On Lipschitz Analysis and Lipschitz Synthesis for the Phase Retrieval Problem

Radu Balan[a], Dongmian Zou[b,*]

[a]*Department of Mathematics, Center for Scientific Computation and Mathematical Modeling, University of Maryland, College Park, MD 20742, USA*
[b]*Department of Mathematics, University of Maryland, College Park, MD 20742, USA*

Abstract

We prove two results with regard to reconstruction from magnitudes of frame coefficients (the so called "phase retrieval problem"). First we show that phase retrievable nonlinear maps are bi-Lipschitz with respect to appropriate metrics on the quotient space. Specifically, if nonlinear analysis maps $\alpha, \beta : \hat{H} \to \mathbb{R}^m$ are injective, with $\alpha(x) = (|\langle x, f_k\rangle|)_{k=1}^m$ and $\beta(x) = (|\langle x, f_k\rangle|^2)_{k=1}^m$, where $\{f_1, \ldots, f_m\}$ is a frame for a Hilbert space H and $\hat{H} = H/T^1$, then α is bi-Lipschitz with respect to the class of "natural metrics" $D_p(x, y) = \min_\varphi \|x - e^{i\varphi}y\|_p$, whereas β is bi-Lipschitz with respect to the class of matrix-norm induced metrics $d_p(x, y) = \|xx^* - yy^*\|_p$. Second we prove that reconstruction can be performed using Lipschitz continuous maps. That is, there exist left inverse maps (synthesis maps) $\omega, \psi : \mathbb{R}^m \to \hat{H}$ of α and β respectively, that are Lipschitz continuous with respect to appropriate metrics. Additionally, we obtain the Lipschitz constants of ω and ψ in terms of the lower Lipschitz constants of α and β, respectively. Surprisingly, the increase in both Lipschitz constants is a relatively small factor, independent of the space dimension or the frame redundancy.

Keywords: Frames, Lipschitz maps, Stability, Phase retrieval
2010 MSC: 15A29, 65H10, 90C26

1. Introduction

Let H be an n-dimensional real or complex Hilbert space. On H we consider the equivalence relation $\sim$ defined by

$$x \sim y \text{ iff there is a scalar } a \text{ of magnitude one, } |a| = 1, \text{ for which } y = ax \ .$$

Let $\hat{H} = H/\sim$ denote the collection of the equivalence classes. We use $\hat{x}$ to denote the equivalence class of x in $\hat{H}$. When there is no ambiguity, we also use x in place of $\hat{x}$ for simplicity.

*Corresponding author.
Email addresses: rvbalan@math.umd.edu (Radu Balan), zou@math.umd.edu (Dongmian Zou)

Assume that $\mathcal{F} = \{f_1, f_2, \dots, f_m\}$ is a frame (that is, a spanning set) for H. Let α and β denote the nonlinear maps

$$\alpha : \hat{H} \to \mathbb{R}^m \quad , \quad \alpha(x) = (|\langle x, f_k\rangle|)_{1\le k\le m} \, , \tag{1}$$

and

$$\beta : \hat{H} \to \mathbb{R}^m \quad , \quad \beta(x) = \left(|\langle x, f_k\rangle|^2\right)_{1\le k\le m} . \tag{2}$$

The *phase retrieval problem*, or the *phaseless reconstruction problem*, refers to analyzing when α (or equivalently, β) is an injective map, and in this case to finding "good" left inverses.

The frame $\mathcal{F}$ is said to be *phase retrievable* if the nonlinear map α (or β) is injective. In this paper we assume α and β are injective maps (hence $\mathcal{F}$ is phase retrievable). The problem is to analyze the stability properties of phaseless reconstruction. We explore this problem by studying Lipschitz properties of these nonlinear maps. A continuous map $f : (X, d_X) \to (Y, d_Y)$, defined between metric spaces X and Y with distances d_X and d_Y respectively, is Lipschitz continuous with Lipschitz constant $\mathrm{Lip}(f)$ if

$$\mathrm{Lip}(f) := \sup_{x_1,x_2\in X} \frac{d_Y(f(x_1), f(x_2))}{d_X(x_1, x_2)} < \infty \, .$$

Further, the map f is called *bi-Lipschitz* with lower Lipschitz constant a and upper Lipschitz constant b if for every $x_1, x_2 \in X$,

$$a\, d_X(x_1, x_2) \le d_Y(f(x_1), f(x_2)) \le b\, d_X(x_1, x_2) \, .$$

Obviously the smallest upper Lipschitz constant is $b = \mathrm{Lip}(f)$. If f is bi-Lipschitz then f is injective.

The space $\hat{H}$ admits two classes of inequivalent distances. We introduce and study them in detail in section 2. In particular, consider the following two distances:

$$D_2(x, y) = \min_{\varphi} \left\|x - e^{i\varphi} y\right\|_2 = \sqrt{\|x\|^2 + \|y\|^2 - 2|\langle x, y\rangle|} \, ,$$

and

$$d_1(x, y) = \|xx^* - yy^*\|_1 = \sqrt{(\|x\|^2 + \|y\|^2)^2 - 4|\langle x, y\rangle|^2} \, .$$

When the frame is phase retrievable the nonlinear maps $\alpha : (\hat{H}, D_2) \to (\mathbb{R}^m, \|\cdot\|_2)$ and $\beta : (\hat{H}, d_1) \to (\mathbb{R}^m, \|\cdot\|_2)$ are shown to be bi-Lipschitz. This statement was previously known for the map β in the real and complex case (see [2, 3, 6]), and for the map α in the real case only (see [13, 6, 8]). In this paper we prove this statement for α in the complex case.

In general, noisy measurements are not in the image of the analysis map $\alpha(\hat{H})$ or $\beta(\hat{H})$. In this paper we prove that the unique left inverses of α and β can be extended from $\alpha(\hat{H})$ and $\beta(\hat{H})$, respectively, to the entire space $\mathbb{R}^m$ while the extended maps remain to be Lipschitz continuous. Specifically, there exist two Lipschitz continuous maps $\omega : (\mathbb{R}^m, \|\cdot\|_2) \to (\hat{H}, D_2)$ and $\psi : (\mathbb{R}^m, \|\cdot\|_2) \to (\hat{H}, d_1)$ so that $\omega(\alpha(x)) = x$ and $\psi(\beta(x)) = x$ for every $x \in H$.

Consider one of the maps α and β, say α (a similar discussion works for β). Assume an additive noise model $y = \alpha(x) + \nu$, where $\nu \in \mathbb{R}^m$ is the noise. For a signal $x_0 \in \hat{H}$, and noise $\nu_1 \in \mathbb{R}^m$, let $y_1 = \alpha(x_0) + \nu_1 \in \mathbb{R}^m$ be the measurement vector, and let $x_1 = \omega(y_1)$ be the reconstructed signal. We have

$$d_1(x_0, x_1) = d_1\left(\omega(\alpha(x_0)), \omega(y_1)\right) \leq \mathrm{Lip}(\omega) \cdot \|\alpha(x_0) - y_1\| = \mathrm{Lip}(\omega) \cdot \|\nu_1\|.$$

Figure 1 is an illustration of this model. In fact, we have stability in a stronger sense. If we have two noisy measurements $y_1 = \alpha(x_0) + \nu_1$ and $y_2 = \alpha(x_0) + \nu_2$ of the signal x_0, then

$$d_1(x_1, x_2) = d_1\left(\omega(y_1), \omega(y_2)\right) \leq \mathrm{Lip}(\omega) \cdot \|y_1 - y_2\| = \mathrm{Lip}(\omega) \cdot \|\nu_1 - \nu_2\|.$$

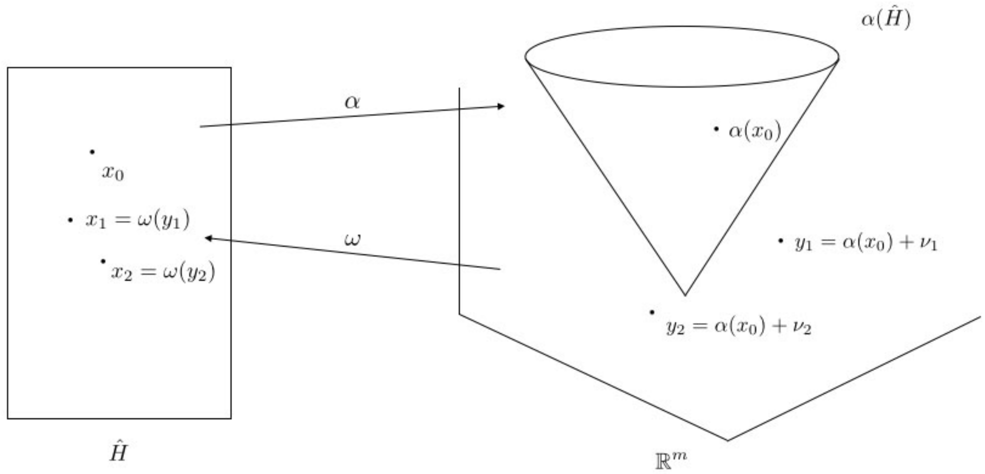

Figure 1: Illustration of the noisy measurement model

Denote by a_α and a_β the lower Lipschitz constants of α and β respectively. In this paper we prove also that the upper Lipschitz constants of these maps obey $\mathrm{Lip}(\omega) \leq \frac{8.25}{a_\alpha}$ and $\mathrm{Lip}(\psi) \leq \frac{8.25}{a_\beta}$. Surprisingly, this shows the Lipschitz constant of these left inverses are just a small factor larger than the minimal Lipschitz constants. Furthermore this factor is independent of dimension n or number of frame vectors m.

The organization of this paper is as follows. Section 2 introduces notations and presents the results for bi-Lipschitz properties. Section 3 presents the results for the extension of the left inverse. Section 4 contains the proof of these results.

2. Bi-Lipschitz Properties for the Analysis Map

2.1. Notations

To study the bi-Lipschitz properties, we need to choose an appropriate distance on $\hat{H}$. We consider two classes of metrics (distances), respectively:

1. the class of *natural metrics*. For every $1 \leq p \leq \infty$ and $x, y \in H$, we define

$$D_p(\hat{x}, \hat{y}) = \min_{|a|=1} \|x - ay\|_p \ .$$

When no subscript is used, $\|\cdot\|$ denotes the Euclidean norm, $\|\cdot\| = \|\cdot\|_2$.
2. the class of *matrix norm induced metrics*. For every $1 \leq p \leq \infty$ and $x, y \in H$, we define

$$d_p(\hat{x}, \hat{y}) = \|xx^* - yy^*\|_p = \begin{cases} (\sum_{k=1}^n (\sigma_k)^p)^{1/p} & for \quad 1 \leq p < \infty \\ \max_{1\leq k\leq n} \sigma_k & for \quad p = \infty \end{cases} , \tag{3}$$

where $(\sigma_k)_{1\leq k\leq n}$ are the singular values of the operator $xx^* - yy^*$, which is of rank at most 2. Here x^* denotes the adjoint of x (see [3] for a detailed discussion), which is the transpose conjugate of x if $H = \mathbb{R}^n$ or $\mathbb{C}^n$.

