi_{2p}(m_p)}(y)] \end{aligned}$$

Also, define

$$F = \cup_{\pi_1, \dots, \pi_{2p} \in S_N} \text{Ran}(L_{\pi_1, \dots, \pi_{2p}}).$$

Notice that F is a finite union of linear subspaces and $E \subset F$. For fixed π_1, \dots, π_{2p} the map $(x, y) \mapsto L_{\pi_1, \dots, \pi_{2p}}(x, y)$ is linear in (x, y) and from the rank-nullity Theorem we have

$$\dim(\text{Ran}(L_{\pi_1, \dots, \pi_{2p}})) \leq 2d.$$

Lemma 3.2. *Assume r, s, M are non-negative integers so that $r + s \leq M$. For any finite collection $\{F_a : a \in [T]\}$, of linear subspaces of dimension at most s , a generic r -dimensional linear subspace K of \mathbb{R}^M , satisfies $K \cap F_a = \{0\}$, $\forall a \in [T]$. Here generic means open and dense with respect to Zarisky topology.*

Proof. Let $\{v_1, \dots, v_r\}$ be a spanning set for K , and $\{w_{1,a}, \dots, w_{M-r,a}\}$ be a linearly independent set of vectors such that $F_a \subset \text{span}\{w_{1,a}, \dots, w_{M-r,a}\}$. Then, $\text{span}\{v_1, \dots, v_r\} \cap \text{span}\{w_{1,a}, \dots, w_{M-r,a}\} = \{0\}$ if, and only if, the set $\{v_1, \dots, v_r, w_{1,a}, \dots, w_{M-r,a}\}$ is linearly independent. Define $R_a(v_1, \dots, v_r) = \det[v_1 | \dots | v_r | w_{1,a} | \dots | w_{M-r,a}]$, and note that $R_a(v_1, \dots, v_r)$ is a polynomial in variables $v_1(1), \dots, v_1(M), \dots, v_r(1), \dots, v_r(M)$. Hence,

$$\begin{aligned} K \cap F_a = \{0\}, \quad \forall a \in [N] &\iff R_a(v_1, \dots, v_r) \neq 0, \quad \forall a \in [N] \\ &\iff \prod_{a=1}^N R_a(v_1, \dots, v_r) \neq 0. \end{aligned}$$

We conclude that

$$\mathbb{U} = \left\{ (v_1, \dots, v_r) : \prod_{a=1}^N R_a(v_1, \dots, v_r) \neq 0 \right\}$$

is an open set with respect to Zariski topology. In order to show that \mathbb{U} is generic we have to find a set $\{v_1, \dots, v_r\}$ such that $\prod_a R_a(v_1, \dots, v_r) \neq 0$.

Let $W_a = \text{span}\{w_{1,a}, \dots, w_{M-r,a}\}$. Notice that $\text{span}(w_{1,a}, \dots, w_{M-r,a})$ are linear subspaces of \mathbb{R}^M each of dimension $M - r$. If $r \geq 1$, each W_a is a proper subspace of \mathbb{R}^M . For a generic $v_1 \neq 0$, $v_1 \notin \cup_{a=1}^N W_a$, and replace W_a with $W_a^1 = \text{span}(W_a, \{v_1\})$. Notice that $\dim(W_a^1) = \dim(W_a) + 1$. If $\dim(W_a^1) < M$, repeat this process and obtain v_2, \dots, v_r until $\dim(W_a^r) = M$. The procedure produces a set of vectors (v_1, \dots, v_r) that satisfy the condition $\prod_a R_a(v_1, \dots, v_r) \neq 0$. This ends the proof of Lemma 3.2. \square

Now we apply this lemma to derive the following corollary for our setup:

Corollary 3.3. *Let $L_{\pi_1, \dots, \pi_{2p}} : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}^m$ be the $(N!)^{2p}$ linear maps introduced before. Then for a generic $\ell : \mathbb{R}^m \rightarrow \mathbb{R}^{2d}$,*

$$\ker(\ell) \cap \bigcup_{\pi_1, \dots, \pi_{2p} \in S_N} \text{Ran}(L_{\pi_1, \dots, \pi_{2p}}) = \{0\}$$

