

Article

Not peer-reviewed version

A Lifting Framework for Constructing Group-Invariant Embeddings

[Guorui Chen](#) *

Posted Date: 6 April 2026

doi: 10.20944/preprints202604.0336.v1

Keywords: group invariance; orbit separation; bi-Lipschitz embeddings; phase retrieval



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

A Lifting Framework for Constructing Group-Invariant Embeddings

Guorui Chen ^{1,2} 

¹ State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190, China; chengurui@lsec.cc.ac.cn

² School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing, 100049, China

Abstract

We study the problem of constructing group-invariant embeddings that faithfully represent data modulo group symmetries, a task that arises naturally in signal processing, physics, and machine learning. A central challenge is to design embeddings that are simultaneously orbit-separating, stable, and computationally tractable. To address this, we develop a general lifting framework for constructing such embeddings. The key idea is to start from a group-invariant embedding defined on a low-dimensional reduced space, and lift it to the ambient space by composing it with a finite family of parameterized linear maps, followed by an aggregation step that produces a global embedding. This framework provides a unified perspective that connects classical problems such as phase retrieval and permutation-invariant embeddings. We demonstrate the effectiveness of this framework in the finite group setting. In this setting, we establish general sufficient conditions for orbit separation and prove that any orbit-separating lifting embedding is automatically bi-Lipschitz. We further extend the bi-Lipschitz result to sparse regimes, and show that, when applied to phase retrieval, it yields an equivalence between uniqueness and stability for real sparse phase retrieval.

Keywords: group invariance; orbit separation; bi-Lipschitz embeddings; phase retrieval

1. Introduction

Data in many areas exhibit intrinsic symmetries that can be naturally described by group actions, where certain transformations of the input do not alter its essential properties or meaning. For instance, in signal processing, complex-valued signals are invariant under global phase shifts, which preserve the amplitude spectrum [1–4]. In image analysis, semantic content is typically invariant under rotations and reflections [5–9]. In 3D geometry and computer vision, point clouds and shapes are invariant under rigid motions described by the Euclidean group $E(3)$ [10–12]. In molecular and physical systems, the potential energy of a molecule is invariant under permutations of identical atoms, as the labeling of atoms carries no physical meaning [13–15].

Across these domains, the central challenge is to represent data in a form that respects the underlying symmetries. The presence of a group action typically introduces substantial redundancy in the ambient space, as multiple elements related by the group action correspond to the same underlying object. As a result, directly performing analysis or learning on the raw data may lead to suboptimal performance or distorted geometric interpretations, since standard notions of distance and similarity fail to account for this equivalence structure. This observation suggests that the true objects of interest are not individual data points themselves, but rather their equivalence classes under the group action. From a mathematical perspective, this setting can be described by a vector space $\mathcal{V}$ over $\mathbb{F}$, where $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, equipped with a group action $G \curvearrowright \mathcal{V}$. Two elements $\mathbf{v}, \mathbf{v}' \in \mathcal{V}$ are considered equivalent if $\mathbf{v}' = g \cdot \mathbf{v}$ for some $g \in G$. The appropriate object of interest is therefore the quotient space $\mathcal{V}/G$, whose elements correspond to group orbits.

While the quotient space $\mathcal{V}/G$ captures the intrinsic structure of the data, it is not directly amenable to computation or learning. It lacks a natural vector space structure and admits no canonical

coordinate representation, rendering it incompatible with standard computational tools—such as linear operations and vector-valued architectures—that underpin most modern machine learning methods. To address this difficulty, one seeks a group-invariant embedding

$$F : \mathcal{V} \rightarrow \mathbb{F}^d, \quad F(g \cdot \mathbf{v}) = F(\mathbf{v}) \text{ for all } g \in G,$$

which serves as a tractable representation of the orbit space within a Euclidean framework. For the embedding to faithfully encode the quotient structure, it must not only be group-invariant but also orbit-separating, that is, distinct orbits map to distinct representations. In addition, to preserve geometric structure and guarantee stability to perturbations, it is desirable for F to be bi-Lipschitz on the orbit space. Finally, practical and theoretical considerations both favor embeddings of low target dimension d , as they provide more compact representations and reduce redundancy. The construction of group-invariant embeddings that are orbit-separating and stable has become an active area of research at the interface of harmonic analysis, geometry, and machine learning [16–23].

A classical approach to constructing group-invariant embeddings relies on *invariant theory*. Let $G \curvearrowright \mathcal{V}$ and consider the ring of invariant polynomials

$$\mathbb{F}[\mathcal{V}]^G = \{p \in \mathbb{F}[\mathcal{V}] : p(g \cdot \mathbf{v}) = p(\mathbf{v}), \forall g \in G, \mathbf{v} \in \mathcal{V}\}.$$

By the Hilbert’s finiteness theorem [24], when G is reductive¹, there exists a finite set of invariant polynomials $\{p_i\}_{i=1}^d$ generating $\mathbb{F}[\mathcal{V}]^G$. The corresponding polynomial map $F(\mathbf{v}) = (p_1(\mathbf{v}), \dots, p_d(\mathbf{v}))$ thus provides a canonical G -invariant embedding. However, identifying such a generating set is often challenging, and even when it can be found, the number d may be prohibitively large. This motivates the study of *separating invariants* [25–29], which aim only to distinguish orbits without generating the full invariant ring. Separating sets are typically easier to construct and often suffice for embedding purposes. While polynomial invariants possess rich algebraic structure, they often lack geometric stability. In particular, polynomially defined group-invariant embeddings generally fail to satisfy Lipschitz continuity. To address this, Cahill, Contreras, and Contreras Hip [30,31] introduced a homogenization technique that modifies polynomial invariants to yield Lipschitz group-invariant embeddings for certain group actions. This technique was later generalized in [32] into a framework that produces a bi-Lipschitz group-invariant embedding on the entire space, in which one first constructs a bi-Lipschitz group-invariant embedding on the unit sphere $\mathbb{S}(\mathcal{V})$ and then extends it to $\mathcal{V}$ via homogenization. However, when restricted to polynomial invariants, achieving bi-Lipschitz embeddings on the sphere requires the group G to act freely² on $\mathbb{S}(\mathcal{V})$, indicating that such polynomial constructions are subject to stringent structural conditions. All the above approaches remain within the polynomial framework. To move beyond it, alternative non-polynomial constructions have been proposed. Dym and Gortler [33] proposed a construction using real semi-algebraic functions that achieves orbit separation with a low target dimension $d = 2 \dim_{\mathbb{R}}(\mathcal{V}) + 1$, though its bi-Lipschitz property remains open. A more recent and conceptually innovative line of work is based on the *max filters*, introduced by Cahill, Iverson, Mixon, and Packer [34]. The max filters form a family of non-polynomial invariants with orbit separation property for general groups and provable bi-Lipschitz behavior in the finite-group setting [35,36]. The max filter framework was later generalized to *coorbit embedding*, which has attracted a growing body of research [37–40]. Although max filters exhibit remarkable theoretical and empirical performance, their computation can become cumbersome in high-dimensional settings. For example, when $\mathcal{V} = \mathbb{R}^{m \times n}$ and the symmetric group $G = S_m$ acts by permuting rows, computing the max filter reduces to a sorting operation for $n = 1$, but requires solving a linear program when $n \geq 2$ (see Section 3.2.1 in [34]).

¹ A reductive group is an algebraic group whose representations are completely reducible.

² A group action $G \curvearrowright \mathcal{X}$ is said to be free if, for any $g \in G$ and $\mathbf{x} \in \mathcal{X}$, $g \cdot \mathbf{x} = \mathbf{x}$ implies g is the identity element. In other words, no nontrivial group element fixes any point in $\mathcal{X}$.

1.1. Our Contributions

The computational tractability of invariants in low-dimensional spaces, as exemplified by max filters, contrasts sharply with the intractability of their high-dimensional counterparts. This disparity motivates a general strategy: instead of constructing invariants directly in high-dimensional spaces, one can exploit efficient constructions on lower-dimensional spaces and subsequently lift them to the ambient space. This paper formalizes this idea into a general lifting framework for building group-invariant embeddings.

Throughout the paper, we focus on the case where $\mathcal{V}$ is a finite-dimensional inner product space over $\mathbb{F}$, and G is a compact group. Our lifting framework starts from a G -invariant embedding f defined on a low-dimensional subspace $\widehat{\mathcal{V}}$, referred to as the *reduced space*. We then introduce a parameterized family of structured linear maps that send elements of $\mathcal{V}$ to $\widehat{\mathcal{V}}$. Composing f with any such linear map yields a G -invariant map on $\mathcal{V}$, which we call an *elementary lifting invariant*. The final lifting group-invariant embedding F is constructed by aggregating a finite collection of these elementary lifting invariants, each associated with a distinct parameter. This framework provides a unified perspective and reveals connections to classical problems such as phase retrieval and permutation-invariant embedding, thereby highlighting its broad applicability.

We establish several results demonstrating the effectiveness of this lifting framework in the finite group setting:

- (1) **Orbit separation:** For a G -invariant semi-algebraic subset $\mathcal{M} \subseteq \mathcal{V}$, we show that a generic collection of $2 \dim_{\mathbb{R}}(\mathcal{M}) / C_{\mathbb{F}} + 1$ elementary lifting invariants suffices to ensure that the resulting embedding is orbit-separating on $\mathcal{M}$, where $C_{\mathbb{F}} = \dim_{\mathbb{R}}(\mathbb{F})$ and $\dim_{\mathbb{R}}(\mathcal{M})$ is the real dimension of $\mathcal{M}$ (see Theorem 1). In particular, taking $\mathcal{M} = \mathcal{V}$ yields that a generic collection of $2D + 1$ elementary lifting invariants is orbit-separating on the whole space, where D is the dimension of $\mathcal{V}$ as an $\mathbb{F}$ -vector space.
- (2) **Refined bounds for orbit separation under structural assumptions:** For $\mathcal{V} = \mathbb{F}^{m \times n}$ with reduced space $\widehat{\mathcal{V}} = \mathbb{F}^m$, we establish a sharper bound, showing that only $mn + m + n$ elementary lifting invariants suffice for orbit separation on the entire space (see Theorem 2). Applied to the permutation-invariant setting, this yields a tighter bound than the $(n - 1)m! + 1$ bound from [41] and the $2mn + 1$ bound from [33].
- (3) **Stability:** We show that any orbit-separating lifting group-invariant embedding is automatically bi-Lipschitz (see Theorem 3). This stability result further extends to sparse regimes, yielding bi-Lipschitz guarantees for both real phase retrieval and permutation-invariant problems under sparsity constraints (see Corollary 4 and Corollary 5).

1.2. Notations

We adopt the following notation throughout this paper. For a complex number $z = z_0 + z_1 i \in \mathbb{C}$, we denote its conjugate by $\bar{z} = z_0 - z_1 i$. For $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, we define $C_{\mathbb{F}} := \dim_{\mathbb{R}}(\mathbb{F})$, that is, $C_{\mathbb{R}} = 1$ and $C_{\mathbb{C}} = 2$. For a matrix $A \in \mathbb{F}^{m \times n}$, we denote by $A^{\top}$ its transpose, by A^* its conjugate transpose, and by $\text{Tr}(A)$ its trace when $m = n$. The identity matrix in $\mathbb{F}^{m \times m}$ is denoted by I_m . The zero matrix in $\mathbb{F}^{m \times n}$ is written as $\mathbf{0}_{m \times n}$, and the boldface symbol $\mathbf{0}$ may be used to denote the corresponding zero element whenever no confusion arises. For vectors $\mathbf{x}, \mathbf{y} \in \mathbb{F}^d$, we define the inner product as $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* \mathbf{y} = \sum_{i=1}^d \bar{x}_i y_i$. The symbol $\|\cdot\|$ denotes the Frobenius norm for matrices and the Euclidean norm for vectors.

For an index set Λ , $(\mathbf{x}_i)_{i \in \Lambda}$ denotes a family indexed by Λ , and we write $(\mathbf{x}_i)_{i \in \Lambda} \subseteq \mathcal{X}$ to mean that $\mathbf{x}_i \in \mathcal{X}$ for all $i \in \Lambda$. When $\Lambda = \{1, \dots, p\}$ is finite, we sometimes regard $(\mathbf{x}_i)_{i=1}^p$ as a p -tuple and use the notation $(\mathbf{x}_i)_{i=1}^p \in \mathcal{X}^p$ to emphasize this viewpoint.

1.3. Outline

The remainder of the paper is organized as follows. In Section 2, we formulate the lifting framework and discuss its connections to classical problems such as phase retrieval and permutation-

invariant embeddings. Section 3 demonstrates the effectiveness of this lifting framework in the finite group setting, establishing rigorous guarantees for orbit separation and bi-Lipschitz stability. Section 4 provides the proofs of the orbit separation results, while Section 5 is devoted to the proofs of the bi-Lipschitz properties. Finally, we conclude the paper in Section 6 with a summary of our contributions and directions for future work.

2. Lifting Framework: Construction and Connections to Related Problems

Let $\mathcal{V}$ be a finite-dimensional inner product space over $\mathbb{F}$, where $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. The inner product is taken to be conjugate-linear in the first argument and linear in the second, i.e., $\langle c\mathbf{v}, \mathbf{w} \rangle = \bar{c}\langle \mathbf{v}, \mathbf{w} \rangle$ and $\langle \mathbf{v}, c\mathbf{w} \rangle = c\langle \mathbf{v}, \mathbf{w} \rangle$. Suppose that a compact group G acts continuously on $\mathcal{V}$ by $\mathbb{F}$ -linear isometries³. The orbit space $\mathcal{V}/G$, consisting of G -orbits $[\mathbf{v}] := G \cdot \mathbf{v}$ for $\mathbf{v} \in \mathcal{V}$, is endowed with the metric

$$d([\mathbf{v}], [\mathbf{w}]) := \min_{g \in G} \|\mathbf{v} - g \cdot \mathbf{w}\|.$$

As discussed in the introduction, our goal is to construct maps $F : \mathcal{V} \rightarrow \mathbb{F}^d$, where d is as small as possible, satisfying the following properties:

- (1) G -invariance: $F(g \cdot \mathbf{v}) = F(\mathbf{v})$ for all $g \in G$ and $\mathbf{v} \in \mathcal{V}$;
- (2) Orbit separation: if $F(\mathbf{v}) = F(\mathbf{w})$, then $[\mathbf{v}] = [\mathbf{w}]$;
- (3) Bi-Lipschitz property: there exist constants $0 < \alpha_F \leq \beta_F < \infty$ such that for all $\mathbf{v}, \mathbf{w} \in \mathcal{V}$,

$$\alpha_F d([\mathbf{v}], [\mathbf{w}]) \leq \|F(\mathbf{v}) - F(\mathbf{w})\| \leq \beta_F d([\mathbf{v}], [\mathbf{w}]).$$

Direct construction of such a map F on $\mathcal{V}$ is challenging. In this section, we develop a framework that starts from an easily constructed G -invariant map f defined on a low-dimensional reduced subspace, and then lifts it to a map F on the entire space $\mathcal{V}$. To motivate our general lifting framework, we first present a classical example that illustrates the key idea in a concrete setting

2.1. A Heuristic Example: Phase Retrieval

Consider the vector space $\mathcal{V} = \mathbb{F}^n$ equipped with the group $G = \{c \in \mathbb{F} : |c| = 1\}$ acting via scalar multiplication. The goal is to construct a G -invariant embedding of $\mathcal{V}$ that is orbit-separating and stable. Rather than attempting to build such an embedding directly in the high-dimensional space, we begin with the one-dimensional case $n = 1$. In this setting, the modulus function $f(x) = |x|$ defines a natural G -invariant embedding of $\hat{\mathcal{V}} = \mathbb{F}$, which we refer to as the reduced space, into $\mathbb{R}_{\geq 0}$. Indeed, for any $x, y \in \mathbb{F}$,

$$\|f(x) - f(y)\| = ||x| - |y|| = \min_{c \in \mathbb{F}, |c|=1} |x - cy| = d([x], [y]),$$

which shows that f is bi-Lipschitz with respect to the orbit metric induced by the group action. In particular, this simple one-dimensional construction completely resolves the group-invariant embedding problem at the level of the reduced space. To lift the map f to higher dimensions, for a general vector $\mathbf{x} = (x_1, \dots, x_n)^\top \in \mathbb{F}^n$, consider a linear combination of its components

$$\Pi_{\alpha}(\mathbf{x}) = \sum_{i=1}^n \alpha_i x_i,$$

³ By Hilbert–Weyl averaging trick, any continuous $\mathbb{F}$ -linear action of a compact group on a finite-dimensional inner product space can be made $\mathbb{F}$ -unitary by averaging the inner product over G using the normalized Haar measure.

where the coefficient vector $\alpha = (\alpha_1, \dots, \alpha_n)^\top \in \mathbb{F}^n$. Each such linear combination reduces the high-dimensional vector $\mathbf{x}$ to a scalar in the reduced space $\hat{\mathcal{V}} = \mathbb{F}$, on which the G -invariant map f can be applied. Composing these two operations yields elementary lifting invariants of the form

$$f \circ \Pi_\alpha(\mathbf{x}) = \left| \sum_{i=1}^n \alpha_i x_i \right| = |\langle \bar{\alpha}, \mathbf{x} \rangle|.$$

By selecting multiple coefficient vectors $\{\alpha^{(j)}\}_{j=1}^N$, one obtains

$$F_{\{\alpha^{(j)}\}_{j=1}^N}(\mathbf{x}) = (f \circ \Pi_{\alpha^{(1)}}(\mathbf{x}), \dots, f \circ \Pi_{\alpha^{(N)}}(\mathbf{x})) = (|\langle \bar{\alpha}^{(1)}, \mathbf{x} \rangle|, \dots, |\langle \bar{\alpha}^{(N)}, \mathbf{x} \rangle|), \quad (1)$$

which is invariant under the action of G . This construction lifts the simple G -invariant embedding f on the reduced space $\mathbb{F}$ to a G -invariant embedding on the ambient space $\mathbb{F}^n$.