Our choice in (3) corresponds to the class of Schatten norms. In particular, d_∞ corresponds to the operator norm $\|\cdot\|_{op}$ in $\mathrm{Sym}(H) = \{T : H \to H \ , \ T = T^*\}$; d_2 corresponds to the Frobenius norm $\|\cdot\|_{Fr}$ in $\mathrm{Sym}(H)$; d_1 corresponds to the nuclear norm $\|\cdot\|_*$ in $\mathrm{Sym}(H)$. Specifically, we have

$$d_\infty(x, y) = \|xx^* - yy^*\|_{op} \ , \ d_2(x, y) = \|xx^* - yy^*\|_{Fr} \ ,$$

$$d_1(x, y) = \|xx^* - yy^*\|_* \ .$$

Note that the Frobenius norm $\|T\|_{Fr} = \sqrt{\mathrm{trace}(TT^*)}$ induces the Euclidean distance on $\mathrm{Sym}(H)$. As a consequence of Lemma 3.8 in [3], we have:

$$d_\infty(x, y) = \frac{1}{2}|\ \|x\|^2 - \|y\|^2\ | + \frac{1}{2}\sqrt{(\|x\|^2 + \|y\|^2)^2 - 4|\ \langle x, y\rangle\ |^2} \ ,$$

$$d_2(x, y) = \sqrt{\|x\|^4 + \|y\|^4 - 2|\ \langle x, y\rangle\ |^2} \ ,$$

$$d_1(x, y) = \sqrt{(\|x\|^2 + \|y\|^2)^2 - 4|\ \langle x, y\rangle\ |^2} \ .$$

To study the above distances it is important to study eigenvalues of symmetric matrices. Let $S^{p,q}(H)$ denote the set of symmetric operators that have at most p strictly positive eigenvalues and q strictly negative eigenvalues. In particular, $S^{1,0}(H)$ is the set of non-negative symmetric operators of rank at most one:

$$S^{1,0}(H) = \{\, xx^*, \quad x \in H \,\} \ . \tag{4}$$

If $H = \mathbb{R}^n$ or $\mathbb{C}^n$, then $\mathrm{Sym}(H)$ is the set of n-dimensional Hermitian matrices. For a matrix $X \in \mathrm{Sym}(\mathbb{R}^n)$ or $\mathrm{Sym}(\mathbb{C}^n)$, we use $\lambda_1(X)$, $\lambda_2(X)$, $\cdots$, $\lambda_n(X)$ to denote its eigenvalues. These eigenvalues are real numbers and we arrange them to satisfy $\lambda_1(X) \geq \lambda_2(X) \geq \cdots \geq \lambda_n(X)$.

To analyze the bi-Lipschitz properties, we define the following three types of Lipschitz bounds for α. Note that the Lipschitz constants are square-roots of those bounds.

(i) The *global lower* and *upper Lipschitz bounds*, respectively:

$$A_0 = \inf_{x,y \in \hat{H}} \frac{\|\alpha(x) - \alpha(y)\|_2^2}{D_2(x,y)^2} \ ,$$

$$B_0 = \sup_{x,y \in \hat{H}} \frac{\|\alpha(x) - \alpha(y)\|_2^2}{D_2(x,y)^2} \ ;$$

(ii) The *type I local lower* and *upper Lipschitz bounds* at $z \in \hat{H}$, respectively:

$$A(z) = \lim_{r \to 0} \inf_{\substack{x,y \in \hat{H} \\ D_2(x,z) < r \\ D_2(y,z) < r}} \frac{\|\alpha(x) - \alpha(y)\|_2^2}{D_2(x,y)^2} \ ,$$

$$B(z) = \sup_{r \to 0} \inf_{\substack{x,y \in \hat{H} \\ D_2(x,z) < r \\ D_2(y,z) < r}} \frac{\|\alpha(x) - \alpha(y)\|_2^2}{D_2(x,y)^2} \ ;$$

(iii) The *type II local lower* and *upper Lipschitz bounds* at $z \in \hat{H}$, respectively:

$$\tilde{A}(z) = \lim_{r \to 0} \inf_{\substack{x \in \hat{H} \\ D_2(x,z) < r}} \frac{\|\alpha(x) - \alpha(z)\|_2^2}{D_2(x,z)^2} \ ,$$

$$\tilde{B}(z) = \sup_{r \to 0} \inf_{\substack{x \in \hat{H} \\ D_2(x,z) < r}} \frac{\|\alpha(x) - \alpha(z)\|_2^2}{D_2(x,y)^2} \ .$$

Similarly, we define the three types of Lipschitz constants for β.

(i) The *global lower* and *upper Lipschitz bounds*, respectively:

$$a_0 = \inf_{x,y \in \hat{H}} \frac{\|\beta(x) - \beta(y)\|_2^2}{d_1(x,y)^2} \ ,$$

$$b_0 = \sup_{x,y \in \hat{H}} \frac{\|\beta(x) - \beta(y)\|_2^2}{d_1(x,y)^2} \ ;$$

(ii) The *type I local lower* and *upper Lipschitz bounds* at $z \in \hat{H}$, respectively:

$$a(z) = \lim_{r \to 0} \inf_{\substack{x,y \in \hat{H} \\ d_1(x,z) < r \\ d_1(y,z) < r}} \frac{\|\beta(x) - \beta(y)\|_2^2}{d_1(x,y)^2} \ ,$$

$$b(z) = \lim_{r \to 0} \sup_{\substack{x,y \in \hat{H} \\ d_1(x,z) < r \\ d_1(y,z) < r}} \frac{\|\beta(x) - \beta(y)\|_2^2}{d_1(x,y)^2} \ ;$$

(iii) The *type II local lower* and *upper Lipschitz bounds* at $z \in \hat{H}$, respectively:

$$\tilde{a}(z) = \lim_{r\to 0} \inf_{\substack{x\in\hat{H} \\ d_1(x,z)<r}} \frac{\|\beta(x) - \beta(z)\|_2^2}{d_1(x,z)^2} ,$$

$$\tilde{b}(z) = \lim_{r\to 0} \sup_{\substack{x\in\hat{H} \\ d_1(x,z)<r}} \frac{\|\beta(x) - \beta(z)\|_2^2}{d_1(x,z)^2} .$$

Due to homogeneity we have $A_0 = A(0)$, $B_0 = B(0)$, $a_0 = a(0)$, $b_0 = b(0)$. Also, for $z \neq 0$, we have $A(z) = A(z/\|z\|)$, $B(z) = B(z/\|z\|)$, $a(z) = a(z/\|z\|)$, $b(z) = b(z/\|z\|)$.

We analyze the bi-Lipschitz properties of α and β by studying these constants.

2.2. Bi-Lipschitz Properties for α

The real case $H = \mathbb{R}^n$ is studied in [6]. We summarize the results as a theorem.

Recall that $\mathcal{F} = \{f_1, \cdots, f_m\}$ is a frame in H if there exist positive constants A and B for which

$$A\|x\|^2 \leq \sum_{k=1}^{m} |\langle x, f_k\rangle|^2 \leq B\|x\|^2 . \tag{5}$$

We say A (resp., B) is the optimal lower (resp., upper) frame bound if A (resp., B) is the largest (resp., smallest) positive number for which the inequality (5) is satisfied.

For any index set $I \subset \{1, 2, \cdots, m\}$, let $\mathcal{F}[I] = \{f_k, k \in I\}$ denote the frame subset indexed by I. Also, let $\sigma_1^2[I]$ and $\sigma_n^2[I]$ denote the upper and lower frame bound of set $\mathcal{F}[I]$, respectively. It is straightforward to see that they respectively correspond to the largest and smallest eigenvalues of $\sum_{k\in I} f_k f_k^*$, that is,

$$\sigma_1^2[I] = \lambda_{\max}\left(\sum_{k\in I} f_k f_k^*\right) \qquad \text{and} \qquad \sigma_n^2[I] = \lambda_{\min}\left(\sum_{k\in I} f_k f_k^*\right) .$$

Theorem 2.1 ([6])**.** *Let $\mathcal{F} \subset \mathbb{R}^n$ be a phase retrievable frame for $\mathbb{R}^n$. Let A and B denote its optimal lower and upper frame bound, respectively. Then*

(i) *For every $0 \neq x \in \mathbb{R}^n$, $A(x) = \sigma_n^2(supp(\alpha(x))$ where $supp(\alpha(x)) = \{k, \langle x, f_k\rangle \neq 0\}$;*
(ii) *For every $x \in \mathbb{R}^n$, $\tilde{A} = A$;*
(iii) *$A_0 = A(0) = \min_{I\subset\{1,2,\cdots,m\}}(\sigma_n^2[I] + \sigma_n^2[I^c])$;*
(iv) *For every $x \in \mathbb{R}^n$, $B(x) = \tilde{B}(x) = B$;*
(v) *$B_0 = B(0) = \tilde{B}(0) = B$.*

Now we consider the complex case $H = \mathbb{C}^n$. We analyze the complex case by doing a realification first. Consider the $\mathbb{R}$-linear map $\mathbf{j} : \mathbb{C}^n \to \mathbb{R}^{2n}$ defined by

$$\mathbf{j}(z) = \begin{bmatrix} \text{real}(z) \\ \text{imag}(z) \end{bmatrix} .$$

This realification is studied in detail in [3]. We call $\mathbf{j}(z)$ the realification of z. For simplicity, in this paper we will denote $\xi = \mathbf{j}(x)$, $\eta = \mathbf{j}(y)$, $\zeta = \mathbf{j}(z)$, $\varphi = \mathbf{j}(f)$, $\delta = \mathbf{j}(d)$, respectively.

For a frame set $\mathcal{F} = \{f_1, f_2, \cdots, f_m\}$, define the symmetric operator

$$\Phi_k = \varphi_k \varphi_k^T + J \varphi_k \varphi_k^T J^T, \quad k = 1, 2, \cdots, m.$$

where

$$J = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix} \tag{6}$$

is a matrix in $\mathbb{R}^{2n \times 2n}$.

Also, define $\mathcal{S} : \mathbb{R}^{2n} \to \mathrm{Sym}(\mathbb{R}^{2n})$ by

$$\mathcal{S}(\xi) = \begin{cases} 0 & , \text{ if } \xi = 0 \\ \sum_{k:\Phi_k\xi \neq 0} \frac{1}{\langle \Phi_k \xi, \xi \rangle} \Phi_k \xi \xi^T \Phi_k & , \text{ if } \xi \neq 0 \end{cases} .$$

We have the following result (proved in Section 4):

Theorem 2.2. *Let $\mathcal{F} \subset \mathbb{C}^n$ be a phase retrievable frame for $\mathbb{C}^n$. Let A and B denote its optimal lower and upper frame bound, respectively. For any $z \in \mathbb{C}^n$, let $\zeta = \mathbf{j}(z)$ be its realification. Then*

(i) *For every $0 \neq z \in \mathbb{C}^n$, $A(z) = \lambda_{2n-1}(\mathcal{S}(\zeta))$;*
(ii) *$A_0 = A(0) > 0$;*
(iii) *For every $z \in \mathbb{C}^n$, $\tilde{A}(z) = \lambda_{2n-1}\left(\mathcal{S}(\zeta) + \sum_{k:\langle z, f_k \rangle = 0} \Phi_k\right)$;*
(iv) *$\tilde{A}(0) = A$;*
(v) *For every $z \in \mathbb{C}^n$, $B(z) = \tilde{B}(z) = \lambda_1\left(\mathcal{S}(\zeta) + \sum_{k:\langle z, f_k \rangle = 0} \Phi_k\right)$;*
(vi) *$B_0 = B(0) = \tilde{B}(0) = B$.*

2.3. Bi-Lipschitz Properties for β

The nonlinear map β naturally induces a linear map between the space $\mathrm{Sym}(H)$ of symmetric operators on H and $\mathbb{R}^m$:

$$\mathcal{A} : \mathrm{Sym}(H) \to \mathbb{R}^m \quad , \quad \mathcal{A}(T) = (\langle T f_k, f_k \rangle)_{1 \leq k \leq m} .$$

This linear map has first been observed in [5] and it has been exploited successfully in various papers e.g. [1, 11, 2]. Note that the map β is injective if and only if $\mathcal{A}$ restricted to $S^{1,0}(H)$ is injective.