Proof. If $m \leq 2d$ then the conclusion is satisfied for any full-rank ℓ . Therefore assume $m > 2d$. A generic linear map $\ell : \mathbb{R}^m \rightarrow \mathbb{R}^{2d}$ is full-rank. Hence $\dim(\text{Ran}(\ell)) = m$, and thus $\dim(\ker(\ell)) = m - 2d$. On the other hand, for a generic linear map ℓ Lemma 3.2 implies

$$\ker(\ell) \cap \text{Ran}(L_{\pi_1, \dots, \pi_{2p}}) = \{0\}$$

for every $\pi_1, \dots, \pi_{2p} \in S_N$. □

Now we can prove Theorem 3.1.

Proof of Theorem 3.1. Corollary 3.3 shows that, a generic linear map $\ell : \mathbb{R}^m \rightarrow \mathbb{R}^{2d}$ satisfies $\ker(\ell) \cap \text{Ran}(D) = \{0\}$. Thus, if $x, y \in \mathcal{V}$ so that $\Psi_{\mathbf{w}, S, \ell}(x) = \Psi_{\mathbf{w}, S, \ell}(y)$ then $\ell(D(x, y)) = 0$. Therefore $D(x, y) = 0$. Since $\Phi_{\mathbf{w}, S}$ is injective it follows $x \sim y$. Thus, $\Psi_{\mathbf{w}, S, \ell} = \ell \circ \Phi_{\mathbf{w}, S}$ is injective on the quotient space. □

The next result establishes the bi-Lipschitz property of the map $\Psi_{\mathbf{w}, S, \ell}$.

Theorem 3.4. *Let $S \subset [N] \times [p]$ and $\mathbf{w} \in \mathcal{V}^p$. Suppose that $\Phi_{\mathbf{w}, S} : \mathcal{V} \rightarrow \mathbb{R}^m$ defined in (2) is injective on the quotient space $\hat{\mathcal{V}}$. Then, for a generic linear map $\ell : \mathbb{R}^m \rightarrow \mathbb{R}^{2d}$, the map $\Psi_{\mathbf{w}, S, \ell} = \ell \circ \Phi_{\mathbf{w}, S}$ is injective and bi-Lipschitz. Here generic means open dense with respect to Zariski topology over the set of matrices.*

The Lipschitz property of $\Psi_{\mathbf{w}, S, \ell}$ is trivial since a composition of Lipschitz maps is Lipschitz and we establish that $\Phi_{\mathbf{w}, S, \ell}$ is Lipschitz. The non-trivial part is to show the lower bound. To do so we need the following lemma.

Lemma 3.5. *Let $\{F_a\}_{a=1}^T$ be a finite collection of r -dimensional subspaces of \mathbb{R}^M , and $\ell : \mathbb{R}^M \rightarrow \mathbb{R}^m$ be a full-rank linear transformation with $M \geq m$. Let Q_a denote the orthogonal projection onto the linear space F_a and Q_ℓ the orthogonal projection onto $\ker \ell$. Let $c_{a, \ell} = (1 - \|Q_a Q_\ell\|^2)^{1/2}$, and $c_\ell = \min_{a \in [T]} c_{a, \ell}$. Set $F = \bigcup_{a=1}^T F_a$. Suppose that $\ker(\ell) \cap F = \{0\}$. Then*

$$\inf_{\substack{x \in F \\ \|x\|=1}} \|\ell(x)\| \geq c_\ell \sigma_m(\ell), \quad (21)$$

where $\sigma_m(\ell)$ is the smallest strictly positive singular value of ℓ (it is the m^{th} singular value).