This procedure coincides with the classical phase retrieval problem, where the goal is to recover a signal $\mathbf{x} \in \mathbb{F}^n$, up to a unimodular constant, from the magnitudes of linear measurements $(|\langle \bar{\alpha}^{(j)}, \mathbf{x} \rangle|)_{j=1}^N$. In this context, the map $F_{\{\alpha^{(j)}\}_{j=1}^N}$ defined in (1) plays the role of the measurement operator. It is known that $F_{\{\alpha^{(j)}\}_{j=1}^N}$ is orbit-separating for generic choices of $\{\alpha^{(j)}\}_{j=1}^N$ whenever $N \geq 2n - 1$ in the real case or $N \geq 4n - 4$ in the complex case [3,4,42]. Moreover, with regard to stability, orbit separation of $F_{\{\alpha^{(j)}\}_{j=1}^N}$ in fact implies its bi-Lipschitzness [43–45].

While the lifted map $F_{\{\alpha^{(j)}\}_{j=1}^N}$ in (1) is G -invariant, this property does not hold automatically. It relies on the fact that the group acts identically on all components of $\mathbf{x}$. Specifically, when $g \in G$ acts on $\mathbf{x}$, each coordinate x_i is multiplied by the same scalar. This uniformity allows linear combinations of the coordinates to behave consistently with respect to the group action. More precisely, we have $\Pi_\alpha(g \cdot \mathbf{x}) = g \cdot \Pi_\alpha(\mathbf{x})$, so that forming linear combination commutes with the group action. Since f is G -invariant, it follows that

$$f \circ \Pi_\alpha(g \cdot \mathbf{x}) = f(g \cdot \Pi_\alpha(\mathbf{x})) = f \circ \Pi_\alpha(\mathbf{x}).$$

Hence, each component of the lifted map is G -invariant, and therefore the lifted embedding $F_{\{\alpha^{(j)}\}_{j=1}^N}$ in (1) is G -invariant. This demonstrates that the preservation of invariance is closely tied to the uniformity of the group action across all components. This principle will guide the choice of the reduced space $\hat{\mathcal{V}}$ in the general setting.

2.2. The General Lifting Construction

The phase retrieval example above illustrates a mechanism for constructing group-invariant embeddings. We now extend this construction to a general setting.

Let $\mathcal{V}$ be a finite-dimensional inner product space over $\mathbb{F}$ equipped with a linear isometric action of a compact group G . To identify an appropriate reduced space $\hat{\mathcal{V}}$, we appeal to representation theory [46]. Since G is compact, the space $\mathcal{V}$ as a representation of G is completely reducible and admits a unitary isotypic decomposition

$$\mathcal{V} = \bigoplus_{\ell=1}^L \mathcal{V}_\ell^{\oplus n_\ell},$$

where $\{\mathcal{V}_\ell\}_{\ell=1}^L$ are pairwise non-isomorphic⁴ irreducible⁵ representations of G , and $n_\ell \in \mathbb{N}$ denotes the multiplicity of $\mathcal{V}_\ell$. Equivalently, each block $\mathcal{V}_\ell^{\oplus n_\ell}$ consists of n_ℓ mutually isomorphic copies of the same irreducible representation $\mathcal{V}_\ell$, and the group acts on each copy via the same representation. Accordingly, any element $\vec{X} \in \mathcal{V}$ admits a unique decomposition

$$\vec{X} = (X_1, \dots, X_L), \quad \text{with } X_\ell = (\mathbf{x}_{\ell 1}, \dots, \mathbf{x}_{\ell n_\ell}) \in \mathcal{V}_\ell^{\oplus n_\ell},$$

⁴ Two representations of G are said to be isomorphic if there exists a bijective linear map between them that intertwines the group action, i.e., they define the same action of G up to a change of coordinates. Otherwise, they are non-isomorphic.

⁵ A representation of G is called irreducible if it contains no nontrivial G -invariant subspaces.

where each $\mathbf{x}_{\ell j} \in \mathcal{V}_\ell$ lies in one copy of the irreducible component $\mathcal{V}_\ell$. As observed in the phase retrieval example, a key requirement for the lifting construction is that the group action be compatible with linear combinations, which is ensured when the group acts in the same way on the components being combined. In the general setting, such a situation arises precisely within each isotypic component $\mathcal{V}_\ell^{\oplus n_\ell}$, where the group acts in the same way on every copy of $\mathcal{V}_\ell$. This allows one to form linear combinations along the multiplicity direction without violating compatibility with the group action. Accordingly, we select a single representative from each isotypic component and define the reduced space as

$$\widehat{\mathcal{V}} := \oplus_{\ell=1}^L \mathcal{V}_\ell,$$

which naturally inherits a G -action from the irreducible representations $\{\mathcal{V}_\ell\}_{\ell=1}^L$.

Suppose there exists a G -invariant map $f : \widehat{\mathcal{V}} \rightarrow \mathbb{R}^{d'}$ that separates orbits and is ideally bi-Lipschitz on $\widehat{\mathcal{V}}/G$. For any parameter vector $\vec{\alpha} = (\alpha_1, \dots, \alpha_L) \in \oplus_{\ell=1}^L \mathbb{R}^{n_\ell}$, with $\alpha_\ell = (\alpha_{\ell 1}, \dots, \alpha_{\ell n_\ell})^\top \in \mathbb{R}^{n_\ell}$, define the linear map

$$\Pi_{\vec{\alpha}} : \mathcal{V} \rightarrow \widehat{\mathcal{V}}, \quad (X_1, \dots, X_L) \mapsto (X_1 \alpha_1, \dots, X_L \alpha_L), \quad (2)$$

where $X_\ell \alpha_\ell := \sum_{i=1}^{n_\ell} \alpha_{\ell i} \mathbf{x}_{\ell i}$. Composing such a linear map with f yields a map of the form $f \circ \Pi_{\vec{\alpha}}$, which we refer to as an *elementary lifting invariant*. Choosing a finite set of parameters⁶ $\{\vec{\alpha}^{(j)}\}_{j=1}^N \subseteq \oplus_{\ell=1}^L \mathbb{R}^{n_\ell}$, we construct

$$F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N} : \mathcal{V} \rightarrow (\mathbb{R}^{d'})^{\oplus N}, \quad F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}(\vec{X}) := (f \circ \Pi_{\vec{\alpha}^{(1)}}(\vec{X}), \dots, f \circ \Pi_{\vec{\alpha}^{(N)}}(\vec{X})). \quad (3)$$

The choice of reduced space $\widehat{\mathcal{V}}$ ensures $g \cdot \Pi_{\vec{\alpha}}(\vec{X}) = \Pi_{\vec{\alpha}}(g \cdot \vec{X})$ for all $g \in G$, which in turn implies that $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ is G -invariant. The core question is:

How to choose the parameters $\{\vec{\alpha}^{(j)}\}_{j=1}^N$ so that the resulting $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ remains injective on orbits and inherits the bi-Lipschitz property?

The remainder of this work is devoted primarily to answering this question in the finite group setting.

Although the above decomposition $\oplus_{\ell=1}^L \mathcal{V}_\ell^{\oplus n_\ell}$ is into irreducibles, the construction and analysis of F do not require irreducibility. In fact, we only need a G -invariant $\mathbb{F}$ -unitary direct sum decomposition. For later reference, we introduce:

Definition 1 (basic lifting structure). *A triple $(G, \mathcal{V} = \oplus_{\ell=1}^L \mathcal{V}_\ell^{\oplus n_\ell}, f)$ is called a basic lifting structure if:*

- $\mathcal{V}$ is a finite-dimensional $\mathbb{F}$ -inner product space equipped with a continuous $\mathbb{F}$ -linear isometric action of a compact group G .
- $\mathcal{V} = \oplus_{\ell=1}^L \mathcal{V}_\ell^{\oplus n_\ell}$ is a G -invariant $\mathbb{F}$ -unitary direct sum decomposition, where each $\mathcal{V}_\ell^{\oplus n_\ell}$ denotes a direct sum of n_ℓ G -invariant subspaces that are mutually isomorphic as G -representations.
- f is a G -invariant map defined on $\widehat{\mathcal{V}} := \oplus_{\ell=1}^L \mathcal{V}_\ell$.

Terminology: Any G -invariant space $\mathcal{X}$ together with a G -invariant map $F : \mathcal{X} \rightarrow \mathbb{R}^d$ naturally induces a map $\hat{F}$ on the orbit space $\mathcal{X}/G$, defined by $\hat{F}([\mathbf{v}]) = F(\mathbf{v})$. Since we exclusively consider G -invariant maps in this work, we adopt the convention of using the same notation F to refer to properties of the induced map $\hat{F}$ when no ambiguity arises. In particular, statements such as F being *injective*, *injective on a G -invariant subset $\mathcal{M} \subseteq \mathcal{X}$* , *bi-Lipschitz*, or *bi-Lipschitz on $\mathcal{M}$* are to be understood as properties of $\hat{F}$, namely, that $\hat{F}$ is injective, injective on $\mathcal{M}/G$, bi-Lipschitz, or bi-Lipschitz on $\mathcal{M}/G$. Under this convention, F is orbit-separating if and only if it is injective.

⁶ For notational uniformity, when $L = 1$, we omit the arrow notation and simply write $\alpha^{(j)}$ instead of $\vec{\alpha}^{(j)}$.

2.3. Connections to Related Problems

The lifting framework introduced above admits natural connections to various classical and modern problems. We highlight a few representative examples below.

Example 1. (Phase retrieval) A representative example is the phase retrieval problem introduced in Section 2.1. As an application of our lifting framework, we establish in Section 3.2 that, for real phase retrieval, if the measurement operator $F_{\{\alpha^{(j)}\}_{j=1}^N}$ defined in (1) is injective on $\mathbb{R}_s^n$, then it is in fact bi-Lipschitz on $\mathbb{R}_s^n$, where

$$\mathbb{R}_s^n = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \text{ has at most } s \text{ nonzero entries}\}.$$

To the best of our knowledge, this provides the first result establishing an equivalence between injectivity and stability in real sparse phase retrieval.

Example 2. (A generalization of phase retrieval to the matrix setting) Let $\mathcal{V} = \mathbb{F}^{m \times n} \cong (\mathbb{F}^m)^{\oplus n}$ and $G = \{Q \in \mathbb{F}^{m \times m} : Q^*Q = I_m\}$ act on $\mathcal{V}$ by left multiplication. The corresponding reduced space is $\hat{\mathcal{V}} = \mathbb{F}^m$, on which the map $f(\mathbf{x}) = \|\mathbf{x}\|$ is bi-Lipschitz. For any $\{\alpha^{(j)}\}_{j=1}^N \subseteq \mathbb{F}^n$, our lifting framework yields

$$F_{\{\alpha^{(j)}\}_{j=1}^N}(X) = (\|X\alpha^{(1)}\|, \dots, \|X\alpha^{(N)}\|), \quad \text{for } X \in \mathbb{F}^{m \times n}.$$

Note that if $m = 1$, this reduces to the classical phase retrieval problem. When $m = 2$ and $\mathbb{F} = \mathbb{R}$, identifying $\mathbb{C}$ with $\mathbb{R}^2$ places this framework in the setting of the conjugate phase retrieval problem, where one aims to recover a complex signal $\mathbf{x} \in \mathbb{C}^n$, up to a global phase and conjugation, from magnitude measurements $\{|\langle \alpha^{(j)}, \mathbf{x} \rangle|\}_{j=1}^N$ with real measurement vectors $\alpha^{(j)} \in \mathbb{R}^n$. It is shown in [47] that generic $N \geq 4n - 6$ real measurement vectors $\{\alpha^{(j)}\}_{j=1}^N \subseteq \mathbb{R}^n$ have the conjugate phase retrieval property, that is, they suffice for injectivity of $F_{\{\alpha^{(j)}\}_{j=1}^N}$.

When $n \geq m$, this framework connects to the low-rank matrix recovery problem. Observe that $\|X\alpha\|^2 = \alpha^* X^* X \alpha = \text{Tr}(X^* X \alpha \alpha^*)$, which transforms the problem into recovering the matrix $X^* X \in \mathbb{F}^{n \times n}$ with $\text{rank}(X^* X) \leq m$ from the linear measurements $\{\text{Tr}(X^* X \alpha^{(j)} \alpha^{(j)*})\}_{j=1}^N$. Injectivity in this case holds for generic parameter vectors $\{\alpha^{(j)}\}_{j=1}^N$ with $N \geq m + 2m(n - m)$ in the real setting and $N \geq 4m(n - m)$ in the complex setting, as shown in [48, 49]. For $\mathbb{F} = \mathbb{C}$, stability properties of $F_{\{\alpha^{(j)}\}_{j=1}^N}$ have been analyzed in [50].

Example 3. (Multiset and permutation invariance) A multiset $\{\{\mathbf{x}_1, \dots, \mathbf{x}_m\}\} \subseteq \mathbb{R}^n$ is a collection of m elements where order is irrelevant and repetitions are allowed. In machine learning, such a multiset is typically represented by a matrix $X = [\mathbf{x}_1 | \dots | \mathbf{x}_m]^\top \in \mathbb{R}^{m \times n}$ whose rows are $\mathbf{x}_i^\top$. Clearly, permuting the rows of X yields the same multiset.

This fits into our framework with $\mathcal{V} = \mathbb{R}^{m \times n} \cong (\mathbb{R}^m)^{\oplus n}$ and $G = S_m$, the symmetric group of $m \times m$ permutation matrices acting on $\mathcal{V}$ by left multiplication. The corresponding reduced space is $\hat{\mathcal{V}} = \mathbb{R}^m$, on which the map $f(\mathbf{x}) = \text{sort}(\mathbf{x})$ is bi-Lipschitz. Here, $\text{sort} : \mathbb{R}^m \rightarrow \mathbb{R}^m$ denotes the operator that arranges the coordinates in nonincreasing order. For any $\{\alpha^{(j)}\}_{j=1}^N \subseteq \mathbb{R}^n$, our lifting procedure gives

$$F_{\{\alpha^{(j)}\}_{j=1}^N}(X) = (\text{sort}(X\alpha^{(1)}), \dots, \text{sort}(X\alpha^{(N)})), \quad \text{for } X \in \mathbb{R}^{m \times n}.$$

Theorem 3.12 in [41] shows that $F_{\{\alpha^{(j)}\}_{j=1}^N}$ is bi-Lipschitz whenever it is injective. Moreover, Proposition 3.1 in [33] ensures that injectivity holds for generic $N \geq 2mn + 1$ parameter vectors $\{\alpha^{(j)}\}_{j=1}^N$ in $\mathbb{R}^n$. In Section 3, we will strengthen these results by showing that for generic $N \geq mn + m + n$ such parameter vectors, the map $F_{\{\alpha^{(j)}\}_{j=1}^N}$ is in fact bi-Lipschitz. We also extend our analysis to the sparse setting.

Example 4. (A cyclic group action) Let $G = \langle e^{2\pi i/r} \rangle$ act on $\mathcal{V} = \mathbb{C}^n \cong \mathbb{C}^{\oplus n}$ by scalar multiplication. The corresponding reduced space is $\hat{\mathcal{V}} = \mathbb{C}$. Example 18 in [32] shows that the map⁷ $f(z) = \frac{z^r}{|z|^{r-1}}$ is bi-Lipschitz on $\mathbb{C}$. Then for any $\{\alpha^{(j)}\}_{j=1}^N \subseteq \mathbb{C}^n$, our lifting procedure gives

$$F_{\{\alpha^{(j)}\}_{j=1}^N}(\mathbf{x}) = \left(\frac{\langle \bar{\alpha}^{(1)}, \mathbf{x} \rangle^r}{|\langle \bar{\alpha}^{(1)}, \mathbf{x} \rangle|^{r-1}}, \dots, \frac{\langle \bar{\alpha}^{(N)}, \mathbf{x} \rangle^r}{|\langle \bar{\alpha}^{(N)}, \mathbf{x} \rangle|^{r-1}} \right), \quad \text{for } \mathbf{x} \in \mathbb{C}^n.$$

Several related constructions appear in the literature. In [30], Cahill, Contreras, and Contreras Hip introduced the map

$$F_1(\mathbf{x}) = \frac{1}{\|\mathbf{x}\|^{r-1}} (x_1^r, \dots, x_n^r, \{x_j x_k^{r-1}\}_{1 \leq j < k \leq n}), \quad \mathbf{x} = (x_1, \dots, x_n)^\top \in \mathbb{C}^n,$$

which is Lipschitz but not lower Lipschitz on $\mathbb{C}^n$. This was later strengthened in Theorem 30 of [32] by adding the additional terms $x_j x_k^{r-1}$ for $j > k$, yielding

$$F_2(\mathbf{x}) = \frac{1}{\|\mathbf{x}\|^{r-1}} (x_1^r, \dots, x_n^r, \{x_j x_k^{r-1}\}_{1 \leq j, k \leq n})$$

which is bi-Lipschitz on $\mathbb{C}^n$. More recently, Theorem 16 in [16] proposed

$$F_3(\mathbf{x}) = \left(\sin\left(\frac{\pi}{2r}\right) \frac{\mathbf{x}^{\otimes r}}{\|\mathbf{x}\|^{r-1}}, \cos\left(\frac{\pi}{2r}\right) \frac{\mathbf{x} \mathbf{x}^*}{\|\mathbf{x}\|} \right),$$

which is bi-Lipschitz on $\mathbb{C}^n$ with the optimal distortion, where $\mathbf{x}^{\otimes r}$ denotes the r -th order tensor of $\mathbf{x}$.