In previous papers [3, 6], the authors establish global bi-Lipschitz results for phase-retrievable frames. We summarize them as follows:

Theorem 2.3 ([3], [6])**.** *Let $\mathcal{F}$ be a phase retrievable frame for $H = \mathbb{C}^n$. Then*

(i) *the global lower Lipschitz bound $a_0 > 0$;*

(ii) *the global upper Lipschitz bound* $b_0 < \infty$, *and*

$$\begin{aligned} b_0 &= \max_{\|x\|=\|y\|=1} \sum_{k=1}^{m} \left(\operatorname{real}\left(\langle x, f_k\rangle \langle f_k, y\rangle\right)\right)^2 \\ &= \max_{\|x\|=1} \sum_{k=1}^{m} |\langle x, f_k\rangle|^4 \\ &= \|T\|^4_{B(H,l^4_m)} \ , \end{aligned}$$

where $T : H \to \mathbb{C}^m$ *is the analysis operator defined by* $x \mapsto (\langle x, f_k\rangle)_{k=1}^m$, *and* $l^4_m := (\mathbb{C}^m, \|\cdot\|_4)$.

Remark 2.4. An upper bound of b_0 is given by

$$b_0 \leq B \left(\max_{1\leq k\leq m} \|f_k\|\right)^2 \leq B^2 \ ,$$

where B is the upper frame bound of $\mathcal{F}$.

We give an expression of the local Lipschitz bounds as well. Define $\mathcal{R} : \mathbb{R}^{2n} \to \mathrm{Sym}(\mathbb{R}^{2n})$ by

$$\mathcal{R}(\xi) = \sum_{k=1}^{m} \Phi_k \xi \xi^T \Phi_k \ .$$

Theorem 2.5. *Let* $\mathcal{F}$ *be a phase retrievable frame for* $H = \mathbb{C}^n$. *For every* $0 \neq z \in H$, *let* $\zeta = \mathbf{j}(z)$ *denote the realification of* z. *Then*

(i) $a(z) = \tilde{a}(z) = \lambda_{2n-1}(\mathcal{R}(\zeta))/\|\zeta\|^2$;
(ii) $b(z) = \tilde{b}(z) = \lambda_1(\mathcal{R}(\zeta))/\|\zeta\|^2$;
(iii) (see [3]) $a(0) = a_0 = \min_{\|\zeta\|=1} \lambda_{2n-1}\left(\mathcal{R}(\zeta)\right)$;
(iv) $\tilde{a}(0) = \min_{\|x\|=1} \sum_{k=1}^m |\langle x, f_k\rangle|^4$;
(v) $b(0) = \tilde{b}(0) = b_0$.

3. Extension of the Inverse Map

The results in this section work for both $H = \mathbb{R}^n$ and $\mathbb{C}^n$. First we show that all metrics D_p and d_p defined in Section 2 induce the same topology in the following result.

Proposition 3.1. *We have the following statements regarding* D_p *and* d_p:

(i) *For each* $1 \leq p \leq \infty$, D_p *and* d_p *are metrics (distances) on* $\hat{H}$.
(ii) $(D_p)_{1\leq p\leq\infty}$ *are equivalent metrics, that is each* D_p *induces the same topology on* $\hat{H}$ *as* D_1. *Additionally, for every* $1 \leq p, q \leq \infty$ *the embedding* $i : (\hat{H}, D_p) \to (\hat{H}, D_q)$, $i(x) = x$, *is Lipschitz with Lipschitz constant*

$$L^D_{p,q,n} = \max(1, n^{\frac{1}{q}-\frac{1}{p}}). \tag{7}$$

(iii) *For* $1 \le p,q \le \infty$, $(d_p)_{1\le p\le\infty}$ *are equivalent metrics, that is each* d_p *induces the same topology on* $\hat{H}$ *as* d_1. *Additionally, for every* $1 \le p,q \le \infty$ *the embedding* $i : (\hat{H}, d_p) \to (\hat{H}, d_q)$, $i(x) = x$, *is Lipschitz with Lipschitz constant*

$$L^d_{p,q,n} = \max(1, 2^{\frac{1}{q}-\frac{1}{p}}). \tag{8}$$

(iv) *The identity map* $i : (\hat{H}, D_p) \to (\hat{H}, d_p)$, $i(x) = x$, *is continuous with continuous inverse. However it is not Lipschitz, nor is its inverse.*

(v) *The metric space* $(\hat{H}, D_p)$ *is Lipschitz isomorphic to* $S^{1,0}(H)$ *endowed with Schatten norm* $\|\cdot\|_p$. *The isomorphism is given by the map*

$$\kappa_\alpha : \hat{H} \to S^{1,0}(H) \quad , \quad \kappa_\alpha(x) = \begin{cases} \frac{1}{\|x\|} xx^* & if \quad x \neq 0 \\ 0 & if \quad x = 0 \end{cases} . \tag{9}$$

The embedding κ_α *is bi-Lipschitz with the lower Lipschitz constant*

$$\min(2^{\frac{1}{2}-\frac{1}{p}}, n^{\frac{1}{p}-\frac{1}{2}})$$

and the upper Lipschitz constant

$$\sqrt{2}\max(n^{\frac{1}{2}-\frac{1}{p}}, 2^{\frac{1}{p}-\frac{1}{2}}) \ .$$

In particular, for $p = 2$, *the lower Lipschitz constant is 1 and the upper Lipschitz constant is* $\sqrt{2}$.

(vi) *The metric space* $(\hat{H}, d_p)$ *is isometrically isomorphic to* $S^{1,0}(H)$ *endowed with Schatten norm* $\|\cdot\|_p$. *The isomorphism is given by the map*

$$\kappa_\beta : \hat{H} \to S^{1,0}(H) \quad , \quad \kappa_\beta(x) = xx^*. \tag{10}$$

In particular the metric space $(\hat{H}, d_1)$ *is isometrically isomorphic to* $S^{1,0}(H)$ *endowed with the nuclear norm* $\|\cdot\|_1$.

(vii) *The nonlinear map* $\iota : (\hat{H}, D_p) \to (\hat{H}, d_p)$ *defined by*

$$\iota(x) = \begin{cases} \frac{x}{\sqrt{\|x\|}} & if \quad x \neq 0 \\ 0 & if \quad x = 0 \end{cases}$$

is bi-Lipschitz with the lower Lipschitz constant $\min(2^{\frac{1}{2}-\frac{1}{p}}, n^{\frac{1}{p}-\frac{1}{2}})$ *and the upper Lipschitz constant* $\sqrt{2}\max(n^{\frac{1}{2}-\frac{1}{p}}, 2^{\frac{1}{p}-\frac{1}{2}})$.

Remark 3.2. (i) Note that the Lipschitz bound $L^D_{p,q,n}$ is equal to the operator norm of the identity between $(\mathbb{C}^n, \|\cdot\|_p)$ and $(\mathbb{C}^n, \|\cdot\|_q)$: $L^D_{p,q,n} = \|I\|_{l^p_n \to l^q_n}$.

(ii) Note the equality $L^d_{p,q,n} = L^D_{p,q,2}$.

The results in Section 2, together with the previous proposition, show that if the frame $\mathcal{F}$ is phase retrievable, then the nonlinear map α (resp., β) is bi-Lipschitz between the metric spaces $(\hat{H}, D_p)$ (resp., $(\hat{H}, d_p)$) and $(\mathbb{R}^m, \|\cdot\|_q)$. Recall that the Lipschitz constants between $(\hat{H}, D_2)$ (resp., $(\hat{H}, d_1)$) and $(\mathbb{R}^m, \|\cdot\| = \|\cdot\|_2)$ are given by $\sqrt{A_0}$ (resp., $\sqrt{a_0}$) and $\sqrt{B_0}$ (resp., $\sqrt{b_0}$):

$$\sqrt{A_0} D_2(x,y) \leq \|\alpha(x) - \alpha(y)\| \leq \sqrt{B_0} D_2(x,y) \ , \tag{11}$$

$$\sqrt{a_0} d_1(x,y) \leq \|\beta(x) - \beta(y)\| \leq \sqrt{b_0} d_1(x,y) \ . \tag{12}$$

Clearly the inverse map defined on the range of α (resp., β) from metric space $(\alpha(\hat{H}), \|\cdot\|)$ (resp., $(\beta(\hat{H}), \|\cdot\|)$) to $(\hat{H}, D_2)$ (resp., $(\hat{H}, d_1)$):

$$\tilde{\omega} : \alpha(\hat{H}) \subset \mathbb{R}^m \to \hat{H} \quad , \quad \tilde{\omega}(c) = x \ \text{ if } \alpha(x) = c \ ; \tag{13}$$

$$\tilde{\psi} : \beta(\hat{H}) \subset \mathbb{R}^m \to \hat{H} \quad , \quad \tilde{\psi}(c) = x \ \text{ if } \beta(x) = c \ . \tag{14}$$

is Lipschitz with Lipschitz constant $1/\sqrt{A_0}$ (resp., $1/\sqrt{a_0}$). We prove that both $\tilde{\omega}$ and $\tilde{\psi}$ can be extended to the entire $\mathbb{R}^m$ as a Lipschitz map, and its Lipschitz constant is increased by a small factor.

The precise statement is given in the following Theorem, which is the main result of this paper.

Theorem 3.3. *Let $\mathcal{F} = \{f_1, \ldots, f_m\}$ be a phase retrievable frame for the n dimensional Hilbert space H, and let $\alpha, \beta : \hat{H} \to \mathbb{R}^m$ denote the injective nonlinear analysis maps as defined in (1) and (2). Let A_0 and a_0 denote the positive constants as in (11) and (12). Then*

(i) *there exists a Lipschitz continuous function $\omega : \mathbb{R}^m \to \hat{H}$ so that $\omega(\alpha(x)) = x$ for all $x \in \hat{H}$. For any $1 \leq p, q \leq \infty$, ω has an upper Lipschitz constant $\mathrm{Lip}(\omega)_{p,q}$ between $(\mathbb{R}^m, \|\cdot\|_p)$ and $(\hat{H}, D_q)$ bounded by:*

$$\mathrm{Lip}(\omega)_{p,q} \leq \begin{cases} \frac{3\sqrt{2}+4}{\sqrt{A_0}} \cdot 2^{\frac{1}{q}-\frac{1}{2}} \cdot \max(1, m^{\frac{1}{2}-\frac{1}{p}}) & \text{for } q \leq 2; \\ \frac{3\sqrt{2}+2^{\frac{3}{2}+\frac{1}{q}}}{\sqrt{A_0}} \cdot n^{\frac{1}{2}-\frac{1}{q}} \cdot \max(1, m^{\frac{1}{2}-\frac{1}{p}}) & \text{for } q > 2. \end{cases} \tag{15}$$

Explicitly this means: for $q \leq 2$ and for all $c, d \in \mathbb{R}^m$:

$$D_q(\omega(c), \omega(d)) \leq \frac{3\sqrt{2}+4}{\sqrt{A_0}} \cdot 2^{\frac{1}{q}-\frac{1}{2}} \cdot \max(1, m^{\frac{1}{2}-\frac{1}{p}}) \|c-d\|_p \ , \tag{16}$$

whereas for $q > 2$ and for all $c, d \in \mathbb{R}^m$:

$$D_q(\omega(c), \omega(d)) \leq \frac{3\sqrt{2}+2^{\frac{3}{2}+\frac{1}{q}}}{\sqrt{A_0}} \cdot n^{\frac{1}{2}-\frac{1}{q}} \cdot \max(1, m^{\frac{1}{2}-\frac{1}{p}}) \|c-d\|_p \ . \tag{17}$$

In particular, for $p = 2$ and $q = 2$ its Lipschitz constant $\mathrm{Lip}(\omega)_{2,2}$ is bounded by $\frac{4+3\sqrt{2}}{\sqrt{a_0}}$:

$$D_2(\omega(c), \omega(d)) \leq \frac{4+3\sqrt{2}}{\sqrt{a_0}} \|c-d\| \ . \tag{18}$$