Proof. Notice that for each $a \in [T]$, the unit sphere of F_a is a compact set. Thus

$$\inf_{\substack{x \in F \\ \|x\|=1}} \|\ell(x)\| = \min_{\substack{x \in F \\ \|x\|=1}} \|\ell(x)\| = \|\ell(y_\infty)\|$$

for some $y_\infty \in F_a \cap S^1(\mathbb{R}^M)$. Let $y_\infty = \sum_{k=1}^M c_k u_k$, where u_j are the normalized right singular vectors of ℓ sorted by singular values $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_m \geq \sigma_{m+1} = \dots = \sigma_M = 0$. Notice that $\sum_{k=1}^M c_k^2 = 1$ and $\sum_{k=1}^m c_k^2 = 1 - \|Q_\ell y_\infty\|^2 \geq 1 - \|Q_a Q_\ell\|^2 \geq c_\ell^2$. Thus

$$\begin{aligned} \|\ell(y_\infty)\|^2 &= \left\| \sum_{k=1}^M c_k \ell(u_k) \right\|^2 = \left\| \sum_{k=1}^m c_k \ell(u_k) \right\|^2 \\ &= \sum_{k=1}^m c_k^2 \sigma_k^2 \geq c_{a,\ell}^2 \sigma_m(\ell)^2 \geq c_\ell^2 \sigma_m(\ell)^2 \end{aligned}$$

which proves this Lemma. \square

Proof Theorem 3.4. Assume without loss of generality that $m \geq 2d$. From Theorem 3.1 we have that, if $\Phi_{\mathbf{w},S}$ is injective then for a generic linear map $\ell : \mathbb{R}^m \rightarrow \mathbb{R}^{2d}$ the map $\Psi_{\mathbf{w},S,\ell} = \ell \circ \Phi_{\mathbf{w},S}$ is injective.

From Theorem 2.1 we have that, if the map $\Phi_{\mathbf{w},S}$ is injective then it is also bi-Lipschitz. Let $a \leq b$ denote its bi-Lipschitz constants.

Compositions of two Lipschitz maps is Lipschitz, hence $\Psi_{\mathbf{w},S,\ell}$ is Lipschitz. Furthermore, an upper Lipschitz constant of $\Psi_{\mathbf{w},S,\ell}$ is $\|\ell\|b$, where $\|\ell\| = \sigma_1(\ell)$ is the largest singular value of ℓ .

Finally from Corollary 3.3 and Lemma 3.5 we have that for a generic linear map ℓ , for all $x, y \in \mathcal{V}$,

$$\|\Psi_{\mathbf{w},S,\ell}(x) - \Psi_{\mathbf{w},S,\ell}(y)\| = \|\ell(D(x, y))\| \geq c_\ell \sigma_{2d}(\ell) \|D(x, y)\| \geq c_\ell \sigma_{2d}(\ell) a \mathbf{d}(x, y)$$

where a is the lower Lipschitz constant of $\Phi_{\mathbf{w},S}$. Therefore the map $\Psi_{\mathbf{w},S,\ell}$ is bi-Lipschitz with a lower Lipschitz constant $c_\ell \sigma_{2d}(\ell) a$. \square

Remark 3.6. We did prove that any linear map $\ell : \mathbb{R}^m \rightarrow \mathbb{R}^{2d}$ so that $\Psi_{\mathbf{w},S,\ell}$ is injective makes $\Psi_{\mathbf{w},S,\ell}$ bi-Lipschitz. It remains an open question whether for any such nonlinear embedding $\Psi_{\mathbf{w},S,\ell}$ if it is injective it is automatically bi-Lipschitz. In general, if the map $f : X \rightarrow Y$ is bi-Lipschitz and the linear map $\ell : Y \rightarrow \mathbb{R}^q$ is so that $\ell \circ f$ is injective, then $\ell \circ f$ may not be bi-Lipschitz. Example: $f : \mathbb{R} \rightarrow \mathbb{R}^2$, $f(t) = (t, t^3)$, $\ell : \mathbb{R}^2 \rightarrow \mathbb{R}$, $\ell(x, y) = y$.