A key consideration in constructing such a G -invariant map $F : \mathcal{V} \rightarrow \mathbb{F}^d$ is to keep the embedding dimension d as small as possible. Whereas the construction F_2 requires $d = O(n^2)$ and F_3 requires $d = O(n^r)$, our approach yields an embedding of dimension $d = N$. Furthermore, as will be demonstrated in Section 3, it suffices to choose $N = 2n + 1 = O(n)$ parameter vectors in order to obtain a bi-Lipschitz embedding $F_{\{\alpha^{(j)}\}_{j=1}^N}$ (see Corollary 3).

3. Orbit Separation and Bi-Lipschitz Stability of the Lifting Framework

In this section, we demonstrate the effectiveness of our lifting framework in the finite group setting.

3.1. Orbit Separation for General Finite Groups

We now treat a general finite group G acting $\mathbb{F}$ -linearly on $\mathcal{V}$. Our aim is to establish general sufficient conditions ensuring orbit separation (i.e., injectivity) of the lifting map $F_{\{\alpha^{(j)}\}_{j=1}^N}$. Recall that any finite-dimensional $\mathbb{F}$ -vector space $\mathcal{X}$ is isomorphic to $\mathbb{F}^{\dim_{\mathbb{F}}(\mathcal{X})}$. Hence, for $\mathcal{V} = \bigoplus_{\ell=1}^L \mathcal{V}_{\ell}^{\oplus n_{\ell}}$, we may, without loss of generality, identify

$$\mathcal{V} = \bigoplus_{\ell=1}^L (\mathbb{F}^{m_{\ell}})^{\oplus n_{\ell}}.$$

This reduction allows us to exploit tools from real algebraic geometry in the analysis that follows. We begin with a few basic definitions.

Definition 2. (Semi-algebraic set) A set $S \subseteq \mathbb{R}^d$ is called semi-algebraic if it can be expressed as a finite union of sets of the form

$$\{\mathbf{x} \in \mathbb{R}^d : p_s(\mathbf{x}) = 0, q_t(\mathbf{x}) > 0 \text{ for } s = 1, \dots, S \text{ and } t = 1, \dots, T\},$$

⁷ Example 18 in [32] also provides another bi-Lipschitz map $\tilde{f}(z) = \left(\frac{z^r}{|z|^{r-1}} \sin \frac{\pi}{2r}, |z| \cos \frac{\pi}{2r} \right)^\top$, which has a smaller distortion than $f(z) = \frac{z^r}{|z|^{r-1}}$.

where $p_s, q_t \in \mathbb{R}[x_1, \dots, x_d]$ are multivariate polynomials with real coefficients.

According to Theorem 2.3.6 in [51], every semi-algebraic set $\mathcal{S} \subseteq \mathbb{R}^d$ can be decomposed into a finite union of sets, each of which is semi-algebraically homeomorphic to an open hypercube $]0, 1[^{d_j}$. The real dimension of $\mathcal{S}$, denoted by $\dim_{\mathbb{R}}(\mathcal{S})$, is defined as the maximum of the d_j .

Definition 3. (Semi-algebraic function) Let $\mathcal{S} \subseteq \mathbb{R}^{d_1}$ and $\mathcal{T} \subseteq \mathbb{R}^{d_2}$ be semi-algebraic sets. A function $f : \mathcal{S} \rightarrow \mathcal{T}$ is called semi-algebraic if its graph

$$\{(\mathbf{x}, \mathbf{y}) \in \mathcal{S} \times \mathcal{T} : \mathbf{y} = f(\mathbf{x})\}$$

is a semi-algebraic set in $\mathbb{R}^{d_1+d_2}$.

In fact, the class of semi-algebraic functions includes many familiar examples, such as polynomial and rational functions. Moreover, all functions f considered in Section 2.3 are semi-algebraic.

To make the notion of semi-algebraic sets and functions well-defined in the complex setting, we adopt the standard identification $\mathbb{C} \cong \mathbb{R}^2$, and interpret complex vector spaces and functions in the real sense. Consequently, whenever we speak of semi-algebraic notions in the complex case, they are always understood as their real counterparts via this identification. For instance, when $\mathbb{F} = \mathbb{C}$, we view $\mathcal{V} = \bigoplus_{\ell=1}^L (\mathbb{C}^{m_\ell})^{\oplus n_\ell}$ as isomorphic to $\bigoplus_{\ell=1}^L (\mathbb{R}^{2m_\ell})^{\oplus n_\ell}$. A subset $\mathcal{M} \subseteq \mathcal{V}$ is then said to be semi-algebraic if, under this identification, it corresponds to a semi-algebraic subset of $\bigoplus_{\ell=1}^L (\mathbb{R}^{2m_\ell})^{\oplus n_\ell}$.

We can now state our general injectivity results, which are proved in Section 4. Our first theorem concerns injectivity on arbitrary G -invariant semi-algebraic subsets of $\mathcal{V}$.

Theorem 1. Let $(G, \mathcal{V} = \bigoplus_{\ell=1}^L (\mathbb{F}^{m_\ell})^{\oplus n_\ell}, f)$ be a basic lifting structure, with G finite. Assume f is semi-algebraic and injective. Let $\mathcal{M} \subseteq \mathcal{V}$ be a G -invariant semi-algebraic subset. Write $C_{\mathbb{F}} = \dim_{\mathbb{R}}(\mathbb{F})$ and assume $N \geq 2 \dim_{\mathbb{R}}(\mathcal{M}) / C_{\mathbb{F}} + 1$. For any $\{\tilde{\alpha}^{(j)}\}_{j=1}^N \subseteq \bigoplus_{\ell=1}^L \mathbb{F}^{n_\ell}$, let $F_{\{\tilde{\alpha}^{(j)}\}_{j=1}^N}$ be defined as in (3). Then for almost every⁸ $(\tilde{\alpha}^{(j)})_{j=1}^N \in (\bigoplus_{\ell=1}^L \mathbb{F}^{n_\ell})^N$, the map $F_{\{\tilde{\alpha}^{(j)}\}_{j=1}^N}$ is injective on $\mathcal{M}$.

In particular, Theorem 1 applies when $\mathcal{M} = \mathcal{V}$. However, in the special case where $\mathcal{V}$ admits the simple decomposition $(\mathbb{F}^m)^{\oplus n}$ (i.e., the $L = 1$ setting), the general bound obtained above can be sharpened. The next theorem provides such an improvement.

Theorem 2. Let $(G, \mathcal{V} = (\mathbb{F}^m)^{\oplus n}, f)$ be a basic lifting structure, with G finite. Assume f is semi-algebraic and injective. For any $\{\alpha^{(j)}\}_{j=1}^N \subseteq \mathbb{F}^n$, let $F_{\{\alpha^{(j)}\}_{j=1}^N}$ be defined as in (3). If $N \geq mn + m + n$, then for almost every $(\alpha^{(j)})_{j=1}^N \in (\mathbb{F}^n)^N$, the map $F_{\{\alpha^{(j)}\}_{j=1}^N}$ is injective.

Remark 1. Theorem 1 is largely based on the argument of Theorem 2.7 in [33]. However, for the special case $\mathcal{V} = (\mathbb{F}^m)^{\oplus n}$ with reduced space $\mathbb{F}^m$, Theorem 2 yields a strictly stronger guarantee than the results in [33], as reflected by the bound $mn + m + n \leq 2 \dim_{\mathbb{R}}(\mathcal{V}) / C_{\mathbb{F}} + 1 = 2mn + 1$.

Applying Theorem 2 to Example 3, we obtain the following result.

Corollary 1. Let $\mathcal{V} = \mathbb{R}^{m \times n}$ and let $G = S_m$ act on $\mathcal{V}$ by permuting the rows of matrices. If $N \geq mn + m + n$, then for almost every $(\alpha^{(j)})_{j=1}^N \in (\mathbb{R}^n)^N$, the map

$$F : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times N}, \quad F(X) = (\text{sort}(X\alpha^{(1)}), \dots, \text{sort}(X\alpha^{(N)})) \quad (4)$$

is injective.

⁸ Throughout this paper, the phrase “for almost every $\mathbf{x} \in \mathcal{X}$ ” is understood with respect to the Hausdorff measure on $\mathcal{X}$.

Remark 2. We are aware of the recent preprint [52], which proves sharper bounds in the permutation-invariant setting. In particular, Theorem 3 in [52] shows that $N \geq m(n-1) + 1$ already suffices for the injectivity of F in (4), improving upon Corollary 1. Moreover, Theorem 4 in [52] demonstrates that if $\lceil N/(n-1) \rceil \leq \log_2(m) + 1$, then the map F in (4) can never be injective. However, these results rely on structural properties specific to the symmetric group S_m , while our approach applies to general finite groups and does not depend on such special features.

3.2. Bi-Lipschitz Stability for General Finite Groups

For practical applications, injectivity alone is not sufficient. Small perturbations in the input should not lead to arbitrarily large deviations in the output. In this subsection, we establish bi-Lipschitz properties for lifting embeddings arising from general finite group actions. Our key observation is that, if injectivity holds on a G -invariant subset $\mathcal{M}$ that is a finite union of linear subspaces, then the bi-Lipschitz property follows automatically. This principle is formalized in the following theorem.

Theorem 3. Let $(G, \mathcal{V} = \bigoplus_{\ell=1}^L \mathcal{V}_\ell^{\oplus n_\ell}, f)$ be a basic lifting structure with G finite, and assume that f is bi-Lipschitz. Let $\mathcal{M} = \bigcup_{i=1}^M \mathcal{L}_i \subseteq \mathcal{V}$ be a G -invariant subset with each $\mathcal{L}_i$ an $\mathbb{F}$ -linear subspace of $\mathcal{V}$. For any $\{\tilde{\alpha}^{(j)}\}_{j=1}^N \subseteq \bigoplus_{\ell=1}^L \mathbb{F}^{n_\ell}$, let $F_{\{\tilde{\alpha}^{(j)}\}_{j=1}^N}$ be defined as in (3). If $F_{\{\tilde{\alpha}^{(j)}\}_{j=1}^N}$ is injective on $\mathcal{M}$, then it is bi-Lipschitz on $\mathcal{M}$.

The proof of Theorem 3 is given in Section 5. In particular, taking $\mathcal{M} = \mathcal{V}$ shows that $F_{\{\tilde{\alpha}^{(j)}\}_{j=1}^N}$ is bi-Lipschitz whenever it is injective. Combining this with Corollary 1, we arrive at the following result for the permutation group action.

Corollary 2. Let $\mathcal{V} = \mathbb{R}^{m \times n}$ and let $G = S_m$ act on $\mathcal{V}$ by permuting the rows of matrices. If $N \geq mn + m + n$, then for almost every $(\alpha^{(j)})_{j=1}^N \in (\mathbb{R}^n)^N$, the map

$$F : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times N}, \quad F(X) = (\text{sort}(X\alpha^{(1)}), \dots, \text{sort}(X\alpha^{(N)}))$$

is bi-Lipschitz.

Applying Theorem 2 and Theorem 3 to Example 4, we immediately obtain the following result.

Corollary 3. Let $\mathcal{V} = \mathbb{C}^n$ and let $G = \langle e^{2\pi i/r} \rangle$ act on $\mathcal{V}$ by scalar multiplication. Then when $N \geq 2n + 1$, for almost every $(\alpha^{(j)})_{j=1}^N \in (\mathbb{C}^n)^N$, the map

$$F_{\{\alpha^{(j)}\}_{j=1}^N}(\mathbf{x}) = \left(\frac{\langle \tilde{\alpha}^{(1)}, \mathbf{x} \rangle^r}{|\langle \tilde{\alpha}^{(1)}, \mathbf{x} \rangle|^{r-1}}, \dots, \frac{\langle \tilde{\alpha}^{(N)}, \mathbf{x} \rangle^r}{|\langle \tilde{\alpha}^{(N)}, \mathbf{x} \rangle|^{r-1}} \right), \quad \text{for } \mathbf{x} \in \mathbb{C}^n.$$

is bi-Lipschitz.

Theorem 3 also provides a natural perspective to study sparse models. In many signal processing problems, one restricts attention to vectors or matrices with only a small number of nonzero entries. The union-of-subspaces structure appearing in sparsity fits precisely into our setting. To make this precise, let

$$\mathbb{R}_s^n := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_0 \leq s\} \quad \text{and} \quad \mathbb{R}_s^{m \times n} := \{X \in \mathbb{R}^{m \times n} : \|X\|_0 \leq s\},$$

where $\|\cdot\|_0$ denotes the number of nonzero entries. The next corollary specializes to a stable formulation of real sparse phase retrieval.

Corollary 4. Let $\mathcal{V} = \mathbb{R}^n$ and let $G = \{\pm 1\}$ act on $\mathcal{V}$ by scalar multiplication. If $N \geq 2s$, then for almost every $(\alpha^{(j)})_{j=1}^N \in (\mathbb{R}^n)^N$, the map

$$F : \mathbb{R}^n \rightarrow \mathbb{R}^N, \quad F(\mathbf{x}) = (|\langle \mathbf{x}, \alpha^{(1)} \rangle|, \dots, |\langle \mathbf{x}, \alpha^{(N)} \rangle|)^\top$$

is bi-Lipschitz on $\mathbb{R}_s^n$.

Proof. By Theorem 2.1 in [53], when $N \geq 2s$ the map F is injective on $\mathbb{R}_s^n$ for almost every $(\alpha^{(j)})_{j=1}^N \in (\mathbb{R}^n)^N$. Note that

$$\mathbb{R}_s^n = \bigcup_{\mathcal{I} \subseteq \{1, \dots, n\} : |\mathcal{I}|=s} \mathbb{R}_{\mathcal{I}}^n, \quad \mathbb{R}_{\mathcal{I}}^n := \{\mathbf{x} = (x_1, \dots, x_n)^\top \in \mathbb{R}^n : x_i = 0 \text{ for } i \notin \mathcal{I}\},$$

and each $\mathbb{R}_{\mathcal{I}}^n$ is a linear subspace of $\mathbb{R}^n$. Obviously, $\mathbb{R}_s^n$ is G -invariant. Thus the conclusion follows from Theorem 3. $\square$

We can further extend this stability guarantee to the permutation group action. The following corollary shows that bi-Lipschitz stability also holds for sparse matrices under row permutations.

Corollary 5. Let $\mathcal{V} = \mathbb{R}^{m \times n}$ and let $G = S_m$ act on $\mathcal{V}$ by permuting the rows of matrices. If $N \geq 2s + 1$, then for almost every $(\alpha^{(j)})_{j=1}^N \in (\mathbb{R}^n)^N$, the map

$$F : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times N}, \quad F(X) = (\text{sort}(X\alpha^{(1)}), \dots, \text{sort}(X\alpha^{(N)}))$$

is bi-Lipschitz on $\mathbb{R}_s^{m \times n}$.

Proof. Applying Theorem 1 with $\mathcal{M} = \mathbb{R}_s^{m \times n}$, we have that F is injective on $\mathbb{R}_s^{m \times n}$ for almost every $(\alpha^{(j)})_{j=1}^N \in (\mathbb{R}^n)^N$ whenever $N \geq 2s + 1$, since $\dim_{\mathbb{R}}(\mathbb{R}_s^{m \times n}) = s$. Using an argument similar to that in the proof of Corollary 4, the set $\mathbb{R}_s^{m \times n}$ is G -invariant and can be written as a finite union of linear subspaces of $\mathbb{R}^{m \times n}$. Thus the conclusion follows from Theorem 3. $\square$

4. Proofs of Theorem 1 and Theorem 2

In this section, we provide detailed proofs of Theorem 1 and Theorem 2. Since our arguments are carried out in the semi-algebraic setting, it is convenient to work with real representations of complex matrices. Define

$$L_{\mathbb{C}} : \mathbb{C}^{m \times n} \rightarrow \mathbb{R}^{2m \times 2n}, \quad X = X_0 + X_1 i \mapsto \begin{bmatrix} X_0 & -X_1 \\ X_1 & X_0 \end{bmatrix}$$

and

$$\ell_{\mathbb{C}} : \mathbb{C}^n \rightarrow \mathbb{R}^{2n}, \quad \mathbf{x} = \mathbf{x}_0 + \mathbf{x}_1 i \mapsto \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{x}_1 \end{bmatrix}.$$

And $L_{\mathbb{R}}$ and $\ell_{\mathbb{R}}$ are taken to be the identity maps on the corresponding spaces. It is straightforward to verify that these maps satisfy the compatibility relation

$$\ell_{\mathbb{F}}(X\alpha) = L_{\mathbb{F}}(X)\ell_{\mathbb{F}}(\alpha) \quad \text{for any } X \in \mathbb{F}^{m \times n}, \alpha \in \mathbb{F}^n.$$

Moreover, applying the singular value decomposition gives

$$\text{rank}(L_{\mathbb{F}}(X)) = C_{\mathbb{F}} \cdot \text{rank}(X) \quad \text{for any } X \in \mathbb{F}^{m \times n}. \quad (5)$$

In addition, we make use of several basic properties of real algebraic dimension.