(ii) *there exists a Lipschitz continuous function* $\psi : \mathbb{R}^m \to \hat{H}$ *so that* $\psi(\beta(x)) = x$ *for all* $x \in \hat{H}$. *For any* $1 \le p, q \le \infty$, ψ *has an upper Lipschitz constant* $\mathrm{Lip}(\psi)_{p,q}$ *between* $(\mathbb{R}^m, \|\cdot\|_p)$ *and* $(\hat{H}, d_q)$ *bounded by:*

$$\mathrm{Lip}(\psi)_{p,q} \le \begin{cases} \frac{3+2\sqrt{2}}{\sqrt{a_0}} \cdot 2^{\frac{1}{q}-\frac{1}{2}} \cdot \max(1, m^{\frac{1}{2}-\frac{1}{p}}) & \text{for } q \le 2; \\ \frac{3+2^{1+\frac{1}{q}}}{\sqrt{a_0}} \max(1, m^{\frac{1}{2}-\frac{1}{p}}) & \text{for } q > 2. \end{cases} \tag{19}$$

Explicitly this means: for $q \le 2$ *and for all* $c, d \in \mathbb{R}^m$*:*

$$d_q(\psi(c), \psi(d)) \le \frac{3+2\sqrt{2}}{\sqrt{a_0}} \cdot 2^{\frac{1}{q}-\frac{1}{2}} \cdot \max(1, m^{\frac{1}{2}-\frac{1}{p}}) \|c-d\|_p \ , \tag{20}$$

whereas for $q > 2$ *and for all* $c, d \in \mathbb{R}^m$*:*

$$d_q(\psi(c), \psi(d)) \le \frac{3+2^{1+\frac{1}{q}}}{\sqrt{a_0}} \max(1, m^{\frac{1}{2}-\frac{1}{p}}) \|c-d\|_p \ . \tag{21}$$

In particular, for $p = 2$ *and* $q = 1$ *its Lipschitz constant* $\mathrm{Lip}(\psi)_{2,1}$ *bounded by* $\frac{4+3\sqrt{2}}{\sqrt{a_0}}$*:*

$$d_1(\psi(c), \psi(d)) \le \frac{4+3\sqrt{2}}{\sqrt{a_0}} \|c-d\| \ . \tag{22}$$

The proof of Theorem 3.3, presented in Section 4, requires the construction of a special Lipschitz map. We believe this particular result is interesting in itself and may be used in other constructions. This construction is given in [7] for the case $p = 2$. Here we consider a general p and give a better bound for the Lipschitz constant. We state it as a lemma.

Lemma 3.4. *Consider the spectral decomposition of any self-adjoint operator* A *in* $\mathrm{Sym}(H)$*, say* $A = \sum_{k=1}^{d} \lambda_{m(k)} P_k$*, where* $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$ *are the* n *eigenvalues including multiplicities, and* $P_1,\ldots,P_d$ *are the orthogonal projections associated to the* d *distinct eigenvalues. Additionally,* $m(1) = 1$ *and* $m(k+1) = m(k) + r(k)$*, where* $r(k) = rank(P_k)$ *is the multiplicity of eigenvalue* $\lambda_{m(k)}$*. Then the map*

$$\pi : \mathrm{Sym}(H) \to S^{1,0}(H) \quad , \quad \pi(A) = (\lambda_1 - \lambda_2) P_1 \tag{23}$$

satisfies the following two properties:

(i) *for* $1 \le p \le \infty$, π *is Lipschitz continuous from* $(\mathrm{Sym}(H), \|\cdot\|_p)$ *to* $(S^{1,0}(H), \|\cdot\|_p)$ *with Lipschitz constant* $\mathrm{Lip}(\pi) \le 3 + 2^{1+\frac{1}{p}}$*;*
(ii) $\pi(A) = A$ *for all* $A \in S^{1,0}(H)$*.*

Remark 3.5. Numerical experiments suggest that the Lipschitz constant of π is smaller than 5 for $p = \infty$. On the other hand it cannot be smaller than 2 as the following example shows.

Example 3.6. If $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$, then $\pi(A) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and $\pi(B) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$. Here we have $\|\pi(A) - \pi(B)\|_\infty = 2$ and $\|A - B\|_\infty = 1$. Thus for this example we have

$$\|\pi(A) - \pi(B)\|_\infty = 2 \|A - B\|_\infty \ .$$

It is unlikely to obtain an isometric extension in Theorem 3.3. Kirszbraun theorem [14] gives a sufficient condition for isometric extensions of Lipschitz maps. The theorem states that isometric extensions are possible when the pair of metric spaces satisfy the Kirszbraun property, or the K property:

Definition 3.7 (The Kirszbraun Property (K)). Let X and Y be two metric spaces with metric d_x and d_y respectively. (X, Y) is said to have *Property (K)* if for any pair of families of closed balls $\{B(x_i, r_i) : i \in I\}$, $\{B(y_i, r_i) : i \in I\}$, such that $d_y(y_i, y_j) \le d_x(x_i, x_j)$ for each $i, j \in I$, it holds that $\bigcap B(x_i, r_i) \neq \emptyset \Rightarrow \bigcap B(y_i, r_i) \neq \emptyset$.

If (X, Y) has Property (K), then by Kirszbraun's Theorem we can extend a Lipschitz mapping defined on a subspace of X to a Lipschitz mapping defined on X while maintaining the Lipschitz constant. Unfortunately, if we consider $(X, d_X) = (\mathbb{R}^m, \|\cdot\|)$ and $Y = \hat{H}$, Property (K) does not hold for either D_p or d_p.

Example 3.8. *Property (K) does not hold for $\hat{H}$ with norm D_p.* Specifically, $(\mathbb{R}^m, \mathbb{R}^n/\sim)$ does not have Property K. We give a counterexample for $m = n = 2, p = 2$: Let $\tilde{y}_1 = (3, 1)$, $\tilde{y}_2 = (-1, 1)$, $\tilde{y}_3 = (0, 1)$ be the representatives of three points y_1, y_2, y_3 in $\mathbb{R}^2/\sim$. Then $D_2(y_1, y_2) = 2\sqrt{2}$, $D_2(y_2, y_3) = 1$ and $D_2(y_1, y_3) = 3$. Consider $x_1 = (0, 0)$, $x_2 = (0, -2\sqrt{2})$, $x_3 = (-1, -2\sqrt{2})$ in $\mathbb{R}^2$ with the Euclidean distance, then we have $\|x_1 - x_2\| = 2\sqrt{2}$, $\|x_2 - x_3\| = 1$ and $\|x_1 - x_3\| = 3$. For $r_1 = \sqrt{6}$, $r_2 = 2 - \sqrt{2}$, $r_3 = \sqrt{6} - \sqrt{3}$, we see that $(1 - \sqrt{2}, 1 + \sqrt{2}) \in \bigcap_{i=1}^3 B(x_i, r_i)$ but $\bigcap_{i=1}^3 B(y_i, r_i) = \emptyset$. To see $\bigcap_{i=1}^3 B(y_i, r_i) = \emptyset$, it suffices to look at the upper half plane in $\mathbb{R}^2$. If we look at the upper half plane H, then $B(y_1, r_1)$ becomes the union of two parts, namely $B(\tilde{y}_1, r_1) \cup H$ and $B(-\tilde{y}_1, r_1) \cup H$, and $B(y_i, r_i)$ becomes $B(\tilde{y}_i, r_i)$ for $i = 2, 3$. But $(B(\tilde{y}_1, r_1) \cup H) \cap B(\tilde{y}_2, r_2) = \emptyset$ and $(B(-\tilde{y}_1, r_1) \cup H) \cap B(\tilde{y}_3, r_3) = \emptyset$. So we obtain that $\bigcap_{i=1}^3 B(y_i, r_i) = \emptyset$.

The following example is given in [7].

Example 3.9. *Property (K) does not hold for $\hat{H}$ with norm d_p.* Specifically, $(\mathbb{R}^m, \mathbb{C}^n/\sim)$ does not have Property K. Let m be any positive integer and $n = 2$, $p = 2$. We want to show that $(X, Y) = (\mathbb{R}^m, \mathbb{C}^n/\sim)$ does not have Property (K). Let $\tilde{y}_1 = (1, 0)$ and $\tilde{y}_2 = (0, \sqrt{3})$ be representitives of y_1, $y_2 \in Y$, respectively. Then $d_1(y_1, y_2) = 4$. Pick any two points x_1, x_2 in X with $\|x_1 - x_2\| = 4$. Then $B(x_1, 2)$ and $B(x_2, 2)$ intersect at $x_3 = (x_1 + x_2)/2 \in X$. It suffices to show that the closed balls $B(y_1, 2)$ and $B(y_2, 2)$ have no intersection in H. Assume on the contrary that the two balls intersect at y_3, then pick a represenitive of y_3, say $\tilde{y}_3 = (a, b)$ where a, $b \in \mathbb{C}$. It can be computed that

$$d_1(y_1, y_3) = |a|^4 + |b|^4 - 2|a|^2 + 2|b|^2 + 2|a|^2|b|^2 + 1 \ , \tag{24}$$

and

$$d_1(y_2, y_3) = |a|^4 + |b|^4 + 6\,|a|^2 - 6\,|b|^2 + 2\,|a|^2\,|b|^2 + 9\ . \tag{25}$$

Set $d_1(y_1, y_3) = d_1(y_2, y_3) = 2$. Take the difference of the right hand side of (24) and (25), we have $|b|^2 - |a|^2 = 1$ and thus $|b|^2 \geq 1$. However, the right hand side of (24) can be rewritten as $(|a|^2 + |b|^2 - 1)^2 + 4\,|b|^2$, so $d_1(y_1, y_3) = 2$ would imply that $|b|^2 \leq 1/2$. This is a contradiction.

Remark 3.10. Using nonlinear functional analysis language ([9]), Lemma 3.4 can be restated by saying that $S^{1,0}(H)$ is a 5-Lipschitz retract in $\mathrm{Sym}(H)$.

Remark 3.11. The Lipschitz inversion results of Theorem 3.3 can be easily extended to systems of quadratic equations, not necessarily of rank-1 matrices from the phase retrieval model considered in this paper.

4. Proof of the Results

4.1. Proof of Theorem 2.2

(i) First we prove the following lemma.

Lemma 4.1. *Fix $x \in \mathbb{C}^n$ and $z \in \mathbb{C}^n$. Let $\xi = \mathbf{j}(x)$ and $\zeta = \mathbf{j}(z)$ be their realifications, respectively. Let $\xi_0 \in \hat{\xi} := \{\mathbf{j}(\tilde{x}) \in \mathbb{R}^{2n} : \tilde{x} \in \hat{x}\}$ be a point in the equivalency class that satisfies $D_2(x, z) = \|\xi_0 - \zeta\|$. Then it is necessary that*

$$\langle \xi_0, J\zeta \rangle = 0 \tag{26}$$

and

$$\langle \xi_0, \zeta \rangle \geq 0\ , \tag{27}$$

where J is defined as in (6).