References

- [1] A.S. Bandeira, J. Cahill, D. Mixon, A.A. Nelson. “Saving phase: Injectivity and Stability for phase retrieval”. In: *Appl. Comp. Harm. Anal.* 37.1 (2014), pp. 106–125.
- [2] B. Alexeev, J. Cahill, and Dustin G. Mixon. “Full Spark Frames”. In: *J. Fourier Anal. Appl* 18 (2012), pp. 1167–1194.
- [3] Benjamin Aslan, Daniel Platt, and David Sheard. “Group invariant machine learning by fundamental domain projections”. In: *NeurIPS Workshop on Symmetry and Geometry in Neural Representations*. PMLR, 2023, pp. 181–218.
- [4] R. Balan. “Frames and Phaseless Reconstruction”. In: vol. Finite Frame Theory: A Complete Introduction to Overcompleteness. Proceedings of Symposia in Applied Mathematics 73. AMS Short Course at the Joint Mathematics Meetings, San Antonio, January 2015 (Ed. K.Okoudjou), 2016, pp. 175–199.
- [5] R. Balan and Y. Wang. “Invertibility and robustness of phaseless reconstruction”. In: *Applied and Comput. Harmon. Analysis* 38.3 (2015), pp. 469–488.
- [6] R. Balan and D. Zou. “On Lipschitz Analysis and Lipschitz Synthesis for the Phase Retrieval Problem”. In: *Linear Algebra and Applications* 496 (2016), pp. 152–181.
- [7] Radu Balan, Naveed Haghani, and Maneesh Singh. “Permutation Invariant Representations with Applications to Graph Deep Learning”. In: *arXiv preprint arXiv:2203.07546* (2022).
- [8] M.M. Bronstein et al. “Geometric Deep Learning: Going Beyond Euclidean Data”. In: *IEEE Signal Processing Magazine* 34.4 (2017), pp. 18–42.
- [9] Jameson Cahill et al. “Group-invariant max filtering”. In: *arXiv:2205.14039 [cs.IT]* (2022), pp. 1–35.
- [10] Emilie Dufresne. “Separating invariants and finite reflection groups”. In: *Advances in Mathematics* 221.6 (2009), pp. 1979–1989. ISSN: 0001-8708. DOI: <https://doi.org/10.1016/j.aim.2009.03.013>. URL: <https://www.sciencedirect.com/science/article/pii/S0001870809330133>.

- [11] Nadav Dym and Steven J Gortler. “Low Dimensional Invariant Embeddings for Universal Geometric Learning”. In: *arXiv preprint arXiv:2205.02956* (2022).
- [12] G. Kemper H. Derksen. *Computational Invariant Theory*. Springer, 2002.
- [13] Morris Hirsch. *Differential Topology*. Springer, 1994.
- [14] A.C. Hip J. Cahill A. Contreras. “Complete Set of translation Invariant Measurements with Lipschitz Bounds”. In: *Appl. Comput. Harm. Anal.* 49.2 (2020), pp. 521–539.
- [15] Martin Larocca et al. “Group-invariant quantum machine learning”. In: *PRX Quantum* 3.3 (2022), p. 030341.
- [16] Haggai Maron et al. “Invariant and Equivariant Graph Networks”. In: *International Conference on Learning Representations*. 2019. URL: <https://openreview.net/forum?id=Syx72jC9tm>.
- [17] Haggai Maron et al. “On the Universality of Invariant Networks”. In: *Proceedings of the 36th International Conference on Machine Learning*. Ed. by Kamalika Chaudhuri and Ruslan Salakhutdinov. Vol. 97. Proceedings of Machine Learning Research. PMLR, June 2019, pp. 4363–4371. URL: <https://proceedings.mlr.press/v97/maron19a.html>.
- [18] Dustin G Mixon and Daniel Packer. “Max filtering with reflection groups”. In: *arXiv preprint arXiv:2212.05104* (2022).
- [19] Dustin G Mixon and Yousef Qaddura. “Injectivity, stability, and positive definiteness of max filtering”. In: *arXiv preprint arXiv:2212.11156* (2022).
- [20] Omri Puny et al. “Frame Averaging for Invariant and Equivariant Network Design”. In: *International Conference on Learning Representations*. 2022. URL: <https://openreview.net/pdf?id=zIUyj55nXR>.
- [21] Akiyoshi Sannai, Yuuki Takai, and Matthieu Cordonnier. *Universal approximations of permutation invariant/equivariant functions by deep neural networks*. 2020. URL: <https://openreview.net/forum?id=HkeZQJBKDB>.
- [22] Dmitry Yarotsky. “Universal approximations of invariant maps by neural networks”. In: *Constructive Approximation* (2021), pp. 1–68.