Lemma 1. ([51,54])

(i) (Proposition 5.28, [54]) If $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathbb{R}^d$ are semi-algebraic sets, then $\dim_{\mathbb{R}}(\mathcal{A}) \leq \dim_{\mathbb{R}}(\mathcal{B})$.

- (ii) (Proposition 2.8.5, [51]) If $\mathcal{A} \subseteq \mathbb{R}^{d_1}$ and $\mathcal{B} \subseteq \mathbb{R}^{d_2}$ are semi-algebraic sets, then $\dim_{\mathbb{R}}(\mathcal{A} \times \mathcal{B}) = \dim_{\mathbb{R}}(\mathcal{A}) + \dim_{\mathbb{R}}(\mathcal{B})$.
- (iii) (Proposition 2.8.5, [51]) If $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^d$ are semi-algebraic sets, then $\dim_{\mathbb{R}}(\mathcal{A} \cup \mathcal{B}) = \max\{\dim_{\mathbb{R}}(\mathcal{A}), \dim_{\mathbb{R}}(\mathcal{B})\}$.
- (iv) (Proposition 2.2.7 and Theorem 2.8.8, [51]) If $\mathcal{S} \subseteq \mathbb{R}^{d_1}$ is semi-algebraic and $\pi : \mathcal{S} \rightarrow \mathbb{R}^{d_2}$ is a semi-algebraic function, then $\pi(\mathcal{S})$ is semi-algebraic and $\dim_{\mathbb{R}}(\pi(\mathcal{S})) \leq \dim_{\mathbb{R}}(\mathcal{S})$, with equality if π is one-to-one on $\mathcal{S}$.
- (v) (Appendix A, [49]) If $\mathcal{S}$ is a semi-algebraic set of dimension d , then it has zero $(d+1)$ -dimensional Hausdorff measure.

We will also repeatedly use the following dimensional estimate.

Lemma 2. (Lemma 2.8, [33]) Let $\mathcal{S} \subseteq \mathbb{R}^{d_1}$ be a semi-algebraic set and $\pi : \mathcal{S} \rightarrow \mathbb{R}^{d_2}$ be a polynomial map. Assume for all $t \in \pi(\mathcal{S})$ we have that $\dim_{\mathbb{R}}(\pi^{-1}(t)) \leq \Delta$, then

$$\dim_{\mathbb{R}}(\mathcal{S}) \leq \dim_{\mathbb{R}}(\pi(\mathcal{S})) + \Delta.$$

4.1. Proof of Theorem 1

We now present the proof of Theorem 1.

Proof of Theorem 1. For each $N \in \mathbb{N}$, consider the “bad set”

$$\mathcal{B}_N := \left\{ ((\vec{X}, \vec{Y}), (\vec{a}^{(j)})_{j=1}^N) \in \mathcal{M}^2 \times (\oplus_{\ell=1}^L \mathbb{F}^{n_{\ell}})^N : \right. \\ \left. [\vec{X}] \neq [\vec{Y}], F_{\{\vec{a}^{(j)}\}_{j=1}^N}(\vec{X}) = F_{\{\vec{a}^{(j)}\}_{j=1}^N}(\vec{Y}) \right\}.$$

Since f is semi-algebraic and $\mathcal{M}$ is a semi-algebraic set, it follows that $\mathcal{B}_N$ is a semi-algebraic set. Let π_1 and π_2 denote the projections from $\mathcal{B}_N$ to $\mathcal{M}^2$ and $(\oplus_{\ell=1}^L \mathbb{F}^{n_{\ell}})^N$, respectively, namely

$$\pi_1((\vec{X}, \vec{Y}), (\vec{a}^{(1)}, \dots, \vec{a}^{(N)})) = (\vec{X}, \vec{Y}), \quad \pi_2((\vec{X}, \vec{Y}), (\vec{a}^{(1)}, \dots, \vec{a}^{(N)})) = (\vec{a}^{(1)}, \dots, \vec{a}^{(N)}).$$

Then $\pi_2(\mathcal{B}_N)$ is precisely the set of parameter tuples for which $F_{\{\vec{a}^{(j)}\}_{j=1}^N}$ fails to be injective on $\mathcal{M}$. Our goal is to show that

$$\dim_{\mathbb{R}}(\mathcal{B}_N) < \dim_{\mathbb{R}}((\oplus_{\ell=1}^L \mathbb{F}^{n_{\ell}})^N), \quad \text{whenever } N \geq 2 \dim_{\mathbb{R}}(\mathcal{M})/C_{\mathbb{F}} + 1. \quad (6)$$

To see why this suffices, observe that

$$\dim_{\mathbb{R}}(\pi_2(\mathcal{B}_N)) \leq \dim_{\mathbb{R}}(\mathcal{B}_N) < \dim_{\mathbb{R}}((\oplus_{\ell=1}^L \mathbb{F}^{n_{\ell}})^N).$$

Then by the conclusion (v) in Lemma 1, we conclude that for almost every parameter choice $\{\vec{a}^{(j)}\}_{j=1}^N$, the lifting map $F_{\{\vec{a}^{(j)}\}_{j=1}^N}$ is injective on $\mathcal{M}$.

We now proceed to establish the dimension estimate (6). By Lemma 2, we have

$$\dim_{\mathbb{R}}(\mathcal{B}_N) \leq \dim_{\mathbb{R}}(\pi_1(\mathcal{B}_N)) + \max_{(\vec{X}, \vec{Y}) \in \pi_1(\mathcal{B}_N)} \dim_{\mathbb{R}}(\pi_1^{-1}(\vec{X}, \vec{Y})). \quad (7)$$

We next estimate $\dim_{\mathbb{R}}(\pi_1^{-1}(\vec{X}, \vec{Y}))$. For any fixed $(\vec{X}, \vec{Y}) \in \pi_1(\mathcal{B}_N)$,

$$\begin{aligned} \pi_1^{-1}(\vec{X}, \vec{Y}) &= \{(\vec{X}, \vec{Y})\} \times \left\{ (\vec{\alpha}^{(j)})_{j=1}^N \in (\oplus_{\ell=1}^L \mathbb{F}^{n_\ell})^N : F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}(\vec{X}) = F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}(\vec{Y}) \right\} \\ &\stackrel{(*)}{=} \{(\vec{X}, \vec{Y})\} \times \left\{ (\vec{\alpha}^{(j)})_{j=1}^N \in (\oplus_{\ell=1}^L \mathbb{F}^{n_\ell})^N : [\Pi_{\vec{\alpha}^{(j)}}(\vec{X})] = [\Pi_{\vec{\alpha}^{(j)}}(\vec{Y})] \text{ for } j = 1, \dots, N \right\} \\ &\subseteq \{(\vec{X}, \vec{Y})\} \times \left(\bigcup_{g \in G} \{ \vec{\alpha} \in \oplus_{\ell=1}^L \mathbb{F}^{n_\ell} : \Pi_{\vec{\alpha}}(\vec{X} - g \cdot \vec{Y}) = \mathbf{0} \} \right)^N \\ &=: \{(\vec{X}, \vec{Y})\} \times \left(\bigcup_{g \in G} \mathcal{W}_g \right)^N, \end{aligned}$$

where the equality $(*)$ follows from the injectivity of f . Since $[\vec{X}] \neq [\vec{Y}]$, we have $\vec{X} - g \cdot \vec{Y} \neq \mathbf{0}$ for all $g \in G$. Thus, for each $g \in G$ there exists some $\ell_0 \in \{1, \dots, L\}$ with $\mathbf{0} \neq X_{\ell_0} - g \cdot Y_{\ell_0} \in \mathbb{F}^{m_{\ell_0} \times n_{\ell_0}}$. Here we write $\vec{X} = (X_1, \dots, X_L) \in \mathcal{M} \subseteq \oplus_{\ell=1}^L (\mathbb{F}^{m_\ell})^{\oplus n_\ell}$ with $X_\ell \in (\mathbb{F}^{m_\ell})^{\oplus n_\ell} = \mathbb{F}^{m_\ell \times n_\ell}$, and similarly $\vec{Y} = (Y_1, \dots, Y_L)$. By the equation (5),

$$\text{rank}(L_{\mathbb{F}}(X_{\ell_0} - g \cdot Y_{\ell_0})) = C_{\mathbb{F}} \cdot \text{rank}(X_{\ell_0} - g \cdot Y_{\ell_0}) \geq C_{\mathbb{F}}.$$

Hence, for $\vec{\alpha} = (\alpha_1, \dots, \alpha_L) \in \mathcal{W}_g$, the constraint $\Pi_{\vec{\alpha}}(\vec{X} - g \cdot \vec{Y}) = \mathbf{0}$ implies $\mathbf{0} = \ell_{\mathbb{F}}((X_{\ell_0} - g \cdot Y_{\ell_0})\alpha_{\ell_0}) = L_{\mathbb{F}}(X_{\ell_0} - g \cdot Y_{\ell_0})\ell_{\mathbb{F}}(\alpha_{\ell_0})$, which forces $\ell_{\mathbb{F}}(\alpha_{\ell_0})$ to lie in a subspace of real codimension at least $C_{\mathbb{F}}$. Therefore

$$\dim_{\mathbb{R}}(\mathcal{W}_g) \leq \dim_{\mathbb{R}}(\oplus_{\ell=1}^L \mathbb{F}^{n_\ell}) - C_{\mathbb{F}}.$$

Since G is finite, it follows that

$$\dim_{\mathbb{R}}(\pi_1^{-1}(\vec{X}, \vec{Y})) \leq N \cdot (\dim_{\mathbb{R}}(\oplus_{\ell=1}^L \mathbb{F}^{n_\ell}) - C_{\mathbb{F}}) = \dim_{\mathbb{R}}((\oplus_{\ell=1}^L \mathbb{F}^{n_\ell})^N) - NC_{\mathbb{F}}. \quad (8)$$

Finally, observe that $\dim_{\mathbb{R}}(\pi_1(\mathcal{B}_N)) \leq 2 \dim_{\mathbb{R}}(\mathcal{M})$. Combining this with (7) and (8), we obtain

$$\dim_{\mathbb{R}}(\mathcal{B}_N) \leq 2 \dim_{\mathbb{R}}(\mathcal{M}) + \dim_{\mathbb{R}}((\oplus_{\ell=1}^L \mathbb{F}^{n_\ell})^N) - NC_{\mathbb{F}}.$$

When $N \geq 2 \dim_{\mathbb{R}}(\mathcal{M})/C_{\mathbb{F}} + 1$, the right-hand side is strictly less than $\dim_{\mathbb{R}}((\oplus_{\ell=1}^L \mathbb{F}^{n_\ell})^N)$, which completes the proof. $\square$

4.2. Proof of Theorem 2

To prove Theorem 2, we need the following auxiliary lemma.

Lemma 3. Let $r \leq \min\{m, n\}$. The set $\mathcal{M}_r := \{X \in \mathbb{F}^{m \times n} : \text{rank}(X) = r\}$ is a semi-algebraic set with real dimension $C_{\mathbb{F}} \cdot r(m + n - r)$, where $C_{\mathbb{F}} = \dim_{\mathbb{R}}(\mathbb{F})$.

Proof. It is well known that $\text{rank}(X) = r$ is equivalent to the vanishing of all $(r+1) \times (r+1)$ minors of X , together with the non-vanishing of at least one $r \times r$ minor. Since minors are polynomial functions of the entries of X , which are regarded as real coordinates via the identification $\mathbb{F}^{m \times n} \cong \mathbb{R}^{C_{\mathbb{F}}mn}$, the set $\mathcal{M}_r$ is semi-algebraic.

To compute its real dimension, consider the action of the group $\text{GL}_m(\mathbb{F}) \times \text{GL}_n(\mathbb{F})$ on $\mathbb{F}^{m \times n}$ given by $(A, B) \cdot X = AXB$. Let $X_0 = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{F}^{m \times n}$. Then $\mathcal{M}_r$ coincides with the orbit $\mathcal{O}_{X_0}$ of X_0 under this action. A direct calculation shows the stabilizer subgroup of X_0 is

$$\begin{aligned} \text{Stab}(X_0) &= \{ (A, B) \in \text{GL}_m(\mathbb{F}) \times \text{GL}_n(\mathbb{F}) : AX_0B = X_0 \} \\ &= \left\{ \left(\begin{bmatrix} A_{11} & A_{12} \\ \mathbf{0}_{(m-r) \times r} & A_{22} \end{bmatrix}, \begin{bmatrix} B_{11} & \mathbf{0}_{r \times (n-r)} \\ B_{21} & B_{22} \end{bmatrix} \right) \in \text{GL}_m(\mathbb{F}) \times \text{GL}_n(\mathbb{F}) : \right. \\ &\quad \left. A_{11}, B_{11} \in \mathbb{F}^{r \times r}, A_{11}B_{11} = I_r \right\}. \end{aligned}$$

Then by Proposition 2.8.14 in [51] and Theorem 3.62 in [55], we have

$$\begin{aligned} \dim_{\mathbb{R}}(\mathcal{M}_r) &= \dim_{\mathbb{R}}(\mathcal{O}_{X_0}) = \dim_{\mathbb{R}}(\text{GL}_m(\mathbb{F}) \times \text{GL}_n(\mathbb{F})) - \dim_{\mathbb{R}}(\text{Stab}(X_0)) \\ &= C_{\mathbb{F}} \cdot (m^2 + n^2) - C_{\mathbb{F}} \cdot (r^2 + m(m-r) + n(n-r)) = C_{\mathbb{F}} \cdot r(m+n-r), \end{aligned}$$

where $\dim_{\mathbb{R}}(\mathcal{O}_{X_0})$ denotes the dimension of the orbit $\mathcal{O}_{X_0}$ as a smooth submanifold of $\mathbb{R}^{C_{\mathbb{F}}mn}$. This completes the proof. $\square$

We now proceed to prove Theorem 2.

Proof of Theorem 2. For each $N \in \mathbb{N}$, consider the “bad set”

$$\begin{aligned} \mathcal{B}_N &:= \left\{ ((X, Y), (\alpha^{(j)})_{j=1}^N) \in (\mathbb{F}^{m \times n})^2 \times (\mathbb{F}^n)^N : \right. \\ &\quad \left. [X] \neq [Y], F_{\{\alpha^{(j)}\}_{j=1}^N}(X) = F_{\{\alpha^{(j)}\}_{j=1}^N}(Y) \right\}. \end{aligned}$$

Since f is semi-algebraic, it follows that $\mathcal{B}_N$ is a semi-algebraic set. As in the proof of Theorem 1, it suffices to show that if $N \geq mn + m + n$, then $\dim_{\mathbb{R}}(\mathcal{B}_N) < \dim_{\mathbb{R}}((\mathbb{F}^n)^N)$, which implies that for almost every parameter choice $\{\alpha^{(j)}\}_{j=1}^N$, the lifting map $F_{\{\alpha^{(j)}\}_{j=1}^N}$ is injective.