Proof. For $\theta \in [0, 2\pi)$ define

$$U(\theta) := \cos(\theta) I + \sin(\theta) J\ .$$

Then it is easy to compute that

$$\mathbf{j}(e^{i\theta} x) = U(\theta)\xi\ .$$

Therefore,

$$D_2(x, z) = \min_{\theta \in [0, 2\pi)} \|U(\theta)\xi - \zeta\|^2 = \|\xi\|^2 + \|\zeta\|^2 - 2 \max_{\theta \in [0, 2\pi)} \langle U(\theta)\xi, \zeta \rangle\ .$$

If $\langle U(\theta)\xi, \zeta \rangle$ is constantly zero, then we are done. Otherwise, note that

$$\max_{\theta \in [0, 2\pi)} \langle U(\theta)\xi, \zeta \rangle = \left(\langle \xi, \zeta \rangle^2 + \langle J\xi, \zeta \rangle^2\right)^{\frac{1}{2}}$$

and the maximum is achieved at $\theta = \theta_0$ if and only if

$$\cos(\theta_0) = \frac{\langle \xi, \zeta \rangle}{\left(\langle \xi, \zeta \rangle^2 + \langle J\xi, \zeta \rangle^2\right)^{\frac{1}{2}}}$$

and

$$\sin(\theta_0) = \frac{\langle J\xi, \zeta \rangle}{\left(\langle \xi, \zeta \rangle^2 + \langle J\xi, \zeta \rangle^2\right)^{\frac{1}{2}}} .$$

Now we can compute

$$\begin{aligned}
\langle \xi_0, J\zeta \rangle &= \langle U(\theta_0)\xi, J\zeta \rangle \\
&= \cos(\theta_0) \langle \xi, J\zeta \rangle + \sin(\theta_0) \langle J\xi, J\zeta \rangle \\
&= \frac{\langle \xi, \zeta \rangle}{\left(\langle \xi, \zeta \rangle^2 + \langle J\xi, \zeta \rangle^2\right)^{\frac{1}{2}}} \langle \xi, J\zeta \rangle + \frac{\langle J\xi, \zeta \rangle}{\left(\langle \xi, \zeta \rangle^2 + \langle J\xi, \zeta \rangle^2\right)^{\frac{1}{2}}} \langle J\xi, J\zeta \rangle \\
&= \frac{\langle \xi, \zeta \rangle}{\left(\langle \xi, \zeta \rangle^2 + \langle J\xi, \zeta \rangle^2\right)^{\frac{1}{2}}} \langle -J\xi, \zeta \rangle + \frac{\langle J\xi, \zeta \rangle}{\left(\langle \xi, \zeta \rangle^2 + \langle J\xi, \zeta \rangle^2\right)^{\frac{1}{2}}} \langle \xi, \zeta \rangle \\
&= 0 .
\end{aligned}$$

So we get (26). (27) is obvious. □

Now we come back to the proof of the theorem. Denote

$$p(x, y) := \frac{\|\alpha(x) - \alpha(y)\|^2}{D_2(x, y)^2}, \qquad x, y \in \mathbb{C}^n,\ \hat{x} \neq \hat{y}. \tag{28}$$

We can represent this quotient in terms of ξ and η. It is easy to compute that

$$p(x, y) = P(\xi, \eta) := \frac{\sum_{k=1}^m \langle \Phi_k \xi, \xi \rangle + \langle \Phi_k \eta, \eta \rangle - 2\sqrt{\langle \Phi_k \xi, \xi \rangle \langle \Phi_k \eta, \eta \rangle}}{\|\xi\|^2 + \|\eta\|^2 - 2\sqrt{\langle \xi, \eta \rangle^2 + \langle \xi, J\eta \rangle^2}} . \tag{29}$$

Fix $r > 0$. Take ξ, $\eta \in \mathbb{R}^{2n}$ that satisfy $D_2(x, z) = \|\xi - \zeta\| < r$ and $D_2(y, z) = \|\eta - \zeta\| < r$. Let $\mu = (\xi + \eta)/2$ and $\nu = (\xi - \eta)/2$. Then $\|\nu\| < r$. Note that for r

small enough we have that $\|\mu\| > \|\nu\|$ and that $\Phi_k\zeta \neq 0 \Rightarrow \Phi_k\mu \neq 0$. Thus

$$
\begin{aligned}
P(\xi,\eta) = &\Bigg(\sum_{k=1}^{m} \langle \Phi_k(\mu+\nu), \mu+\nu\rangle + \langle \Phi_k(\mu-\nu), \mu-\nu\rangle - \\
&\qquad 2\sqrt{\langle \Phi_k(\mu+\nu), \mu+\nu\rangle \langle \Phi_k(\mu-\nu), \mu-\nu\rangle}\Bigg) \cdot \\
&\qquad \left(\|\mu+\nu\|^2 + \|\mu-\nu\|^2 - 2\sqrt{\langle \mu+\nu, \mu-\nu\rangle^2 + \langle \mu+\nu, J(\mu-\nu)\rangle^2}\right)^{-1} \\
= &\left(\sum_{k=1}^{m} \langle \Phi_k\mu, \mu\rangle + \langle \Phi_k\nu, \nu\rangle - \sqrt{(\langle \Phi_k\mu, \mu\rangle + \langle \Phi_k\nu, \nu\rangle)^2 - 4\langle \Phi_k\mu, \nu\rangle^2}\right) \cdot \\
&\qquad \left(\|\mu\|^2 + \|\nu\|^2 - \sqrt{\|\mu\|^4 + \|\nu\|^4 - 2\|\mu\|^2\|\nu\|^2 + 4\langle \mu, J\nu\rangle^2}\right)^{-1} \\
\geq &\left(\sum_{k:\Phi_k\zeta\neq 0} \langle \Phi_k\mu, \mu\rangle + \langle \Phi_k\nu, \nu\rangle - \sqrt{(\langle \Phi_k\mu, \mu\rangle + \langle \Phi_k\nu, \nu\rangle)^2 - 4\langle \Phi_k\mu, \nu\rangle^2}\right) \cdot \\
&\qquad \left(\|\mu\|^2 + \|\nu\|^2 - \sqrt{\|\mu\|^4 + \|\nu\|^4 - 2\|\mu\|^2\|\nu\|^2}\right)^{-1} \\
= &\frac{1}{2\|\nu\|^2} \sum_{k:\Phi_k\zeta\neq 0} \langle \Phi_k\mu, \mu\rangle + \langle \Phi_k\nu, \nu\rangle - \sqrt{(\langle \Phi_k\mu, \mu\rangle + \langle \Phi_k\nu, \nu\rangle)^2 - 4\langle \Phi_k\mu, \nu\rangle^2} \\
= &\frac{1}{2\|\nu\|^2} \sum_{k:\Phi_k\zeta\neq 0} \langle \Phi_k\mu, \mu\rangle + \langle \Phi_k\nu, \nu\rangle - \\
&\qquad \langle \Phi_k\mu, \mu\rangle \sqrt{\left(1 + \frac{\langle \Phi_k\nu, \nu\rangle}{\langle \Phi_k\mu, \mu\rangle}\right)^2 - 4\frac{\langle \Phi_k\mu, \nu\rangle^2}{\langle \Phi_k\mu, \mu\rangle^2}} \\
= &\frac{1}{2\|\nu\|^2} \sum_{k:\Phi_k\zeta\neq 0} \langle \Phi_k\mu, \mu\rangle + \langle \Phi_k\nu, \nu\rangle - \\
&\qquad \langle \Phi_k\mu, \mu\rangle \sqrt{1 + 2\frac{\langle \Phi_k\nu, \nu\rangle}{\langle \Phi_k\mu, \mu\rangle} + \frac{\langle \Phi_k\nu, \nu\rangle^2}{\langle \Phi_k\mu, \mu\rangle^2} - 4\frac{\langle \Phi_k\mu, \nu\rangle^2}{\langle \Phi_k\mu, \mu\rangle^2}} \\
= &\frac{1}{2\|\nu\|^2} \sum_{k:\Phi_k\zeta\neq 0} \langle \Phi_k\mu, \mu\rangle + \langle \Phi_k\nu, \nu\rangle - \\
&\qquad \langle \Phi_k\mu, \mu\rangle \left(1 + \frac{\langle \Phi_k\nu, \nu\rangle}{\langle \Phi_k\mu, \mu\rangle} - 2\frac{\langle \Phi_k\mu, \nu\rangle^2}{\langle \Phi_k\mu, \mu\rangle^2}\right) + O(\|\nu\|^4) \\
= &\sum_{k:\Phi_k\zeta\neq 0} \frac{\langle \Phi_k\mu, \nu\rangle^2}{\langle \Phi_k\mu, \mu\rangle \|\nu\|^2} + O(\|\nu\|^2) \\
= &\frac{1}{\|\nu\|^2} \langle \mathcal{S}(\mu)\nu, \nu\rangle + O(\|\nu\|^2) \,.
\end{aligned}
$$

Note that

$$|\langle J\mu,\nu\rangle| = |\langle J\mu,\nu\rangle - \langle J\zeta,\nu\rangle| \le \|J\mu - J\zeta\|\,\|\nu\| = \|\mu-\zeta\|\,\|\nu\| \tag{30}$$

since $\langle J\zeta,\nu\rangle = 0$ by Lemma 4.1. Also, $\|\mu-\zeta\| < r$. Therefore,

$$\|P_{J\mu}\nu\| = \frac{|\langle J\mu,\nu\rangle|}{\|J\mu\|} = \frac{|\langle J\mu,\nu\rangle|}{\|\mu\|} \le \frac{r\,\|\nu\|}{\|\mu\|}$$

and thus

$$\left\|P_{J\mu}^{\perp}\nu\right\|^2 \ge \left(1 - \frac{r^2}{\|\mu\|^2}\right)\|\nu\|^2 \ .$$

As a consequence, we have

$$\begin{aligned} P(\xi,\eta) &= \frac{1}{\|\nu\|^2}\left\langle \mathcal{S}(\mu)P_{J\mu}^{\perp}\nu, P_{J\mu}^{\perp}\nu\right\rangle + O(\|\nu\|^2) \\ &\ge \frac{1}{\left\|P_{J\mu}^{\perp}\nu\right\|^2}\left\langle \mathcal{S}(\mu)P_{J\mu}^{\perp}\nu, P_{J\mu}^{\perp}\nu\right\rangle \left(1 - \frac{r^2}{\|\mu\|^2}\right) + O(r^2) \\ &\ge \left(1 - \frac{r^2}{\|\mu\|^2}\right)\lambda_{2n-1}\left(\mathcal{S}(\mu)\right) + O(r^2) \ . \end{aligned}$$

Take $r \to 0$, by the continuity of eigenvalues with respect to matrix entries we have that

$$A(z) \ge \lambda_{2n-1}(\mathcal{S}(\zeta)) \ . \tag{31}$$

On the other hand, take E_{2n-1} to be the unit-norm eigenvector correspondent to $\lambda_{2n-1}(\mathcal{S}(\zeta))$. For each $r > 0$, take $\xi = \zeta + \frac{r}{2}E_{2n-1}$ and $\eta = \zeta - \frac{r}{2}E_{2n-1}$. Then

$$p(x,y) = P(\xi,\eta) = \lambda_{2n-1}(\mathcal{S}(\zeta)) \ .$$

Hence

$$A(z) \le \lambda_{2n-1}(\mathcal{S}(\zeta)) \ .$$

Together with (31) we have

$$A(z) = \lambda_{2n-1}(\mathcal{S}(\zeta)) \ .$$

(ii) Assume on the contrary that $A_0 = 0$, then for any $N \in \mathbb{N}$, there exist x_N, $y_N \in H$ for which

$$p(x_N,y_N) = \frac{\|\alpha(x_N) - \alpha(y_N)\|^2}{D_2(x_N,y_N)^2} \le \frac{1}{N}. \tag{32}$$

Without loss of generality we assume that $\|x_N\| \ge \|y_N\|$ for each N, for otherwise we can just swap the role of x_N and y_N. Also due to homogeneity we assume $\|x_N\| = 1$. By compactness of the closed ball $\mathcal{B}_1(0) = \{x \in H : \|x\| \le 1\}$ in $H = \mathbb{C}^n$, there exist convergent subsequences of $\{x_N\}_{N\in\mathbb{N}}$ and $\{y_N\}_{N\in\mathbb{N}}$, which to avoid overuse of notations we still denote as $\{x_N\}_{N\in\mathbb{N}} \to x_0 \in H$ and $\{y_N\}_{N\in\mathbb{N}} \to y_0 \in H$.

Since $\|x_0\| = 1$ we have from (i) that $A(x_0) > 0$. Note that $D_2(x_N, y_N) \leq \|x_N\| + \|y_N\| \leq 2$, so by (32) we have $\|\alpha(x_N) - \alpha(y_N)\| \to 0$. That is, $\|\alpha(x_0) - \alpha(y_0)\| = 0$. By injectivity we have $x_0 = y_0$ in $\hat{H}$. By Proposition 2.2(i),

$$p(x_N, y_N) \geq A(x_0) - 1/N > 1/N$$

for N large enough. This is a contradiction with (32).

(iii) The case $z = 0$ is an easy computation. We now present the proof for $z \neq 0$. First we consider $p(x, z) = P(\xi, \zeta)$ as defined in (29). Fix $r > 0$. Take $\xi \in \mathbb{R}^{2n}$ that satisfy $D_2(x, z) = \|\xi - \zeta\| < r$. Let $d = x - z$ and $\delta = \mathbf{j}(d) = \xi - \zeta$. Note that

$$P(\xi,\zeta) = \frac{\sum_{k=1}^{m} \langle \Phi_k\xi, \xi\rangle + \langle \Phi_k\zeta,\zeta\rangle - 2\sqrt{\langle \Phi_k\xi,\xi\rangle \langle \Phi_k\zeta,\zeta\rangle}}{\|\xi\|^2 + \|\zeta\|^2 - 2\sqrt{\langle \xi,\zeta\rangle^2 + \langle \xi, J\zeta\rangle^2}} \; .$$

We can compute its numerator

$$\begin{aligned}
&\sum_{k=1}^{m} \langle \Phi_k\xi, \xi\rangle + \langle \Phi_k\zeta,\zeta\rangle - 2\sqrt{\langle \Phi_k\xi,\xi\rangle \langle \Phi_k\zeta,\zeta\rangle} \\
=&\sum_{k=1}^{m} \langle \Phi_k\zeta,\zeta\rangle + 2\langle \Phi_k\zeta,\delta\rangle + \langle \Phi_k\delta,\delta\rangle + \langle \Phi_k\zeta,\zeta\rangle - \\
&\qquad 2\sqrt{(\langle \Phi_k\zeta,\zeta\rangle + 2\langle \Phi_k\zeta,\delta\rangle + \langle \Phi_k\delta,\delta\rangle)\cdot \langle \Phi_k\zeta,\zeta\rangle} \\
=&\sum_{k:\Phi_k\zeta\neq 0} 2\langle \Phi_k\zeta,\zeta\rangle + 2\langle \Phi_k\zeta,\delta\rangle + \langle \Phi_k\delta,\delta\rangle + \\
&\qquad 2\langle \Phi_k\zeta,\zeta\rangle \left(1 + \frac{\langle \Phi_k\zeta,\zeta\rangle\langle \Phi_k\zeta,\delta\rangle + \frac{1}{2}\langle \Phi_k\zeta,\zeta\rangle\langle \Phi_k\delta,\delta\rangle}{\langle \Phi_k\zeta,\zeta\rangle^2} - \right. \\
&\qquad \left. \frac{1}{8}\cdot\frac{4\langle \Phi_k\zeta,\zeta\rangle^2\langle \Phi_k\zeta,\delta\rangle^2}{\langle \Phi_k\zeta,\zeta\rangle^4} + O\left(\|\delta\|^3\right)\right) + \sum_{k:\Phi_k\zeta=0}\langle \Phi_k\delta,\delta\rangle \\
=&\sum_{k:\Phi_k\zeta\neq 0}\frac{\langle \Phi_k\zeta,\delta\rangle^2}{\langle \Phi_k\zeta,\zeta\rangle} + \sum_{k:\Phi_k\zeta=0}\langle \Phi_k\delta,\delta\rangle + O\left(\|\delta\|^3\right) \;;
\end{aligned}$$

and its denominator

$$\begin{aligned}
&\|\xi\|^2 + \|\zeta\|^2 - 2\sqrt{\langle \xi,\zeta\rangle^2 + \langle \xi, J\zeta\rangle^2} \\
= \;& 2\|\zeta\|^2 + \|\delta\|^2 + 2\langle \zeta,\delta\rangle - 2\|\zeta\|^2 \left(1 + \right. \\
&\qquad \left. \frac{\|\zeta\|^2\langle \zeta,\delta\rangle + \frac{1}{2}\langle \zeta,\delta\rangle + \frac{1}{2}\langle J\zeta,\delta\rangle^2}{\|\zeta\|^4} - \frac{4\|\zeta\|^4\langle \zeta,\delta\rangle^2}{8\|\zeta\|^8} + O\left(\|\delta\|^3\right)\right) \\
= \;& \|\delta\|^2 + O\left(\|\delta\|^3\right) \; .
\end{aligned}$$

We used Lemma 4.1 to get $\langle J\zeta, \delta\rangle = 0$ in the above.

Take $r \to 0$, we see that

$$\tilde{A}(z) \geq \lambda_{2n-1}\left(\mathcal{S}(\zeta) + \sum_{k:\langle z, f_k\rangle = 0} \Phi_k\right).$$

Let $\tilde{E}_{2n-1}$ be the unit-norm eigenvector correspondening to

$$\lambda_{2n-1}\left(\mathcal{S}(\zeta) + \sum_{k:\langle z, f_k\rangle = 0} \Phi_k\right).$$

Note that $\left\langle J\zeta, \tilde{E}_{2n-1}\right\rangle = 0$ since $\mathcal{S}(\zeta)J\zeta = 0$ and $\Phi_k J\zeta = J\Phi_k\zeta = 0$ for each k with $\langle z, f_k\rangle = 0$. Take $\xi = \zeta + \frac{r}{2}\tilde{E}_{2n-1}$ for each r, we again also have

$$\tilde{A}(z) \leq \lambda_{2n-1}\left(\mathcal{S}(\zeta) + \sum_{k:\langle z, f_k\rangle = 0} \Phi_k\right).$$

Therefore

$$\tilde{A}(z) = \lambda_{2n-1}\left(\mathcal{S}(\zeta) + \sum_{k:\langle z, f_k\rangle = 0} \Phi_k\right).$$

(iv) Take $z = 0$ in (iii).

(v) $\tilde{B}(z)$ can be computed in a similar way as in (iii) (in particular, the expansion for $P(\xi, \zeta)$ is exactly the same). We compute $B(z)$. $B(0)$ is computed in [8], Lemma 16. Now we consider $z \neq 0$. Use the same notations as in (29). Fix $r > 0$. Again, take $\xi, \eta \in \mathbb{R}^{2n}$ that satisfy $D_2(x, z) = \|\xi - \zeta\| < r$ and $D_2(y, z) = \|\eta - \zeta\| < r$. Let $\mu = (\xi + \eta)/2$ and $\nu = (\xi - \eta)/2$. Also let $\delta_1 = \xi - \zeta$ and $\delta_2 = \eta - \zeta$. Recall that

$$\begin{aligned} P(\xi, \eta) &= \frac{\sum_{k=1}^m \langle \Phi_k\xi, \xi\rangle + \langle \Phi_k\eta, \eta\rangle - 2\sqrt{\langle \Phi_k\xi, \xi\rangle \langle \Phi_k\eta, \eta\rangle}}{\|\xi\|^2 + \|\eta\|^2 - 2\sqrt{\langle \xi, \eta\rangle^2 + \langle \xi, J\eta\rangle^2}} \\ &= \sum_{k=1}^m \frac{\langle \Phi_k\xi, \xi\rangle + \langle \Phi_k\eta, \eta\rangle - 2\sqrt{\langle \Phi_k\xi, \xi\rangle \langle \Phi_k\eta, \eta\rangle}}{\|\xi\|^2 + \|\eta\|^2 - 2\sqrt{\langle \xi, \eta\rangle^2 + \langle \xi, J\eta\rangle^2}}. \end{aligned}$$

Now we compute it as $\sum_{k=1}^m = \sum_{k:\Phi_k\zeta \neq 0} + \sum_{k:\Phi_k\zeta = 0}$. Again,

$$\begin{aligned} &\sum_{k:\Phi_k\zeta\neq 0} \frac{\langle \Phi_k\xi, \xi\rangle + \langle \Phi_k\eta, \eta\rangle - 2\sqrt{\langle \Phi_k\xi, \xi\rangle \langle \Phi_k\eta, \eta\rangle}}{\|\xi\|^2 + \|\eta\|^2 - 2\sqrt{\langle \xi, \eta\rangle^2 + \langle \xi, J\eta\rangle^2}} \\ =& \sum_{k:\Phi_k\zeta\neq 0} \frac{\langle \Phi_k\mu, \mu\rangle + \langle \Phi_k\nu, \nu\rangle - \sqrt{(\langle \Phi_k\mu, \mu\rangle + \langle \Phi_k\nu, \nu\rangle)^2 - 4\langle \Phi_k\mu, \nu\rangle^2}}{\|\mu\|^2 + \|\nu\|^2 - \sqrt{\|\mu\|^4 + \|\nu\|^4 - 2\|\mu\|^2\|\nu\|^2 + 4\langle \mu, J\nu\rangle^2}}. \end{aligned} \tag{33}$$

Using the same computation as in (i), we get that the numerator is

$$\begin{aligned}
&\sum_{k:\Phi_k\zeta\neq 0} \langle \Phi_k\mu,\mu\rangle + \langle \Phi_k\nu,\nu\rangle - \sqrt{(\langle \Phi_k\mu,\mu\rangle + \langle \Phi_k\nu,\nu\rangle)^2 - 4\langle \Phi_k\mu,\nu\rangle^2}\\
&= 2\langle \mathcal{S}(\mu)\nu,\nu\rangle + O(\|\nu\|^4)\ .
\end{aligned}$$

Since $\mu \neq 0$, the denominator is

$$\begin{aligned}
&\|\mu\|^2 + \|\nu\|^2 - \sqrt{\|\mu\|^4 + \|\nu\|^4 - 2\|\mu\|^2\|\nu\|^2 + 4\langle \mu, J\nu\rangle^2}\\
&= \|\mu\|^2 + \|\nu\|^2 - \|\mu\|^2\sqrt{1 + \frac{\|\nu\|^4}{\|\mu\|^4} - \frac{2\|\nu\|^2}{\|\mu\|^2} + \frac{4\langle \mu, J\nu\rangle^2}{\|\mu\|^4}}\\
&= \|\mu\|^2 + \|\nu\|^2 - \|\mu\|^2\left(1 - \frac{\|\nu\|^2}{\|\mu\|^2} + \frac{2\langle \mu, J\nu\rangle^2}{\|\mu\|^4}\right) + O(\|\nu\|^4) \qquad (34)\\
&= 2\|\nu\|^2 - \frac{2\langle J\mu,\nu\rangle^2}{\|\mu\|^2} + O(\|\nu\|^4)\\
&= 2\|\nu\|^2 + O(\|\nu\|^4) \quad \text{by (30).}
\end{aligned}$$

Also we can compute using the denominator as above (note that $\nu = (\delta_1 - \delta_2)/2$) that

$$\begin{aligned}
&\sum_{k:\Phi_k\zeta=0} \frac{\langle \Phi_k\xi,\xi\rangle + \langle \Phi_k\eta,\eta\rangle - 2\sqrt{\langle \Phi_k\xi,\xi\rangle\langle \Phi_k\eta,\eta\rangle}}{\|\xi\|^2 + \|\eta\|^2 - 2\sqrt{\langle \xi,\eta\rangle^2 + \langle \xi, J\eta\rangle^2}}\\
&= \sum_{k:\Phi_k\zeta=0} \frac{\left(\left\|\Phi_k^{1/2}\delta_1\right\| - \left\|\Phi_k^{1/2}\delta_2\right\|\right)^2}{\|\delta_1 - \delta_2\|^2 + O(\|\nu\|^4)}\ . \qquad (35)
\end{aligned}$$