Since f is injective, the equality $F_{\{\alpha^{(j)}\}_{j=1}^N}(X) = F_{\{\alpha^{(j)}\}_{j=1}^N}(Y)$ implies that for each j there exists $g_j \in G$ such that $(X - g_j \cdot Y)\alpha^{(j)} = \mathbf{0}$. Therefore

$$\begin{aligned} \mathcal{B}_N &\subseteq \bigcup_{(g_j)_{j=1}^N \in G^N} \left\{ ((X, Y), (\alpha^{(j)})_{j=1}^N) \in (\mathbb{F}^{m \times n})^2 \times (\mathbb{F}^n)^N : \right. \\ &\quad \left. [X] \neq [Y], (X - g_j \cdot Y)\alpha^{(j)} = \mathbf{0} \text{ for } j = 1, \dots, N \right\} \\ &\subseteq \bigcup_{(g_j)_{j=1}^N \in G^N} \bigcup_{(r_j)_{j=1}^N \in \{1, \dots, \min(m, n)\}^N} \left\{ ((X, Y), (\alpha^{(j)})_{j=1}^N) \in (\mathbb{F}^{m \times n})^2 \times (\mathbb{F}^n)^N : [X] \neq [Y], \right. \\ &\quad \left. \text{rank}(X - g_j \cdot Y) = r_j, (X - g_j \cdot Y)\alpha^{(j)} = \mathbf{0} \text{ for } j = 1, \dots, N \right\} \\ &=: \bigcup_{(g_j)_{j=1}^N \in G^N} \bigcup_{(r_j)_{j=1}^N \in \{1, \dots, \min(m, n)\}^N} \mathcal{S}_{(g_j)_{j=1}^N, (r_j)_{j=1}^N}. \end{aligned}$$

Now for each fixed choice of $(g_j)_{j=1}^N$ and $(r_j)_{j=1}^N$, we estimate the real dimension of $\mathcal{S}_{(g_j)_{j=1}^N, (r_j)_{j=1}^N}$. Let π_1 denote the projection from $\mathcal{S}_{(g_j)_{j=1}^N, (r_j)_{j=1}^N}$ to $(\mathbb{F}^{m \times n})^2$, namely, $\pi_1((X, Y), (\alpha^{(j)})_{j=1}^N) = (X, Y)$. By Lemma 2,

$$\begin{aligned} \dim_{\mathbb{R}}(\mathcal{S}_{(g_j)_{j=1}^N, (r_j)_{j=1}^N}) &\leq \dim_{\mathbb{R}}(\pi_1(\mathcal{S}_{(g_j)_{j=1}^N, (r_j)_{j=1}^N})) \\ &\quad + \max_{(X, Y) \in \pi_1(\mathcal{S}_{(g_j)_{j=1}^N, (r_j)_{j=1}^N})} \dim_{\mathbb{R}}(\pi_1^{-1}(X, Y)). \end{aligned} \quad (9)$$

Let $j_* = \arg \min_j r_j$. We first estimate $\dim_{\mathbb{R}}(\pi_1(\mathcal{S}_{(g_j)_{j=1}^N, (r_j)_{j=1}^N}))$. Observe that

$$\begin{aligned} \pi_1(\mathcal{S}_{(g_j)_{j=1}^N, (r_j)_{j=1}^N}) &\subseteq \{(X, Y) \in (\mathbb{F}^{m \times n})^2 : [X] \neq [Y], \text{rank}(X - g_j \cdot Y) = r_j \text{ for } j = 1, \dots, N\} \\ &\subseteq \{(X, Y) \in (\mathbb{F}^{m \times n})^2 : \text{rank}(X - g_{j_*} \cdot Y) = r_{j_*}\} =: \mathcal{Q}_*. \end{aligned}$$

Consider the map $h : (X, Y) \rightarrow (X, X - g_{j_*} \cdot Y)$. It is easy to verify that h is a one-to-one semi-algebraic map. Moreover, $h(\mathcal{Q}_*) = \{(X, Z) \in (\mathbb{F}^{m \times n})^2 : \text{rank}(Z) = r_{j_*}\}$. Combining the conclusion (iv) in Lemma 1 with Lemma 3, we have

$$\begin{aligned} \dim_{\mathbb{R}}(\pi_1(\mathcal{S}_{(g_j)_{j=1}^N, (r_j)_{j=1}^N})) &\leq \dim_{\mathbb{R}}(\mathcal{Q}_*) = \dim_{\mathbb{R}}(h(\mathcal{Q}_*)) \\ &= C_{\mathbb{F}} \cdot mn + C_{\mathbb{F}} \cdot r_{j_*} (m + n - r_{j_*}). \end{aligned} \quad (10)$$

Next we estimate $\dim_{\mathbb{R}}(\pi_1^{-1}(X, Y))$. For each $(X, Y) \in \pi_1(\mathcal{S}_{(g_j)_{j=1}^N, (r_j)_{j=1}^N})$,

$$\begin{aligned} \dim_{\mathbb{R}}(\pi_1^{-1}(X, Y)) &= \dim_{\mathbb{R}}\left(\prod_{j=1}^N \{\alpha \in \mathbb{F}^n : (X - g_j \cdot Y)\alpha = \mathbf{0}\}\right) \\ &= \sum_{j=1}^N C_{\mathbb{F}} \cdot (n - r_j) \leq \sum_{j=1}^N C_{\mathbb{F}} \cdot (n - r_{j_*}) = \dim_{\mathbb{R}}((\mathbb{F}^n)^N) - C_{\mathbb{F}} \cdot N \cdot r_{j_*}. \end{aligned} \quad (11)$$

Combining (9)–(11), we obtain

$$\dim_{\mathbb{R}}(\mathcal{S}_{(g_j)_{j=1}^N, (r_j)_{j=1}^N}) \leq \dim_{\mathbb{R}}((\mathbb{F}^n)^N) + C_{\mathbb{F}} \cdot (mn + r_{j_*} (m + n - r_{j_*}) - Nr_{j_*}).$$

When $N \geq mn + m + n$, since $r_{j_*} \geq 1$, it follows that $mn + r_{j_*} (m + n - r_{j_*}) - Nr_{j_*} \leq mn + m + n - 1 - N < 0$. Therefore, $\dim_{\mathbb{R}}(\mathcal{S}_{(g_j)_{j=1}^N, (r_j)_{j=1}^N}) < \dim_{\mathbb{R}}((\mathbb{F}^n)^N)$. Since G is finite, we conclude that

$$\dim_{\mathbb{R}}(\mathcal{B}_N) = \max \dim_{\mathbb{R}}(\mathcal{S}_{(g_j)_{j=1}^N, (r_j)_{j=1}^N}) < \dim_{\mathbb{R}}((\mathbb{F}^n)^N),$$

which completes the proof. $\square$

5. Proof of Theorem 3

In this section, we prove the bi-Lipschitz property of $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ stated in Theorem 3. We first establish Lipschitz continuity, which holds under any compact group action. We then verify the lower Lipschitz property, which relies on the finiteness of the group and the structural properties of $\mathcal{M}$.

5.1. Lipschitz Property

We begin by showing that the lifting map $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ is Lipschitz. The following lemma provides the key estimate used in establishing this property.

Lemma 4. Let $\mathcal{V} = \bigoplus_{\ell=1}^L \mathcal{V}_{\ell}^{\oplus n_{\ell}}$ be an inner product space over $\mathbb{F}$, where the decomposition is unitary. For any $\vec{\alpha} \in \bigoplus_{\ell=1}^L \mathbb{F}^{n_{\ell}}$ and any $\vec{X} \in \mathcal{V}$, it holds

$$\|\Pi_{\vec{\alpha}}(\vec{X})\| \leq \|\vec{\alpha}\| \|\vec{X}\|,$$

where $\Pi_{\vec{\alpha}}$ is defined as in (2).

Proof. Write $\vec{X} = (X_1, \dots, X_L) \in \mathcal{V} = \bigoplus_{\ell=1}^L \mathcal{V}_{\ell}^{\oplus n_{\ell}}$, where $X_{\ell} = (\mathbf{x}_{\ell 1}, \dots, \mathbf{x}_{\ell n_{\ell}}) \in \mathcal{V}_{\ell}^{\oplus n_{\ell}}$. Similarly, write $\vec{\alpha} = (\alpha_1, \dots, \alpha_L) \in \bigoplus_{\ell=1}^L \mathbb{F}^{n_{\ell}}$, where $\alpha_{\ell} = (\alpha_{\ell 1}, \dots, \alpha_{\ell n_{\ell}})^{\top} \in \mathbb{F}^{n_{\ell}}$. Then

$$\begin{aligned} \|\Pi_{\vec{\alpha}}(\vec{X})\|^2 &= \sum_{\ell=1}^L \left\| \sum_{i=1}^{n_{\ell}} \alpha_{\ell i} \mathbf{x}_{\ell i} \right\|^2 \leq \sum_{\ell=1}^L \left(\sum_{i=1}^{n_{\ell}} |\alpha_{\ell i}| \|\mathbf{x}_{\ell i}\| \right)^2 \leq \sum_{\ell=1}^L \left(\sum_{i=1}^{n_{\ell}} |\alpha_{\ell i}|^2 \right) \left(\sum_{i=1}^{n_{\ell}} \|\mathbf{x}_{\ell i}\|^2 \right) \\ &= \sum_{\ell=1}^L \|\alpha_{\ell}\|^2 \|X_{\ell}\|^2 \leq \|\vec{\alpha}\|^2 \sum_{\ell=1}^L \|X_{\ell}\|^2 = \|\vec{\alpha}\|^2 \|\vec{X}\|^2. \end{aligned}$$

This completes the proof. $\square$

We now deduce the Lipschitz property of the lifting map.

Theorem 4. Let $(G, \mathcal{V} = \bigoplus_{\ell=1}^L \mathcal{V}_{\ell}^{\oplus n_{\ell}}, f)$ be a basic lifting structure. Assume that f is Lipschitz. Then for any choice of parameters $\{\vec{\alpha}^{(j)}\}_{j=1}^N \subseteq \bigoplus_{\ell=1}^L \mathbb{F}^{n_{\ell}}$, the lifting map $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ defined as in (3) is also Lipschitz.

Proof. For any $\vec{X}, \vec{Y} \in \mathcal{V}$, choose $g_* \in \arg \min_{g \in G} \|\vec{X} - g \cdot \vec{Y}\|$. Let β_f be a Lipschitz constant of f . Then

$$\begin{aligned} \|F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}(\vec{X}) - F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}(\vec{Y})\|^2 &= \sum_{j=1}^N \|f \circ \Pi_{\vec{\alpha}^{(j)}}(\vec{X}) - f \circ \Pi_{\vec{\alpha}^{(j)}}(\vec{Y})\|^2 \\ &\leq \beta_f^2 \cdot \sum_{j=1}^N \min_{g \in G} \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X} - g \cdot \vec{Y})\|^2 \leq \beta_f^2 \cdot \sum_{j=1}^N \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X} - g_* \cdot \vec{Y})\|^2 \\ &\stackrel{(*)}{\leq} \beta_f^2 \cdot \max_{j \in \{1, \dots, N\}} \|\vec{\alpha}^{(j)}\|^2 \cdot \|\vec{X} - g_* \cdot \vec{Y}\|^2 = \beta_f^2 \cdot \max_{j \in \{1, \dots, N\}} \|\vec{\alpha}^{(j)}\|^2 \cdot d^2([\vec{X}], [\vec{Y}]), \end{aligned}$$

where the inequality $(*)$ follows from Lemma 4. This shows that $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ is Lipschitz. $\square$

5.2. Lower Lipschitz Property

We now turn to the proof that the lifting map $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ is lower Lipschitz on $\mathcal{M}$. Our approach follows the methodology developed in [37,40].

5.2.1. Properties of Stabilizer Subgroups

Before proceeding with the proof, we introduce some preliminaries concerning the stabilizer subgroup, which will be used throughout the proof. Since the subsequent arguments sometimes refer to the ambient space $\mathcal{V}$ and sometimes to the reduced space $\hat{\mathcal{V}}$, we introduce the general space $\mathcal{X}$ to provide a unified perspective for discussing stabilizer subgroups. Throughout this subsection (Section 5.2.1), we assume that $\mathcal{X}$ is a finite-dimensional inner product space over $\mathbb{F}$, and that a **finite** group G acts on $\mathcal{X}$ by $\mathbb{F}$ -linear isometries. For any $\mathbf{x} \in \mathcal{X}$, define its stabilizer subgroup by $\text{Stab}(\mathbf{x}) := \{g \in G : g \cdot \mathbf{x} = \mathbf{x}\}$. We further define the quantities

$$\delta(\mathbf{x}) := \begin{cases} \min_{g \in G \setminus \text{Stab}(\mathbf{x})} \|g \cdot \mathbf{x} - \mathbf{x}\|, & \text{if } \text{Stab}(\mathbf{x}) \neq G \\ 1, & \text{if } \text{Stab}(\mathbf{x}) = G \end{cases} \quad \text{and} \quad \Delta(\mathbf{x}) := \max_{g \in G} \|g \cdot \mathbf{x} - \mathbf{x}\|. \quad (12)$$

Since G is finite, it follows from the definition that $\delta(\mathbf{x}) > 0$ for any $\mathbf{x} \in \mathcal{X}$.⁹ We now present some key lemmas about stabilizer subgroups that will be used later.

Lemma 5. Let $\mathbf{x}, \mathbf{y} \in \mathcal{X}$ satisfy $\Delta(\mathbf{y}) < \delta(\mathbf{x})$. Then $\text{Stab}(\mathbf{x} + \mathbf{y}) \subseteq \text{Stab}(\mathbf{x})$.

⁹ In fact, this is the main reason why the finiteness of G is required in the proof of the lower Lipschitz property.

Proof. On one hand, if $\text{Stab}(\mathbf{x}) = G$, the statement is trivial since $\text{Stab}(\mathbf{x} + \mathbf{y}) \subseteq G = \text{Stab}(\mathbf{x})$. On the other hand, if $\text{Stab}(\mathbf{x}) \neq G$, take any $g_0 \notin \text{Stab}(\mathbf{x})$. Then

$$\|g_0 \cdot (\mathbf{x} + \mathbf{y}) - (\mathbf{x} + \mathbf{y})\| \geq \|g_0 \cdot \mathbf{x} - \mathbf{x}\| - \|g_0 \cdot \mathbf{y} - \mathbf{y}\| \geq \delta(\mathbf{x}) - \Delta(\mathbf{y}) > 0.$$

Hence $g_0 \notin \text{Stab}(\mathbf{x} + \mathbf{y})$, and the conclusion follows. $\square$

Lemma 6. Let $p \in \mathbb{N}$, and let $(\mathbf{x}_k)_{k=1}^p \subseteq \mathcal{X}$ satisfy $\Delta(\mathbf{x}_{s+1}) < \delta(\sum_{k=1}^s \mathbf{x}_k)$ for $s = 1, \dots, p-1$. Then any $g \in \text{Stab}(\sum_{k=1}^p \mathbf{x}_k)$ satisfies $g \cdot \mathbf{x}_s = \mathbf{x}_s$ for all $s = 1, \dots, p$.

Proof. The condition $\Delta(\mathbf{x}_{s+1}) < \delta(\sum_{k=1}^s \mathbf{x}_k)$ for $s = 1, \dots, p-1$ allows us to apply Lemma 5 repeatedly, which yields the chain of inclusions

$$\text{Stab}\left(\sum_{k=1}^p \mathbf{x}_k\right) \subseteq \text{Stab}\left(\sum_{k=1}^{p-1} \mathbf{x}_k\right) \subseteq \dots \subseteq \text{Stab}(\mathbf{x}_1).$$

Therefore any $g \in \text{Stab}(\sum_{k=1}^p \mathbf{x}_k)$ fixes each partial sum. For each $s = 1, \dots, p$, we then have

$$g \cdot \mathbf{x}_s = g \cdot \left(\sum_{k=1}^s \mathbf{x}_k - \sum_{k=1}^{s-1} \mathbf{x}_k\right) = g \cdot \left(\sum_{k=1}^s \mathbf{x}_k\right) - g \cdot \left(\sum_{k=1}^{s-1} \mathbf{x}_k\right) = \sum_{k=1}^s \mathbf{x}_k - \sum_{k=1}^{s-1} \mathbf{x}_k = \mathbf{x}_s,$$

where we interpret $\sum_{k=1}^0 \mathbf{x}_k = \mathbf{0}$. This completes the proof. $\square$

Lemma 7. Let $p \in \mathbb{N}$, and let $(\mathbf{x}_k)_{k=1}^p \subseteq \mathcal{X}$ and $(c_k)_{k=1}^p \subseteq \mathbb{F}$ satisfy

$$\Delta\left(\sum_{k=1}^p (c_k - 1)\mathbf{x}_k\right) < \delta\left(\sum_{k=1}^p \mathbf{x}_k\right) \quad \text{and} \quad \Delta(\mathbf{x}_{s+1}) < \delta\left(\sum_{k=1}^s \mathbf{x}_k\right) \quad \text{for all } s = 1, \dots, p-1.$$

Then

$$\text{Stab}\left(\sum_{k=1}^p c_k \mathbf{x}_k\right) = \text{Stab}\left(\sum_{k=1}^p \mathbf{x}_k\right).$$

Proof. We first prove $\text{Stab}(\sum_{k=1}^p c_k \mathbf{x}_k) \subseteq \text{Stab}(\sum_{k=1}^p \mathbf{x}_k)$. Observe that $\sum_{k=1}^p c_k \mathbf{x}_k = \sum_{k=1}^p \mathbf{x}_k + \sum_{k=1}^p (c_k - 1)\mathbf{x}_k$. Since $\Delta(\sum_{k=1}^p (c_k - 1)\mathbf{x}_k) < \delta(\sum_{k=1}^p \mathbf{x}_k)$, Lemma 5 implies

$$\text{Stab}\left(\sum_{k=1}^p c_k \mathbf{x}_k\right) = \text{Stab}\left(\sum_{k=1}^p \mathbf{x}_k + \sum_{k=1}^p (c_k - 1)\mathbf{x}_k\right) \subseteq \text{Stab}\left(\sum_{k=1}^p \mathbf{x}_k\right).$$

For the reverse inclusion, note that $\Delta(\mathbf{x}_{s+1}) < \delta(\sum_{k=1}^s \mathbf{x}_k)$ for all $s = 1, \dots, p-1$. By Lemma 6, every $g_0 \in \text{Stab}(\sum_{k=1}^p \mathbf{x}_k)$ satisfies $g_0 \cdot \mathbf{x}_s = \mathbf{x}_s$ for all $s = 1, \dots, p$. Consequently, g_0 also fixes $\sum_{k=1}^p c_k \mathbf{x}_k$. This proves the reverse inclusion. Combining the two inclusions yields the desired equality. $\square$

Lemma 8. Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{X}$. If $\|\mathbf{x} - \mathbf{z}\| < (1/4) \cdot \delta(\mathbf{z})$ and $\|\mathbf{y} - \mathbf{z}\| < (1/4) \cdot \delta(\mathbf{z})$, then

$$\arg \min_{g \in G} \|\mathbf{x} - g \cdot \mathbf{y}\| \subseteq \text{Stab}(\mathbf{z}).$$

Proof. If $\text{Stab}(\mathbf{z}) = G$, the statement is trivial. Otherwise, suppose for contradiction that there exists $g_* \notin \text{Stab}(\mathbf{z})$ achieving the minimum distance $\|\mathbf{x} - g_* \cdot \mathbf{y}\| = \min_{g \in G} \|\mathbf{x} - g \cdot \mathbf{y}\|$. By the triangle inequality, we have

$$\begin{aligned} \|\mathbf{x} - g_* \cdot \mathbf{y}\| &\geq \|\mathbf{z} - g_* \cdot \mathbf{z}\| - \|\mathbf{x} - \mathbf{z}\| - \|g_* \cdot \mathbf{y} - g_* \cdot \mathbf{z}\| \\ &> \delta(\mathbf{z}) - \frac{1}{4}\delta(\mathbf{z}) - \frac{1}{4}\delta(\mathbf{z}) = \frac{1}{2}\delta(\mathbf{z}) > \|\mathbf{x} - \mathbf{z}\| + \|\mathbf{y} - \mathbf{z}\| \geq \|\mathbf{x} - \mathbf{y}\|, \end{aligned}$$

which contradicts the minimality of g_* . Thus $\arg \min_{g \in G} \|\mathbf{x} - g \cdot \mathbf{y}\| \subseteq \text{Stab}(\mathbf{z})$. $\square$

5.2.2. A Heuristic Proof Sketch and Its Limitations

To better understand the idea behind the formal proof, we begin with a heuristic proof sketch that follows the general structure of the final argument but ultimately breaks down due to several subtle issues. These issues necessitate a more refined analysis and motivate the modifications incorporated in the formal proof.