Now put together (33), (34) and (35), we get

$$P(\xi,\eta) = \frac{\langle \mathcal{S}(\mu)\nu,\nu\rangle + O(\|\nu\|^4)}{\|\nu\|^2 + O(\|\nu\|^4)} + \sum_{k:\Phi_k\zeta=0} \frac{\left(\left\|\Phi_k^{1/2}\delta_1\right\| - \left\|\Phi_k^{1/2}\delta_2\right\|\right)^2}{\|\delta_1 - \delta_2\|^2 + O(\|\nu\|^4)}\ .$$

Note that

$$\left(\left\|\Phi_k^{1/2}\delta_1\right\| - \left\|\Phi_k^{1/2}\delta_2\right\|\right)^2 \leq \langle \Phi_k(\delta_1 - \delta_2), \delta_1 - \delta_2\rangle$$

since it is equivalent to

$$\langle \Phi_k\delta_1,\delta_1\rangle\langle \Phi_k\delta_2,\delta_2\rangle \geq (\langle \Phi_k\delta_1,\delta_2\rangle)^2\ , \qquad (36)$$

which is the Cauchy-Schwarz inequality. Therefore, we have that

$$P(\xi,\eta) \leq \frac{\left\langle (\mathcal{S}(\mu) + \sum_{k:\Phi_k\zeta=0}\Phi_k)\nu,\nu\right\rangle + O(\|\nu\|^4)}{\|\nu\|^2 + O(\|\nu\|^4)} \leq \lambda_1\left(\mathcal{S}(\mu) + \sum_{k:\Phi_k\zeta=0}\Phi_k\right) + O(r^2)\ .$$

Take $r \to 0$ we have that

$$B(z) \le \lambda_1 \left(\mathcal{S}(\zeta) + \sum_{k:\Phi_k \zeta = 0} \Phi_k \right) .$$

Again we get the other direction of the above inequality by taking $\xi = \zeta + \frac{r}{2} E_1$ and $\eta = \zeta - \frac{r}{2} E_1$ for each $r > 0$ where E_1 is the unit-norm eigenvector correspondent to $\lambda_1 \left(\mathcal{S}(\zeta) + \sum_{k:\langle z, f_k \rangle = 0} \Phi_k \right)$. Note that for each r, the equality in (36) holds for this pair of ξ and η.

(vi) Take $z = 0$ in (v).

4.2. Proof of Theorem 2.5

Only the first two parts are nontrivial. We prove them as follows.

Fix $z \in \mathbb{C}^n$. Take $x = z + d_1$ and $y = z + d_2$ with $\|d_1\| < r$ and $\|d_2\| < r$ for r small. Let $u = x + y = 2z + d_1 + d_2$ and $v = x - y = d_1 - d_2$. Let $\mu = 2\zeta + \delta_1 + \delta_2 \in \mathbb{R}^{2n}$ and $\nu = \delta_1 - \delta_2 \in \mathbb{R}^{2n}$ be the realification of u and v, respectively. Define

$$\rho(x, y) = \frac{\|\beta(x) - \beta(y)\|^2}{d_1(x, y)^2} .$$

By the same computation as in [3], Section 4.1, we get

$$\rho(x, y) = Q(\zeta; \delta_1, \delta_2) := \frac{\langle \mathcal{R}(2\zeta + \delta_1 + \delta_2)(\delta_1 - \delta_2), \delta_1 - \delta_2 \rangle}{\|2\zeta + \delta_1 + \delta_2\|^2 \left\langle P^{\perp}_{J(2\zeta + \delta_1 + \delta_2)}(\delta_1 - \delta_2), \delta_1 - \delta_2 \right\rangle} .$$

Since $J(2\zeta + \delta_1 + \delta_2) \in \ker \mathcal{R}(2\zeta + \delta_1 + \delta_2)$, we have

$$Q(\zeta; \delta_1, \delta_2) = \frac{\left\langle \mathcal{R}(2\zeta + \delta_1 + \delta_2) P^{\perp}_{J(2\zeta + \delta_1 + \delta_2)}(\delta_1 - \delta_2), P^{\perp}_{J(2\zeta + \delta_1 + \delta_2)}(\delta_1 - \delta_2) \right\rangle}{\|2\zeta + \delta_1 + \delta_2\|^2 \left\langle P^{\perp}_{J(2\zeta + \delta_1 + \delta_2)}(\delta_1 - \delta_2), \delta_1 - \delta_2 \right\rangle} .$$

Now let $\delta = \delta_1 + \delta_2$ and $\nu = \delta_1 - \delta_2$. Note the set inclusion relation

$$\begin{aligned} &\left\{ \delta_1, \delta_2 \in \mathbb{R}^{2n} : \quad \|\delta\| < \frac{r}{2},\ \|\nu\| < \frac{r}{2},\ \nu \perp J(2\zeta + \delta) \right\} \\ \subset &\left\{ \delta_1, \delta_2 \in \mathbb{R}^{2n} : \quad \|\delta_1\| < r,\ \|\delta_2\| < r,\ \nu \perp J(2\zeta + \delta) \right\} \\ \subset &\left\{ \delta_1, \delta_2 \in \mathbb{R}^{2n} : \quad \|\delta\| < 2r,\ \|\nu\| < 2r,\ \nu \perp J(2\zeta + \delta) \right\} . \end{aligned}$$

Thus we have

$$\inf_{\substack{\|\delta\| < 2r \\ \|\nu\| < 2r \\ \nu \perp J(2\zeta + \delta)}} Q(\zeta; \delta_1, \delta_2) \quad \le \inf_{\substack{\|\delta_1\| < r \\ \|\delta_2\| < r \\ \nu \perp J(2\zeta + \delta)}} Q(\zeta; \delta_1, \delta_2) \quad \le \inf_{\substack{\|\delta\| < r/2 \\ \|\nu\| < r/2 \\ \nu \perp J(2\zeta + \delta)}} Q(\zeta; \delta_1, \delta_2) .$$

That is,

$$\inf_{\|\delta\|<2r} \frac{\lambda_{2n-1}(\mathcal{R}(2\zeta+\delta))}{\|2\zeta+\delta\|^2} \quad \leq \inf_{\substack{\|\delta_1\|<r \\ \|\delta_2\|<r \\ \nu\perp J(2\zeta+\delta)}} Q(\zeta;\delta_1,\delta_2) \quad \leq \inf_{\|\delta\|<r/2} \frac{\lambda_{2n-1}(\mathcal{R}(2\zeta+\delta))}{\|2\zeta+\delta\|^2} \ .$$

Take $r \to 0$, by the continuity of eigenvalues with respect to the matrix entries, we have

$$\lambda_{2n-1}(\mathcal{R}(\zeta))/\left\|\zeta\right\|^2 \leq a(z) \leq \lambda_{2n-1}(\mathcal{R}(\zeta))/\left\|\zeta\right\|^2 \ .$$

That is,

$$a(z) = \lambda_{2n-1}(\mathcal{R}(\zeta))/\left\|\zeta\right\|^2 \ .$$

Now consider

$$\rho(x,z) = \frac{\|\beta(x)-\beta(z)\|^2}{d_1(x,z)^2} \ .$$

For simplicity write $\delta = \delta_1$. We can compute that

$$\rho(x,z) = Q(\zeta;\delta) = \frac{\langle \mathcal{R}(2\zeta+\delta)\delta,\delta\rangle}{\|2\zeta+\delta\|^2 \left\langle P^{\perp}_{J(2\zeta+\delta)}\delta,\delta\right\rangle} = \frac{\left\langle \mathcal{R}(2\zeta+\delta)P^{\perp}_{J(2\zeta+\delta)}\delta, P^{\perp}_{J(2\zeta+\delta)}\delta\right\rangle}{\|2\zeta+\delta\|^2 \left\langle P^{\perp}_{J(2\zeta+\delta)}\delta,\delta\right\rangle} \ .$$

Note that

$$\inf_{\substack{\|\delta\|<r \\ \delta\perp J(2\zeta+\delta)}} Q(\zeta;\delta) \quad \geq \quad \inf_{\|\sigma\|<r} \inf_{\substack{\|\delta\|<r \\ \delta\perp J(2\zeta+\delta)}} Q(\zeta;\delta) \quad = \quad \inf_{\|\sigma\|<r} \lambda_{2n-1}(\mathcal{R}(2\zeta+\delta)) \ .$$

Take $r \to 0$ we have that

$$\tilde{a}(z) \geq \lambda_{2n-1}(\mathcal{R}(2\zeta))/\left\|2\zeta\right\|^2 = \lambda_{2n-1}(\mathcal{R}(\zeta))/\left\|\zeta\right\|^2 \ .$$

On the other hand, take $\tilde{e}_{2n-1}$ to be a unit-norm eigenvector correspondent to $\lambda_{2n-1}(\mathcal{R}(2\zeta))$. Then by the continuity of eigenvalues with respect to the matrix entries, for any $\varepsilon > 0$, there exists $t > 0$ so that $\delta = t\tilde{e}_{2n-1}$ satisfy

$$\frac{\langle \mathcal{R}(2\zeta+\delta)\delta,\delta\rangle}{\left\langle P^{\perp}_{J(2\zeta+\delta)}\delta,\delta\right\rangle} \leq \lambda_{2n-1}(\mathcal{R}(2\zeta)) + \varepsilon$$

and from there we have

$$\tilde{a}(z) \leq \lambda_{2n-1}(\mathcal{R}(2\zeta))/\left\|2\zeta\right\|^2 = \lambda_{2n-1}(\mathcal{R}(\zeta))/\left\|\zeta\right\|^2 \ .$$

Therefore,

$$\tilde{a}(z) = \lambda_{2n-1}(\mathcal{R}(\zeta))/\left\|\zeta\right\|^2 \ .$$

In a similar way (replacing infimum by supremum) we also get $b(z)$ and $\tilde{b}(z)$ as stated in the theorem.

4.3. Proof of Proposition 3.1

(i) Obviously $D_p(\hat{x},\hat{y}) \geq 0$ for any $\hat{x}$, $\hat{y} \in \hat{H}$ and $D_p(\hat{x},\hat{y}) = 0$ if and only if $\hat{x} = \hat{y}$. Also $D_p(\hat{x},\hat{y}) = D_p(\hat{y},\hat{x})$ since $\|x - ay\|_p = \|y - a^{-1}x\|_p$ for any x, $y \in H$, $|a| = 1$. Moreover, for any $\hat{x}$, $\hat{y}$, $\hat{z} \in \hat{H}$, fix $D_p(\hat{x},\hat{y}) = \|x - ay\|_p$, $D_p(\hat{y},\hat{z}) = \|z - by\|$, then

$$\begin{aligned} D_p(\hat{x},\hat{z}) &\leq \left\|x - ab^{-1}z\right\|_p = \|bx - az\|_p \\ &\leq \|bx - aby\|_p + \|aby - az\|_p = D_p(\hat{x},\hat{y}) + D_p(\hat{y},\hat{z}) \ . \end{aligned}$$

Therefore D_p is a metric. d_p is also a metric since $\|\cdot\|_p$ in the definition of d_p is the standard Schatten p-norm of a matrix.