Consider the simple case where $\mathcal{M}$ is a linear subspace. Suppose that $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ fails to be lower Lipschitz on $\mathcal{M}$. This assumption will lead, in the formal proof, to the existence of sequences $(\vec{X}_t)_{t \in \mathbb{N}}, (\vec{Y}_t)_{t \in \mathbb{N}} \subseteq \mathcal{M}$ such that

$$\lim_{t \rightarrow \infty} \frac{\sum_{j=1}^N \min_{g \in G} \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t)\|^2}{\min_{g \in G} \|\vec{X}_t - g \cdot \vec{Y}_t\|^2} = 0, \quad (13)$$

and $\lim_{t \rightarrow \infty} \vec{X}_t = \lim_{t \rightarrow \infty} \vec{Y}_t = \vec{Z}$ for some $\vec{Z} \in \mathcal{M}$. By Lemma 8, the convergence of both sequences to $\vec{Z}$ suggests that the minimizing group elements in (13) lie in $\text{Stab}(\vec{Z})$ and $\text{Stab}(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z}))$ (this is precisely why the stabilizer subgroup analysis was developed earlier). Moreover, this convergence leads one to expect that

$$\text{Stab}(\vec{X}_t) = \text{Stab}(\vec{Y}_t) = \text{Stab}(\vec{Z}), \quad \text{Stab}(\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t)) = \text{Stab}(\Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t)) = \text{Stab}(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z})), \quad (14)$$

provided t is sufficiently large. If (14) were valid, it would follow that

$$\min_{g \in G} \|\vec{X}_t - g \cdot \vec{Y}_t\| = \min_{g \in \text{Stab}(\vec{Z})} \|\vec{X}_t - g \cdot \vec{Y}_t\| = \|\vec{X}_t - \vec{Y}_t\|,$$

and similarly,

$$\min_{g \in G} \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t)\| = \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t)\| = \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t - \vec{Y}_t)\|.$$

Let $\vec{P}_t = (\vec{X}_t - \vec{Y}_t) / \|\vec{X}_t - \vec{Y}_t\| \in \mathcal{M}$. Then, by (13), we obtain $\lim_{t \rightarrow \infty} \sum_{j=1}^N \|\Pi_{\vec{\alpha}^{(j)}}(\vec{P}_t)\|^2 = 0$. Since $(\vec{P}_t)_{t \in \mathbb{N}}$ is bounded, there exists a subsequence (still denoted by $(\vec{P}_t)_{t \in \mathbb{N}}$) such that $\vec{P}_t \rightarrow \vec{P}_\infty \in \mathcal{M}$ with $\|\vec{P}_\infty\| = 1$. Passing to the limit yields $\sum_{j=1}^N \|\Pi_{\vec{\alpha}^{(j)}}(\vec{P}_\infty)\|^2 = \lim_{t \rightarrow \infty} \sum_{j=1}^N \|\Pi_{\vec{\alpha}^{(j)}}(\vec{P}_t)\|^2 = 0$, which implies that $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}(\vec{P}_\infty) = F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}(\mathbf{0})$. However, since $[\vec{P}_\infty] \neq [\mathbf{0}]$, this contradicts the injectivity of $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ on $\mathcal{M}$.

A key step in the above heuristic argument is the validity of (14). In general, however, convergence alone is insufficient to guarantee these stabilizer identities. In the formal proof, we will therefore construct the sequences $(\vec{X}_t)_{t \in \mathbb{N}}$ and $(\vec{Y}_t)_{t \in \mathbb{N}}$ via successive refinements to ensure that the stabilizer relations in (14) hold. Finally, we note that the above heuristic argument is presented in the special case where $\mathcal{M}$ is a linear subspace. In the more general setting where $\mathcal{M}$ is a finite union of linear subspaces, additional modifications are required, which will be addressed in the formal proof.

5.2.3. Formal Proof of the Lower Lipschitz Property

We now give the formal proof of the lower Lipschitz property. The proof proceeds by contradiction and is organized around three key lemmas. The first lemma extracts, from the failure of the lower Lipschitz property, a pair of sequences exhibiting a vanishing ratio as in (13), thereby identifying a limiting configuration. The second lemma shows that, if these sequences are confined to a finite-dimensional subspace generated by an orthogonal family satisfying a suitable hierarchical separation condition, then the stabilizer relations (14) hold, which in turn lead to the failure of injectivity. The third lemma iteratively refines these sequences by building a hierarchical decomposition while progressively confining them to an associated subspace, until the rigidity required in the second lemma is attained.

Lemma 9. (Base step) Assume the hypotheses of Theorem 3 hold. Suppose the lifting map $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ is injective but fails to be lower Lipschitz on $\mathcal{M} = \cup_{i=1}^M \mathcal{L}_i$. Then there exist indices $i_1, i_2 \in \{1, \dots, M\}$, sequences $(\vec{X}_t)_{t \in \mathbb{N}} \subseteq \mathcal{L}_{i_1}$, $(\vec{Y}_t)_{t \in \mathbb{N}} \subseteq \mathcal{L}_{i_2}$ and a nonzero point $\vec{Z} \in \mathcal{L}_{i_1} \cap \mathcal{L}_{i_2}$ such that $[\vec{X}_t] \neq [\vec{Y}_t]$, $\lim_{t \rightarrow \infty} \vec{X}_t = \lim_{t \rightarrow \infty} \vec{Y}_t = \vec{Z}$ and

$$\lim_{t \rightarrow \infty} \frac{\sum_{j=1}^N \min_{g \in G} \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t)\|^2}{\min_{g \in G} \|\vec{X}_t - g \cdot \vec{Y}_t\|^2} = 0.$$

Proof. Since $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ is not lower Lipschitz on $\mathcal{M}$, there exist sequences $(\vec{X}_t)_{t \in \mathbb{N}}, (\vec{Y}_t)_{t \in \mathbb{N}} \in \mathcal{M}$ with $[\vec{X}_t] \neq [\vec{Y}_t]$ such that

$$\begin{aligned} 0 &= \lim_{t \rightarrow \infty} \frac{\|F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}(\vec{X}_t) - F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}(\vec{Y}_t)\|^2}{d^2([\vec{X}_t], [\vec{Y}_t])} = \lim_{t \rightarrow \infty} \frac{\sum_{j=1}^N \|f \circ \Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - f \circ \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t)\|^2}{\min_{g \in G} \|\vec{X}_t - g \cdot \vec{Y}_t\|^2} \\ &\geq \alpha_f^2 \cdot \lim_{t \rightarrow \infty} \frac{\sum_{j=1}^N \min_{g \in G} \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t)\|^2}{\min_{g \in G} \|\vec{X}_t - g \cdot \vec{Y}_t\|^2} \geq 0, \end{aligned}$$

where $\alpha_f > 0$ is the lower Lipschitz constant of f . Thus we have

$$\lim_{t \rightarrow \infty} \frac{\sum_{j=1}^N \min_{g \in G} \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t)\|^2}{\min_{g \in G} \|\vec{X}_t - g \cdot \vec{Y}_t\|^2} = 0. \quad (15)$$

We now perform a series of adjustments to the sequences $(\vec{X}_t)_{t \in \mathbb{N}}$ and $(\vec{Y}_t)_{t \in \mathbb{N}}$ that preserve (15), after which the adjusted sequences satisfy the remaining conclusions of the lemma.

Since $\mathcal{M}$ is finite union of linear subspace, there exists at least one subspace $\mathcal{L}_{i_1}$ containing infinitely many elements of $(\vec{X}_t)_{t \in \mathbb{N}}$. Without loss of generality, we may assume that $(\vec{X}_t)_{t \in \mathbb{N}} \subseteq \mathcal{L}_{i_1}$. Note that scaling $(\vec{X}_t, \vec{Y}_t)$ does not affect (15). Hence we further normalize the sequences so that $\max(\|\vec{X}_t\|, \|\vec{Y}_t\|) = 1$. It follows that $\min_{g \in G} \|\vec{X}_t - g \cdot \vec{Y}_t\| \leq 2$. Combining this with (15), we have

$$\lim_{t \rightarrow \infty} \sum_{j=1}^N \min_{g \in G} \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t)\|^2 = 0. \quad (16)$$

Moreover, the normalization $\max(\|\vec{X}_t\|, \|\vec{Y}_t\|) = 1$ ensures that both sequences $(\vec{X}_t)_{t \in \mathbb{N}}$ and $(\vec{Y}_t)_{t \in \mathbb{N}}$ are bounded. Therefore, there exist the subsequences along which $(\vec{X}_t)_{t \in \mathbb{N}}$ and $(\vec{Y}_t)_{t \in \mathbb{N}}$ converge, respectively. By passing to convergent subsequences if necessary, we may assume that

$$\vec{X}_t \rightarrow \vec{X}_\infty \in \mathcal{L}_{i_1}, \quad \vec{Y}_t \rightarrow \vec{Y}_\infty \in \mathcal{M} \quad \text{as } t \rightarrow \infty.$$

Since $\max(\|\vec{X}_t\|, \|\vec{Y}_t\|) = 1$, $\vec{X}_\infty$ and $\vec{Y}_\infty$ can not be zero simultaneously. By the continuity of $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ (which follows from its Lipschitz property established in Theorem 4), we obtain

$$\begin{aligned} 0 &\leq \|F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}(\vec{X}_\infty) - F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}(\vec{Y}_\infty)\|^2 = \lim_{t \rightarrow \infty} \|F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}(\vec{X}_t) - F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}(\vec{Y}_t)\|^2 \\ &\leq \beta_f^2 \cdot \lim_{t \rightarrow \infty} \sum_{j=1}^N \min_{g \in G} \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t)\|^2 = 0, \end{aligned} \quad (17)$$

where β_f is the Lipschitz constant of f , and the last equality follows from (16). The injectivity of $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ then implies the existence of some $g_* \in G$ such that $\vec{X}_\infty = g_* \cdot \vec{Y}_\infty$. Since $\vec{X}_\infty$ and $\vec{Y}_\infty$ cannot both be zero, this implies that $\vec{X}_\infty \neq \mathbf{0}$. Again, since $\mathcal{M}$ is a finite union, there exists a subspace $\mathcal{L}_{i_2}$ containing infinitely many elements of the sequence $(g_* \cdot \vec{Y}_t)_{t \in \mathbb{N}}$. Without loss of generality, we may assume that $(g_* \cdot \vec{Y}_t)_{t \in \mathbb{N}} \subseteq \mathcal{L}_{i_2}$. It follows that $\vec{X}_\infty = g_* \cdot \vec{Y}_\infty \in \mathcal{L}_{i_1} \cap \mathcal{L}_{i_2} \setminus \{\mathbf{0}\}$.

Setting $\vec{Z} = \vec{X}_\infty$ and replacing $\vec{Y}_t$ with $g_* \cdot \vec{Y}_t$, we obtain the desired conclusion. $\square$

Lemma 10. Assume the hypotheses of Theorem 3 hold. Fix a linear subspace $\mathcal{L}_{i_0} \subseteq \mathcal{M}$, and let $p \in \mathbb{N}$, $p \leq \dim_{\mathbb{F}}(\mathcal{L}_{i_0})$. Suppose $(\vec{Z}_k)_{k=1}^p \subseteq \mathcal{L}_{i_0} \setminus \{0\}$ is an orthogonal sequence satisfying

$$\Delta(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_{s+1})) < \delta \left(\sum_{k=1}^s \Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_k) \right) \quad \text{for all } j \in \{1, \dots, N\}, s \in \{1, \dots, p-1\}, \quad (18)$$

and set $\vec{Z} := \sum_{k=1}^p \vec{Z}_k$. If there exist sequences $(\vec{X}_t)_{t \in \mathbb{N}}, (\vec{Y}_t)_{t \in \mathbb{N}} \subseteq \text{span}\{\vec{Z}_k\}_{k=1}^p$ such that $\lim_{t \rightarrow \infty} \vec{X}_t = \lim_{t \rightarrow \infty} \vec{Y}_t = \vec{Z}$ and

$$\lim_{t \rightarrow \infty} \frac{\sum_{j=1}^N \min_{g \in G} \left\| \Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t) \right\|^2}{\min_{g \in G} \left\| \vec{X}_t - g \cdot \vec{Y}_t \right\|^2} = 0,$$

then $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ is not injective on $\mathcal{M}$.

Proof. Write the sequences $\vec{X}_t$ and $\vec{Y}_t$ in the basis $\{\vec{Z}_k\}_{k=1}^p$ as $\vec{X}_t = \sum_{k=1}^p c_{tk} \vec{Z}_k$ and $\vec{Y}_t = \sum_{k=1}^p d_{tk} \vec{Z}_k$, where $c_{tk}, d_{tk} \in \mathbb{F}$. Consequently, for each $j = 1, \dots, N$,

$$\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) = \sum_{k=1}^p c_{tk} \Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_k) \quad \text{and} \quad \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t) = \sum_{k=1}^p d_{tk} \Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_k).$$

We first show that for all $j = 1, \dots, N$, it holds

$$\text{Stab}(\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t)) = \text{Stab}(\Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t)) = \text{Stab}(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z})) \quad (19)$$

provided t is sufficiently large. To see this, since $\vec{X}_t, \vec{Y}_t \rightarrow \vec{Z}$, we have $\sum_{k=1}^p \Pi_{\vec{\alpha}^{(j)}}((c_{tk} - 1)\vec{Z}_k) \rightarrow 0$ and $\sum_{k=1}^p \Pi_{\vec{\alpha}^{(j)}}((d_{tk} - 1)\vec{Z}_k) \rightarrow 0$ for each j , as $t \rightarrow \infty$. By definition (12), it holds $\delta(\sum_{k=1}^p \Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_k)) > 0$ for each j . Consequently, for sufficiently large t ,

$$\Delta\left(\sum_{k=1}^p \Pi_{\vec{\alpha}^{(j)}}((c_{tk} - 1)\vec{Z}_k)\right) < \delta\left(\sum_{k=1}^p \Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_k)\right)$$

and the same inequality holds for the d_{tk} 's. Combining this with condition (18) and Lemma 7, we obtain (19).

Next we show that for each $j = 1, \dots, N$,

$$\arg \min_{g \in G} \left\| \Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t) \right\| \subseteq \text{Stab}(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z})) \quad (20)$$

provided t is sufficiently large. Since both $\vec{X}_t$ and $\vec{Y}_t$ converge to $\vec{Z}$, for each j we have

$$\left\| \Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - \Pi_{\vec{\alpha}^{(j)}}(\vec{Z}) \right\| < \frac{1}{4} \delta(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z})) \quad \text{and} \quad \left\| \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t) - \Pi_{\vec{\alpha}^{(j)}}(\vec{Z}) \right\| < \frac{1}{4} \delta(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z}))$$

provided t is sufficiently large. Then by Lemma 8, we obtain (20).

Combining (19) and (20), we obtain, for all sufficiently large t and each $j = 1, \dots, N$,

$$\min_{g \in G} \left\| \Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t) \right\| = \left\| \Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t) \right\|.$$

Define the normalized difference $\vec{P}_t = (\vec{X}_t - \vec{Y}_t) / \|\vec{X}_t - \vec{Y}_t\| \in \mathcal{L}_{i_0}$. Since the unit sphere in $\text{span}\{\vec{Z}_k\}_{k=1}^p$ is compact, there exists a subsequence of $(\vec{P}_t)_{t \in \mathbb{N}}$ that converges to some $\vec{P}_\infty \in \mathcal{L}_{i_0}$ with $\|\vec{P}_\infty\| = 1$. Taking limits gives

$$\begin{aligned} 0 &\leq \sum_{j=1}^N \left\| \Pi_{\vec{\alpha}^{(j)}}(\vec{P}_\infty) \right\|^2 = \lim_{t \rightarrow \infty} \frac{\sum_{j=1}^N \left\| \Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t) \right\|^2}{\left\| \vec{X}_t - \vec{Y}_t \right\|^2} \\ &= \lim_{t \rightarrow \infty} \frac{\sum_{j=1}^N \min_{g \in G} \left\| \Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t) \right\|^2}{\left\| \vec{X}_t - \vec{Y}_t \right\|^2} \\ &\leq \lim_{t \rightarrow \infty} \frac{\sum_{j=1}^N \min_{g \in G} \left\| \Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t) \right\|^2}{\min_{g \in G} \left\| \vec{X}_t - g \cdot \vec{Y}_t \right\|^2} = 0. \end{aligned}$$

Hence $\Pi_{\vec{\alpha}^{(j)}}(\vec{P}_\infty) = \mathbf{0}$ for all j , so that $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}(\vec{P}_\infty) = F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}(\mathbf{0})$. Since $[\vec{P}_\infty] \neq [\mathbf{0}]$ and $\vec{P}_\infty \in \mathcal{L}_{i_0} \subseteq \mathcal{M}$, it follows that $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ is not injective on $\mathcal{M}$. $\square$

Lemma 11. (Induction step) Assume the hypotheses of Theorem 3 hold. Fix two linear subspaces $\mathcal{L}_{i_1}, \mathcal{L}_{i_2} \subseteq \mathcal{M}$. Let $p \in \mathbb{N}$, $p \leq \dim_{\mathbb{F}}(\mathcal{L}_{i_1})$, and let $(\vec{Z}_k)_{k=1}^p \subseteq \mathcal{L}_{i_1} \cap \mathcal{L}_{i_2} \setminus \{\mathbf{0}\}$ be an orthogonal sequence satisfying

$$\Delta(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_{s+1})) < \delta \left(\sum_{k=1}^s \Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_k) \right) \quad \text{and} \quad \Delta(\vec{Z}_{s+1}) < \delta \left(\sum_{k=1}^s \vec{Z}_k \right)$$

for all $j \in \{1, \dots, N\}, s \in \{1, \dots, p-1\}$. Set $\vec{Z} := \sum_{k=1}^p \vec{Z}_k$, and suppose sequences $(\vec{X}_t)_{t \in \mathbb{N}} \subseteq \mathcal{L}_{i_1}, (\vec{Y}_t)_{t \in \mathbb{N}} \subseteq \mathcal{L}_{i_2}$ exist with $\lim_{t \rightarrow \infty} \vec{X}_t = \lim_{t \rightarrow \infty} \vec{Y}_t = \vec{Z}$ and

$$\lim_{t \rightarrow \infty} \frac{\sum_{j=1}^N \min_{g \in G} \left\| \Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t) \right\|^2}{\min_{g \in G} \left\| \vec{X}_t - g \cdot \vec{Y}_t \right\|^2} = 0.$$

If $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ is injective on $\mathcal{M}$, then $p < \dim_{\mathbb{F}}(\mathcal{L}_{i_1})$. Moreover, there exists a new linear subspace $\mathcal{L}_{i_2}^{\text{new}} \subseteq \mathcal{M}$ containing $(\vec{Z}_k)_{k=1}^p$, and a nonzero $\vec{Z}_{p+1} \in \mathcal{L}_{i_1} \cap \mathcal{L}_{i_2}^{\text{new}} \cap (\text{span}\{\vec{Z}_k\}_{k=1}^p)^\perp$ such that

$$\Delta(\vec{Z}_{p+1}) < \delta \left(\sum_{k=1}^p \vec{Z}_k \right), \quad \Delta(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_{p+1})) < \delta \left(\sum_{k=1}^p \Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_k) \right) \quad \text{for all } j = 1, \dots, N.$$

Furthermore, there exist sequences $(\vec{X}_t^{\text{new}})_{t \in \mathbb{N}} \subseteq \mathcal{L}_{i_1}, (\vec{Y}_t^{\text{new}})_{t \in \mathbb{N}} \subseteq \mathcal{L}_{i_2}^{\text{new}}$ that converge to $\vec{Z} + \vec{Z}_{p+1}$ and satisfy the vanishing ratio property

$$\lim_{t \rightarrow \infty} \frac{\sum_{j=1}^N \min_{g \in G} \left\| \Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t^{\text{new}}) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t^{\text{new}}) \right\|^2}{\min_{g \in G} \left\| \vec{X}_t^{\text{new}} - g \cdot \vec{Y}_t^{\text{new}} \right\|^2} = 0.$$

Proof. Write $\vec{X}_t = \sum_{k=1}^p c_{tk} \vec{Z}_k + \vec{V}_t$ and $\vec{Y}_t = \sum_{k=1}^p d_{tk} \vec{Z}_k + \vec{W}_t$, where $c_{tk}, d_{tk} \in \mathbb{F}$, $\vec{V}_t \in \mathcal{L}_{i_1} \cap (\text{span}\{\vec{Z}_k\}_{k=1}^p)^\perp$ and $\vec{W}_t \in \mathcal{L}_{i_2} \cap (\text{span}\{\vec{Z}_k\}_{k=1}^p)^\perp$. If $\vec{V}_t = \vec{W}_t = \mathbf{0}$, then by Lemma 10, $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ is not injective on $\mathcal{M}$, contradicting the assumption. Thus $\max(\|\vec{V}_t\|, \|\vec{W}_t\|) > 0$. Define

$$\vec{U}_t := \vec{Y}_t - \vec{X}_t + \vec{V}_t = \sum_{k=1}^p (d_{tk} - c_{tk}) \vec{Z}_k + \vec{W}_t \in \mathcal{L}_{i_2}.$$

Clearly, $\|\vec{U}_t\| \geq \|\vec{W}_t\|$. Let

$$\eta_t := \frac{\epsilon}{\max(\|\vec{V}_t\|, \|\vec{U}_t\|)} \min \left\{ \frac{1}{4} \delta(\vec{Z}), \frac{1}{4} \min_{j \in \{1, \dots, N\}} \frac{\delta(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z}))}{\|\vec{\alpha}^{(j)}\|} \right\} =: \frac{1}{\max(\|\vec{V}_t\|, \|\vec{U}_t\|)} \cdot c_{\#}, \quad (21)$$

where $\epsilon \in (0, 1)$ is a fixed number and $c_{\#} > 0$ denotes the resulting constant independent of $t \in \mathbb{N}$. Define

$$\vec{X}_t^{\text{new}} := \sum_{k=1}^p \vec{Z}_k + \eta_t \vec{V}_t = \vec{Z} + \eta_t \vec{V}_t \in \mathcal{L}_{i_1} \quad \text{and} \quad \vec{Y}_t^{\text{gen}} := \sum_{k=1}^p \vec{Z}_k + \eta_t \vec{U}_t = \vec{Z} + \eta_t \vec{U}_t \in \mathcal{L}_{i_2}.$$

We assert that

$$\min_{g \in G} \|\vec{X}_t^{\text{new}} - g \cdot \vec{Y}_t^{\text{gen}}\| \geq \eta_t \min_{g \in G} \|\vec{X}_t - g \cdot \vec{Y}_t\| \quad (22)$$

and

$$\min_{g \in G} \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t^{\text{new}}) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t^{\text{gen}})\| \leq \eta_t \min_{g \in G} \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t)\| \quad (23)$$

for all $j = 1, \dots, N$, provided t is sufficiently large. From (22) and (23), it follows that

$$\begin{aligned} 0 &\leq \lim_{t \rightarrow \infty} \frac{\sum_{j=1}^N \min_{g \in G} \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t^{\text{new}}) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t^{\text{gen}})\|^2}{\min_{g \in G} \|\vec{X}_t^{\text{new}} - g \cdot \vec{Y}_t^{\text{gen}}\|^2} \\ &\leq \lim_{t \rightarrow \infty} \frac{\sum_{j=1}^N \min_{g \in G} \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t)\|^2}{\min_{g \in G} \|\vec{X}_t - g \cdot \vec{Y}_t\|^2} = 0. \end{aligned} \quad (24)$$

We now analyze the limiting properties of the sequences $(\vec{X}_t^{\text{new}})_{t \in \mathbb{N}}$ and $(\vec{Y}_t^{\text{gen}})_{t \in \mathbb{N}}$. By the definition (21) of η_t , the sequences $(\eta_t \vec{V}_t)_{t \in \mathbb{N}}$ and $(\eta_t \vec{U}_t)_{t \in \mathbb{N}}$ both are bounded. Thus there exist subsequences along which $(\eta_t \vec{V}_t)_{t \in \mathbb{N}}$ and $(\eta_t \vec{U}_t)_{t \in \mathbb{N}}$ converge, respectively. Passing to a subsequence if necessary, assume

$$\eta_t \vec{V}_t \rightarrow \vec{V}_{\infty} \in \mathcal{L}_{i_1} \cap (\text{span}\{\vec{Z}_k\}_{k=1}^p)^{\perp}, \quad \eta_t \vec{U}_t \rightarrow \vec{U}_{\infty} \in \mathcal{L}_{i_2}. \quad (25)$$

By the construction (21) of η_t , $\max(\|\eta_t \vec{V}_t\|, \|\eta_t \vec{U}_t\|) = c_{\#} > 0$ for the constant $c_{\#}$ independent of $t \in \mathbb{N}$. Hence $\vec{V}_{\infty}$ and $\vec{U}_{\infty}$ are not simultaneously zero. As in the analysis of (17) in the proof of Lemma 10, it follows from (24) that $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}(\vec{Z} + \vec{V}_{\infty}) = F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}(\vec{Z} + \vec{U}_{\infty})$. The injectivity of $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ implies the existence of some $g_* \in G$ such that $\vec{Z} + \vec{V}_{\infty} = g_* \cdot \vec{Z} + g_* \cdot \vec{U}_{\infty}$. We claim that $g_* \in \text{Stab}(\vec{Z})$, otherwise by the definition (21) of η_t we have

$$\frac{1}{2} \delta(\vec{Z}) > 2 \max(\|\vec{V}_{\infty}\|, \|\vec{U}_{\infty}\|) \geq \|\vec{V}_{\infty} - g_* \cdot \vec{U}_{\infty}\| = \|g_* \cdot \vec{Z} - \vec{Z}\| \geq \delta(\vec{Z}),$$

which is a contradiction. Thus we have $\vec{V}_{\infty} = g_* \cdot \vec{U}_{\infty}$. Since $\vec{V}_{\infty}$ and $\vec{U}_{\infty}$ are not simultaneously zero, it follows that $\vec{V}_{\infty}$ and $\vec{U}_{\infty}$ are both not zero. Then we have $p < \dim_{\mathbb{F}}(\mathcal{L}_{i_1})$. Otherwise $\vec{V}_t = \mathbf{0}$ for all t , which implies $\vec{V}_{\infty} = \mathbf{0}$.

Since g_* acts $\mathbb{F}$ -linearly, the set $g_* \cdot \mathcal{L}_{i_2}$ is a linear subspace. Because $\mathcal{M}$ is G -invariant, it follows that $g_* \cdot \mathcal{L}_{i_2} \subseteq \mathcal{M}$. Hence, there exists a linear subspace $\mathcal{L}_{i_2}^{\text{new}} \subseteq \mathcal{M}$ such that $g_* \cdot \mathcal{L}_{i_2} \subseteq \mathcal{L}_{i_2}^{\text{new}}$. By the conditions $\Delta(\vec{Z}_{s+1}) < \delta(\sum_{k=1}^s \vec{Z}_k)$ for $s = 1, \dots, p-1$, it follows from Lemma 6 that if $g \in \text{Stab}(\vec{Z}) = \text{Stab}(\sum_{k=1}^p \vec{Z}_k)$ then $g \in \text{Stab}(\vec{Z}_s)$ for all $s = 1, \dots, p$. In particular, this implies $(\vec{Z}_k)_{k=1}^p = g_* \cdot (\vec{Z}_k)_{k=1}^p \subseteq \mathcal{L}_{i_2}^{\text{new}}$. Moreover, by (25), we have

$$\vec{V}_{\infty} = g_* \cdot \vec{U}_{\infty} \in \mathcal{L}_{i_1} \cap \mathcal{L}_{i_2}^{\text{new}} \cap (\text{span}\{\vec{Z}_k\}_{k=1}^p)^{\perp} \setminus \{\mathbf{0}\}.$$

Let $\vec{Z}_{p+1} = \vec{V}_\infty$ and $\vec{Y}_t^{\text{new}} = g_* \cdot \vec{Y}_t^{\text{gen}} \in \mathcal{L}_{i_2}^{\text{new}}$. Note that replacing $\vec{Y}_t^{\text{gen}}$ with $\vec{Y}_t^{\text{new}}$ does not affect the vanishing ratio property in (24). Furthermore, by the definition (21) of η_t , we have

$$\Delta(\vec{Z}_{p+1}) \leq 2\|\vec{V}_\infty\| = 2 \lim_{t \rightarrow \infty} \|\vec{V}_t \eta_t\| \leq \frac{1}{2} \delta(\vec{Z}) < \delta(\vec{Z})$$

and

$$\Delta(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_{p+1})) \leq 2\|\vec{Z}_{p+1}\| \cdot \|\vec{\alpha}^{(j)}\| \leq \frac{1}{2} \delta(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z})) < \delta(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z})), \quad \text{for } j = 1, \dots, N,$$

where we use the inequality $\|\Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_{p+1})\| \leq \|\vec{Z}_{p+1}\| \cdot \|\vec{\alpha}^{(j)}\|$, which follows from Lemma 4. Thus we obtain the conclusion.

The only remaining task is to prove the assertions (22) and (23). We begin with (22). By the definition (21) of η_t ,

$$\|\vec{X}_t^{\text{new}} - \vec{Z}\| = \|\eta_t \vec{V}_t\| < \frac{1}{4} \delta(\vec{Z}) \quad \text{and} \quad \|\vec{Y}_t^{\text{gen}} - \vec{Z}\| = \|\eta_t \vec{U}_t\| < \frac{1}{4} \delta(\vec{Z}).$$

Lemma 8 then implies that

$$\arg \min_{g \in G} \|\vec{X}_t^{\text{new}} - g \cdot \vec{Y}_t^{\text{gen}}\| \subseteq \text{Stab}(\vec{Z}). \quad (26)$$

Since $\vec{X}_t, \vec{Y}_t \rightarrow \vec{Z}$ as $t \rightarrow \infty$, we have $\Delta(\sum_{k=1}^p (c_{tk} - 1) \vec{Z}_k) \leq 2\|\sum_{k=1}^p (c_{tk} - 1) \vec{Z}_k\| \rightarrow 0$ and $\Delta(\sum_{k=1}^p (d_{tk} - 1) \vec{Z}_k) \leq 2\|\sum_{k=1}^p (d_{tk} - 1) \vec{Z}_k\| \rightarrow 0$ as $t \rightarrow \infty$. Hence, for t sufficiently large,

$$\Delta\left(\sum_{k=1}^p (c_{tk} - 1) \vec{Z}_k\right) < \delta\left(\sum_{k=1}^p \vec{Z}_k\right) \quad \text{and} \quad \Delta\left(\sum_{k=1}^p (d_{tk} - 1) \vec{Z}_k\right) < \delta\left(\sum_{k=1}^p \vec{Z}_k\right).$$

Together with the condition $\Delta(\vec{Z}_{s+1}) < \delta(\sum_{k=1}^s \vec{Z}_k)$ for each $s = 1, \dots, p-1$, Lemma 7 yields

$$\text{Stab}\left(\sum_{k=1}^p c_{tk} \vec{Z}_k\right) = \text{Stab}\left(\sum_{k=1}^p d_{tk} \vec{Z}_k\right) = \text{Stab}\left(\sum_{k=1}^p \vec{Z}_k\right). \quad (27)$$

Choose $g_0 \in \arg \min_{g \in G} \|\vec{X}_t^{\text{new}} - g \cdot \vec{Y}_t^{\text{gen}}\|$. By (26) and (27), we have

$$g_0 \in \text{Stab}(\vec{Z}) = \text{Stab}\left(\sum_{k=1}^p c_{tk} \vec{Z}_k\right) = \text{Stab}\left(\sum_{k=1}^p d_{tk} \vec{Z}_k\right).$$

Therefore,

$$\begin{aligned} \min_{g \in G} \|\vec{X}_t^{\text{new}} - g \cdot \vec{Y}_t^{\text{gen}}\| &= \|\vec{X}_t^{\text{new}} - g_0 \cdot \vec{Y}_t^{\text{gen}}\| = \eta_t \|\vec{V}_t - g_0 \cdot \vec{U}_t\| \\ &= \eta_t \left\| \vec{V}_t - g_0 \cdot \left(\sum_{k=1}^p (d_{tk} - c_{tk}) \vec{Z}_k + \vec{W}_t \right) \right\| = \eta_t \left\| \sum_{k=1}^p c_{tk} \vec{Z}_k + \vec{V}_t - g_0 \cdot \left(\sum_{k=1}^p d_{tk} \vec{Z}_k + \vec{W}_t \right) \right\| \\ &= \eta_t \|\vec{X}_t - g_0 \cdot \vec{Y}_t\| \geq \eta_t \min_{g \in G} \|\vec{X}_t - g \cdot \vec{Y}_t\|, \end{aligned}$$

which completes the proof of (22).

Next we prove (23). Since $\vec{X}_t, \vec{Y}_t \rightarrow \vec{Z}$, we have that for all sufficiently large t and each $j = 1, \dots, N$,

$$\|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - \Pi_{\vec{\alpha}^{(j)}}(\vec{Z})\| < \frac{1}{4} \delta(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z})) \quad \text{and} \quad \|\Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t) - \Pi_{\vec{\alpha}^{(j)}}(\vec{Z})\| < \frac{1}{4} \delta(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z})).$$

By Lemma 8, this implies that

$$\arg \min_{g \in G} \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t)\| \subseteq \text{Stab}(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z})) \quad (28)$$

for sufficiently large t and each $j = 1, \dots, N$. Meanwhile, the convergence $\vec{X}_t, \vec{Y}_t \rightarrow \vec{Z}$ implies that for each j and all sufficiently large t ,

$$\Delta\left(\sum_{k=1}^p (c_{tk} - 1)\Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_k)\right) < \delta(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z})) \quad \text{and} \quad \Delta\left(\sum_{k=1}^p (d_{tk} - 1)\Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_k)\right) < \delta(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z})).$$

Combining these estimates with the condition that $\Delta(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_{s+1})) < \delta(\sum_{k=1}^s \Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_k))$ for $s = 1, \dots, p-1$, and applying Lemma 7, we deduce that for all $j = 1, \dots, N$ and sufficiently large t ,

$$\text{Stab}\left(\sum_{k=1}^p c_{tk}\Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_k)\right) = \text{Stab}\left(\sum_{k=1}^p d_{tk}\Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_k)\right) = \text{Stab}\left(\sum_{k=1}^p \Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_k)\right). \quad (29)$$

For each j , choose $\tilde{g}_0 \in \arg \min_{g \in G} \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t)\|$. By (28) and (29),

$$\tilde{g}_0 \in \text{Stab}(\Pi_{\vec{\alpha}^{(j)}}(\vec{Z})) = \text{Stab}\left(\sum_{k=1}^p c_{tk}\Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_k)\right) = \text{Stab}\left(\sum_{k=1}^p d_{tk}\Pi_{\vec{\alpha}^{(j)}}(\vec{Z}_k)\right).$$

Then we have

$$\begin{aligned} \eta_t \min_{g \in G} \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t)\| &= \eta_t \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t) - \tilde{g}_0 \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t)\| \\ &= \eta_t \|\Pi_{\vec{\alpha}^{(j)}}(\vec{V}_t) - \tilde{g}_0 \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{U}_t)\| = \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t^{\text{new}}) - \tilde{g}_0 \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t^{\text{gen}})\| \\ &\geq \min_{g \in G} \|\Pi_{\vec{\alpha}^{(j)}}(\vec{X}_t^{\text{new}}) - g \cdot \Pi_{\vec{\alpha}^{(j)}}(\vec{Y}_t^{\text{gen}})\|, \end{aligned}$$

which completes the proof of (23). $\square$

With these key lemmas established, we are now in a position to prove Theorem 3.

Proof of Theorem 3. The Lipschitz continuity of $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ was established in Theorem 4. It remains to show that $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ is lower Lipschitz on $\mathcal{M}$. We argue by contradiction. Suppose that $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ fails to be lower Lipschitz on $\mathcal{M}$. Then, by Lemma 9, the assumptions of Lemma 11 are satisfied for $p = 1$. Applying Lemma 11 iteratively, one can construct an orthogonal sequence that eventually spans the entire subspace $\mathcal{L}_{i_1}$ for some i_1 . However, this contradicts the conclusion of Lemma 11, which asserts that p must always remain strictly smaller than $\dim_{\mathbb{R}}(\mathcal{L}_{i_1})$ under the injectivity assumption. Therefore, the assumption that $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ fails to be lower Lipschitz cannot hold. We conclude that $F_{\{\vec{\alpha}^{(j)}\}_{j=1}^N}$ is indeed lower Lipschitz on $\mathcal{M}$. $\square$

6. Conclusions and Discussion

In this paper, we have developed a lifting framework for constructing group-invariant embeddings that are orbit-separating, stable, and computationally efficient. Starting from a group-invariant map on a low-dimensional reduced space, the framework lifts it via a finite family of parameterized linear maps, providing a unified perspective that connects classical problems such as phase retrieval, permutation-invariant learning, and low-rank matrix recovery.

We established the effectiveness of this framework for finite groups. We proved that a generic collection of elementary lifting invariants yields an orbit-separating embedding on any G -invariant semi-algebraic set $\mathcal{M}$ when the number of parameters satisfies $N \geq 2 \dim_{\mathbb{R}}(\mathcal{M})/C_{\mathbb{F}} + 1$. For the special case $\mathcal{V} = (\mathbb{F}^m)^{\oplus n}$ with reduced space $\hat{\mathcal{V}} = \mathbb{F}^m$, we obtained the sharper bound $N \geq mn + m + n$. Moreover, we showed that any orbit-separating lifting embedding is automatically bi-Lipschitz, ensuring stability against perturbations, and this property extends naturally to sparse regimes. These results demonstrate that the lifting framework provides a systematic and computationally efficient approach to constructing parsimonious and stable group-invariant representations.

Several directions remain open for future investigation. Extending the theoretical guarantees from finite groups to general compact groups, particularly to continuous groups such as $O(m)$ or $U(m)$, would significantly broaden the applicability of the framework. Additionally, developing efficient algorithms for selecting the parameter vectors $\{\vec{\alpha}^{(j)}\}_{j=1}^N$ in practice, beyond generic random choices, could further enhance the computational performance of the resulting embeddings. Finally, exploring the connections between this lifting framework and modern equivariant neural network architectures may yield new insights into the design of group-equivariant models that are both expressive and stable.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: The author gratefully acknowledges Prof. Zhiqiang Xu for his guidance and valuable advice during the preparation of this work

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wang, Y.; Xu, Z. Generalized phase retrieval: Measurement number, matrix recovery and beyond. *Appl. Comput. Harmon. Anal.* **2019**, *47*, 423–446. <https://doi.org/10.1016/j.acha.2017.09.003>.
2. Huang, M.; Rong, Y.; Wang, Y.; Xu, Z. Almost everywhere generalized phase retrieval. *Appl. Comput. Harmon. Anal.* **2021**, *50*, 16–33. <https://doi.org/10.1016/j.acha.2020.08.002>.
3. Balan, R.; Casazza, P.; Edidin, D. On signal reconstruction without phase. *Appl. Comput. Harmon. Anal.* **2006**, *20*, 345–356. <https://doi.org/10.1016/j.acha.2005.07.001>.
4. Conca, A.; Edidin, D.; Hering, M.; Vinzant, C. An algebraic characterization of injectivity in phase retrieval. *Appl. Comput. Harmon. Anal.* **2015**, *38*, 346–356. <https://doi.org/10.1016/j.acha.2014.06.005>.
5. Cohen, T.S.; Geiger, M.; Weiler, M. A general theory of equivariant CNNs on homogeneous spaces. In Proceedings of the 33rd Conference on Advances in Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada, 8–14 December 2019.
6. Cohen, T.; Weiler, M.; Kicanaoglu, B.; Welling, M. Gauge equivariant convolutional networks and the icosahedral CNN. In Proceedings of the 36th International Conference on Machine Learning (ICML 2019), Long Beach, California, USA, 9–15 June 2019; pp. 1321–1330.
7. Weiler, M.; Cesa, G. General E(2)-equivariant steerable CNNs. In Proceedings of the 33rd Conference on Advances in Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada, 8–14 December 2019.
8. Cohen, T.; Welling, M. Group equivariant convolutional networks. In Proceedings of the 33rd International Conference on Machine Learning (ICML 2016), New York, NY, USA, 19–24 June 2016; pp. 2990–2999.
9. Chidester, B.; Zhou, T.; Do, M.N.; Ma, J. Rotation equivariant and invariant neural networks for microscopy image analysis. *Bioinformatics* **2019**, *35*, i530–i537. <https://doi.org/10.1093/bioinformatics/btz353>.
10. Satorras, V.G.; Hoogeboom, E.; Welling, M. E(n) equivariant graph neural networks. In Proceedings of the 38th International Conference on Machine Learning (ICML 2021), Virtual Event, 18–24 July 2021; pp. 9323–9332.
11. Fuchs, F.; Worrall, D.; Fischer, V.; Welling, M. SE(3)-transformers: 3D roto-translation equivariant attention networks. In Proceedings of the 34th Conference on Neural Information Processing Systems (NeurIPS 2020), Vancouver, Canada, 6–12 December 2020; pp. 1970–1981.
12. Thomas, N.; Smidt, T.; Kearnes, S.; Yang, L.; Li, L.; Kohlhoff, K.; Riley, P. Tensor field networks: Rotation-and translation-equivariant neural networks for 3D point clouds. *arXiv preprint arXiv:1802.08219* **2018**.
13. Behler, J.; Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401. <https://doi.org/10.1103/PhysRevLett.98.146401>.

14. Gasteiger, J.; Groß, J.; Günnemann, S. Directional Message Passing for Molecular Graphs. *arXiv preprint arXiv:2003.03123* **2020**.
15. Gilmer, J.; Schoenholz, S.S.; Riley, P.F.; Vinyals, O.; Dahl, G.E. Neural message passing for quantum chemistry. In Proceedings of the 34th International Conference on Machine Learning (ICML 2017), Sydney, Australia, 6–11 August 2017; pp. 1263–1272.
16. Blum-Smith, B.; Derksen, H.; Mixon, D.G.; Qaddura, Y.; Vose, B. Estimating the Euclidean distortion of an orbit space. *arXiv preprint arXiv:2506.04425* **2025**.
17. Qaddura, Y. A max filtering local stability theorem with application to weighted phase retrieval and cryo-EM. *Appl. Comput. Harmon. Anal.* **2025**, p. 101821. <https://doi.org/https://doi.org/10.1016/j.acha.2025.101821>.
18. He, P.; Heile, Z.; Tran, J.; Wang, A.; Chand, S. Rigid-invariant sliced Wasserstein via independent embeddings. *arXiv preprint arXiv:2510.10233* **2025**.
19. Wang, Y.; Wang, T.; Li, S.; He, X.; Li, M.; Wang, Z.; Zheng, N.; Shao, B.; Liu, T.Y. Enhancing geometric representations for molecules with equivariant vector-scalar interactive message passing. *Nat. Commun.* **2024**, *15*, 313. <https://doi.org/https://doi.org/10.1038/s41467-023-43720-2>.
20. Villar, S.; Hogg, D.W.; Storey-Fisher, K.; Yao, W.; Blum-Smith, B. Scalars are universal: Equivariant machine learning, structured like classical physics. In Proceedings of the 35th Conference on Neural Information Processing Systems (NeurIPS 2021), Virtual Event, 6–14 December 2021; pp. 28848–28863.
21. Aslan, B.; Platt, D.; Sheard, D. Group invariant machine learning by fundamental domain projections. In Proceedings of the NeurIPS Workshop on Symmetry and Geometry in Neural Representations, New Orleans, Louisiana, USA, 16 December 2023; pp. 181–218.
22. Blum-Smith, B.; Villar, S. Machine learning and invariant theory. *Notices Amer. Math. Soc.* **2023**, *70*. <https://doi.org/https://doi.org/10.1090/noti2760>.
23. Yarotsky, D. Universal approximations of invariant maps by neural networks. *Constr. Approx.* **2022**, *55*, 407–474. <https://doi.org/https://doi.org/10.1007/s00365-021-09546-1>.
24. Hilbert, D. Über die Theorie der algebraischen Formen. *Math. Ann.* **1890**, *36*, 473–534. <https://doi.org/https://doi.org/10.1007/BF01208503>.
25. Derksen, H.; Kemper, G. *Computational Invariant Theory*; Springer: Berlin, Heidelberg, Germany, 2015.
26. Kemper, G. Computing invariants of reductive groups in positive characteristic. *Transform. Groups* **2003**, *8*, 159–176. <https://doi.org/https://doi.org/10.1007/s00031-003-0305-1>.
27. Kemper, G. Separating invariants. *J. Symbolic Comput.* **2009**, *44*, 1212–1222. <https://doi.org/https://doi.org/10.1016/j.jsc.2008.02.012>.
28. Draisma, J.; Kemper, G.; Wehlau, D. Polarization of separating invariants. *Canadian J. Math.* **2008**, *60*, 556–571. <https://doi.org/https://doi.org/10.4153/CJM-2008-027-2>.
29. Domokos, M. Typical separating invariants. *Transform. Groups* **2007**, *12*, 49–63. <https://doi.org/https://doi.org/10.1007/s00031-005-1131-4>.
30. Cahill, J.; Contreras, A.; Contreras Hip, A. Complete set of translation invariant measurements with Lipschitz bounds. *Appl. Comput. Harmon. Anal.* **2020**, *49*, 521–539. <https://doi.org/https://doi.org/10.1016/j.acha.2020.05.007>.
31. Cahill, J.; Contreras, A.; Contreras Hip, A. Stable separation of orbits for finite Abelian group actions. *J. Fourier Anal. Appl.* **2024**, *30*, 12. <https://doi.org/https://doi.org/10.1007/s00041-024-10069-z>.
32. Cahill, J.; Iverson, J.W.; Mixon, D.G. Towards a bilipschitz invariant theory. *Appl. Comput. Harmon. Anal.* **2024**, *72*, 101669. <https://doi.org/https://doi.org/10.1016/j.acha.2024.101669>.
33. Dym, N.; Gortler, S.J. Low-dimensional invariant embeddings for universal geometric learning. *Found. Comput. Math.* **2025**, *25*, 375–415. <https://doi.org/https://doi.org/10.1007/s10208-024-09641-2>.
34. Cahill, J.; Iverson, J.W.; Mixon, D.G.; Packer, D. Group-invariant max filtering. *Found. Comput. Math.* **2025**, *25*, 1047–1084. <https://doi.org/https://doi.org/10.1007/s10208-024-09656-9>.
35. Mixon, D.G.; Qaddura, Y. Injectivity, stability, and positive definiteness of max filtering. *Constr. Approx.* **2025**, pp. 1–38. <https://doi.org/https://doi.org/10.1007/s00365-025-09707-6>.
36. Mixon, D.G.; Packer, D. Max filtering with reflection groups. *Adv. Comput. Math.* **2023**, *49*, 82. <https://doi.org/https://doi.org/10.1007/s10444-023-10084-6>.
37. Balan, R.; Tsoukanis, E.; Wellershoff, M. Stability of sorting based embeddings. *arXiv preprint arXiv:2410.05446* **2024**.
38. Mixon, D.G.; Qaddura, Y. Stable coorbit embeddings of orbifold quotients. *arXiv preprint arXiv:2403.14042* **2024**.

39. Balan, R.; Tsoukanis, E. G-invariant representations using coorbits: Injectivity properties. *arXiv preprint arXiv:2310.16365* **2023**.
40. Balan, R.; Tsoukanis, E. G-invariant representations using coorbits: Bi-Lipschitz properties. *arXiv preprint arXiv:2308.11784* **2023**.
41. Balan, R.; Haghani, N.; Singh, M. Permutation-invariant representations with applications to graph deep learning. *Appl. Comput. Harmon. Anal.* **2025**, *79*, 101798. <https://doi.org/https://doi.org/10.1016/j.acha.2025.101798>.
42. Bodmann, B.G.; Hammen, N. Stable phase retrieval with low-redundancy frames. *Adv. Comput. Math.* **2015**, *41*, 317–331. <https://doi.org/https://doi.org/10.1007/s10444-014-9359-y>.
43. Bandeira, A.S.; Cahill, J.; Mixon, D.G.; Nelson, A.A. Saving phase: Injectivity and stability for phase retrieval. *Appl. Comput. Harmon. Anal.* **2014**, *37*, 106–125. <https://doi.org/https://doi.org/10.1016/j.acha.2013.10.002>.
44. Balan, R.; Wang, Y. Invertibility and robustness of phaseless reconstruction. *Appl. Comput. Harmon. Anal.* **2015**, *38*, 469–488. <https://doi.org/https://doi.org/10.1016/j.acha.2014.07.003>.
45. Balan, R.; Zou, D. On Lipschitz analysis and Lipschitz synthesis for the phase retrieval problem. *Linear Algebra Appl.* **2016**, *496*, 152–181. <https://doi.org/https://doi.org/10.1016/j.laa.2015.12.029>.
46. Simon, B. *Representations of Finite and Compact Groups*; American Mathematical Society: Providence, RI, USA, 1996.
47. Evans, L.; Lai, C.K. Conjugate phase retrieval on $\mathbb{C}^M$ by real vectors. *Linear Algebra Appl.* **2020**, *587*, 45–69. <https://doi.org/https://doi.org/10.1016/j.laa.2019.11.001>.
48. Xu, Z. The minimal measurement number for low-rank matrix recovery. *Appl. Comput. Harmon. Anal.* **2018**, *44*, 497–508. <https://doi.org/https://doi.org/10.1016/j.acha.2017.01.005>.
49. Kech, M.; Wolf, M.M. Constrained quantum tomography of semi-algebraic sets with applications to low-rank matrix recovery. *Inf. Inference* **2017**, *6*, 171–195. <https://doi.org/https://doi.org/10.1093/imaiai/iaw019>.
50. Balan, R.; Dock, C.B. Lipschitz analysis of generalized phase retrievable matrix frames. *SIAM J. Matrix Anal. Appl.* **2022**, *43*, 1518–1571. <https://doi.org/https://doi.org/10.1137/21M1435446>.
51. Bochnak, J.; Coste, M.; Roy, M.F. *Real Algebraic Geometry*; Springer: Berlin, Heidelberg, Germany, 2013.
52. Dym, N.; Wellershoff, M.; Tsoukanis, E.; Levy, D.; Balan, R. Quantitative bounds for sorting-based permutation-invariant embeddings. *arXiv preprint arXiv:2510.22186* **2025**.
53. Wang, Y.; Xu, Z. Phase retrieval for sparse signals. *Appl. Comput. Harmon. Anal.* **2014**, *37*, 531–544. <https://doi.org/https://doi.org/10.1016/j.acha.2014.04.001>.
54. Basu, S.; Pollack, R.; Roy, M.F. *Algorithms in Real Algebraic Geometry*; Springer: Berlin, Heidelberg, Germany, 2006.
55. Warner, F.W. *Foundations of Differentiable Manifolds and Lie Groups*; Vol. 94, Springer: New York, NY, USA, 1983.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.