(ii) For $p \leq q$, by Hölder's inequality we have for any $x = (x_1, x_2, ..., x_n) \in H$ that $\sum_{i=1}^n |x_i|^p \leq n^{(\frac{1}{p}-\frac{1}{q})}(\sum_{i=1}^n |x_i|^q)^{\frac{p}{q}}$. Thus $\|x\|_p \leq n^{(\frac{1}{p}-\frac{1}{q})}\|x\|_q$. Also, since $\|\cdot\|_p$ is homogeneous, we can assume $\|x\|_p = 1$. Then $\sum_{i=1}^n |x_i|^q \leq \sum_{i=1}^n |x_i|^p = 1$. Thus $\|x\|_q \leq \|x\|_p$. Therefore, we have $D_q(\hat{x},\hat{y}) = \|x - a_1 y\|_q \geq n^{(\frac{1}{p}-\frac{1}{q})}\|x - a_1 y\|_p \geq n^{(\frac{1}{p}-\frac{1}{q})} D_p(\hat{x},\hat{y})$ and $D_p(\hat{x},\hat{y}) = \|x - a_2 y\|_p \geq \|x - a_2 y\|_q \geq D_q(\hat{x},\hat{y})$ for some a_1, a_2 with magnitude 1. Hence

$$D_q(\hat{x},\hat{y}) \leq D_p(\hat{x},\hat{y}) \leq n^{(\frac{1}{p}-\frac{1}{q})} D_q(\hat{x},\hat{y}) \ .$$

We see that $(D_p)_{1\leq p\leq\infty}$ are equivalent. The second part follows then immediately.

(iii) The proof is similar to (ii). Note that there are at most 2 σ_i's that are nonzero, so we have $2^{(\frac{1}{p}-\frac{1}{q})}$ instead of $n^{(\frac{1}{p}-\frac{1}{q})}$.

(iv) To prove that D_p and d_q are equivalent, we need only to show that each open ball with respect to D_p contains an open ball with respect to d_p, and vice versa. By (ii) and (iii), it is sufficient to consider the case when $p = q = 2$.
First, we fix $x \in H = \mathbb{C}^n$, $r > 0$. Let $R = \min(1, rn^{-2}(2\|x\|_\infty + 1)^{-1})$. Then for any $\hat{y}$ such that $D_2(\hat{x},\hat{y}) < R$, we take y such that $\|x - y\| < R$, then $\forall 1 \leq i, j \leq n$, $|x_i\overline{x_j} - y_i\overline{y_j}| = |x_i(\overline{x_j} - \overline{y_j}) + (x_i - y_i)\overline{y_j}| < |x_i|R + R(|x_i| + R) = R(2|x_i| + R) \leq R(2|x_i| + 1) \leq \frac{r}{n^2}$. Hence $d_2(\hat{x},\hat{y}) = \|xx^* - yy^*\|_2 < n^2 \cdot \frac{r}{n^2} = r$.
On the other hand, we fix $x \in H = \mathbb{C}^n$, $R > 0$. Let $r = R^2/\sqrt{2}$. Then for any $\hat{y}$ such that $d_2(\hat{x},\hat{y}) < r$, we have

$$(d_2(\hat{x},\hat{y}))^2 = \|x\|^4 + \|y\|^4 - 2|\langle x,y\rangle|^2 < r^2 = \frac{R^4}{2} \ .$$

But we also have

$$(D_2(\hat{x},\hat{y}))^2 = \min_{|a|=1}\|x - ay\|^2 = \left\|x - \frac{\langle x,y\rangle}{|\langle x,y\rangle|}y\right\|^2 = \|x\|^2 + \|y\|^2 - 2|\langle x,y\rangle| \ ,$$

so

$$(D_2(\hat{x},\hat{y}))^4 = \|x\|^4 + \|y\|^4 + 2\|x\|^2\|y\|^2 - 4(\|x\|^2 + \|y\|^2)|\langle x,y\rangle| + 4|\langle x,y\rangle|^2 \ .$$

Since $|\langle x,y\rangle| \leq \|x\|\,\|y\| \leq (\|x\|^2+\|y\|^2)/2$, we can easily check that $(D_2(\hat{x},\hat{y}))^4 \leq 2(d_2(\hat{x},\hat{y}))^2 < R^4$. Hence $D_2(\hat{x},\hat{y}) < R$.
Thus D_2 and d_2 are indeed equivalent metrics. Therefore D_p and d_q are equivalent. Also, the imbedding i is not Lipschitz: if we take $x=(x_1,0,\ldots,0)\in\mathbb{C}^n$, then $D_2(\hat{x},0)=|x_1|$, $d_2(\hat{x},0)=|x_1|^2$.

(v) First, for $p=2$, for $\hat{x}\neq\hat{y}$ in $\hat{H}-\{0\}$, we compute the quotient

$$\begin{aligned}
\rho(x,y) &= \frac{\|\kappa_\alpha(x)-\kappa_\alpha(y)\|^2}{D_2(x,y)^2}\\
&= \frac{\left\|\|x\|^{-1}xx^* - \|y\|^{-1}yy^*\right\|^2}{\|x\|^2+\|y\|^2-2\,|\langle x,y\rangle|}\\
&= \frac{\|xx^*\|^2\,\|y\|^2+\|x\|^2\,\|yy^*\|^2-2\,\|x\|\,\|y\|\,\mathrm{trace}(xx^*yy^*)}{\|x\|^4\,\|y\|^2+\|x\|^2\,\|y\|^4-2\,\|x\|^2\,\|y\|^2\,|x^*y|}\\
&= 1+\frac{2\,\|x\|\,\|y\|\,(\|x\|\,\|y\|\,|x^*y|-\mathrm{trace}(xx^*yy^*))}{\|x\|^4\,\|y\|^2+\|x\|^2\,\|y\|^4-2\,\|x\|^2\,\|y\|^2\,|x^*y|}\\
&= 1+\frac{2\,(\|x\|\,\|y\|\,|x^*y|-\mathrm{trace}(xx^*yy^*))}{\|x\|^3\,\|y\|+\|x\|\,\|y\|^3-2\,\|x\|\,\|y\|\,|x^*y|}\ ,
\end{aligned}$$

where we used $\|xx^*\|=\|x\|^2$. For simplicity write $a=\|x\|$, $b=\|y\|$ and $t=|\langle x,y\rangle|\cdot(\|x\|\,\|y\|)^{-1}$. We have $a>0$, $b>0$ and $0\leq t\leq 1$.
Now

$$\rho(x,y)=1+\frac{2(abt-abt^2)}{a^2+b^2-2abt}\ .$$

Obviously $\rho(x,y)\geq 1$. Now we prove that $\rho(x,y)\leq 2$. Note that

$$1+\frac{2(abt-abt^2)}{a^2+b^2-2abt}\leq 2 \;\Leftrightarrow\; a^2+b^2-4abt+2abt^2\geq 0\ ,$$

but

$$a^2+b^2-4abt+2abt^2\geq 2ab-4abt+2abt^2=2ab(t-1)^2\geq 0,$$

so we are done. Note that take any x, y with $\langle x,y\rangle=0$ we would have $\rho(x,y)=1$. On the other hand, taking $\|x\|=\|y\|$ and let $t\to 1$ we see that $\rho(x,y)=2-\varepsilon$ is achievable for any small $\varepsilon>0$. Therefore the constants are optimal. The case where one of x and y is zero would not break the constraint of these two constants. Therefore after taking the square root, we get lower Lipschitz constant 1 and upper Lipschitz constant $\sqrt{2}$.
For other p, we use the results in (ii) and (iii) to get that the lower Lipschitz constant for κ_α is $\min(2^{\frac{1}{2}-\frac{1}{p}}, n^{\frac{1}{p}-\frac{1}{2}})$ and the upper Lipschitz constant is $\sqrt{2}\max(n^{\frac{1}{2}-\frac{1}{p}}, 2^{\frac{1}{p}-\frac{1}{2}})$.

(vi) This follows directly from the construction of the map.

(vii) This follows directly from (v) and (vi).

4.4. Proof of Lemma 3.4

(ii) follows directly from the expression of π. We prove (i) below.

Let A, $B \in \mathrm{Sym}(H)$ where $A = \sum_{k=1}^{d} \lambda_{m(k)} P_k$ and $B = \sum_{k'=1}^{d'} \mu_{m(k')} Q_{k'}$. We now show that

$$\|\pi(A) - \pi(B)\|_p \leq (3 + 2^{1+\frac{1}{p}}) \|A - B\|_p \ . \tag{37}$$

Assume $\lambda_1 - \lambda_2 \leq \mu_1 - \mu_2$. Otherwise switch the notations for A and B. If $\mu_1 - \mu_2 = 0$ then $\pi(A) = \pi(B) = 0$ and the inequality (37) is satisfied. Assume now $\mu_1 - \mu_2 > 0$. Thus Q_1 is of rank 1 and $\|Q_1\|_p = 1$ for all p.

First we consider the case $\lambda_1 - \lambda_2 > 0$. In this case P_1 is of rank 1, and we have

$$\pi(A) - \pi(B) = (\lambda_1 - \lambda_2)P_1 - (\mu_1 - \mu_2)Q_1 = (\lambda_1 - \lambda_2)(P_1 - Q_1) + (\lambda_1 - \mu_1 - (\lambda_2 - \mu_2))Q_1 \ . \tag{38}$$

Here $\|P_1\|_\infty = \|Q_1\|_\infty = 1$. Therefore we have $\|P_1 - Q_1\|_\infty \leq 1$ since P_1, $Q_1 \geq 0$. From that we have $\|P_1 - Q_1\|_p \leq 2^{\frac{1}{p}}$.

Also, by Weyl's inequality we have $|\lambda_i - \mu_i| \leq \|A - B\|_\infty$ for each i. Apply this to $i = 1$, 2 we get $|\lambda_1 - \mu_1 - (\lambda_2 - \mu_2)| \leq |\lambda_1 - \mu_1| + |\lambda_2 - \mu_2| \leq 2\|A - B\|_\infty$. Thus $|\lambda_1 - \mu_1| + |\lambda_2 - \mu_2| \leq 2\|A - B\|_\infty \leq 2\|A - B\|_p$.

Let $g := \lambda_1 - \lambda_2$, $\delta := \|A - B\|_p$, then apply the above inequality to (38) we get

$$\|\pi(A) - \pi(B)\|_p \leq g \|P_1 - Q_1\|_p + 2\delta \leq 2^{\frac{1}{p}} g + 2\delta \ . \tag{39}$$

If $0 < g \leq (2 + 2^{-\frac{1}{p}})\delta$, then $\|\pi(A) - \pi(B)\|_p \leq (2^{1+\frac{1}{p}} + 3)\delta$ and we are done.

Now we consider the case where $g > (2 + 2^{-\frac{1}{p}})\delta$. Note that in this case we have $\delta < g/2$. Thus we have $|\lambda_1 - \mu_1| < g/2$ and $|\lambda_2 - \mu_2| < g/2$. That means $\mu_1 > (\lambda_1 + \lambda_2)/2$ and $\mu_2 < (\lambda_1 + \lambda_2)/2$. Therefore, we can use holomorphic functional calculus and put

$$P_1 = -\frac{1}{2\pi i} \oint_\gamma R_A dz$$

and

$$Q_1 = -\frac{1}{2\pi i} \oint_\gamma R_B dz$$

where $R_A = (A - zI)^{-1}$, $R_B = (B - zI)^{-1}$, and $\gamma = \gamma(t)$ is the contour given in Figure 2 (note that γ encloses μ_1 but not μ_2) and used also by [15]. Therefore we have

$$\|P_1 - Q_1\|_p \leq \frac{1}{2\pi} \int_I \|(R_A - R_B)(\gamma(t))\|_p |\gamma'(t)| dt \ . \tag{40}$$

Now we have

$$(R_A - R_B)(z) = R_A(z) - (I + R_A(z)(B - A))^{-1} R_A(z) = \sum_{n \geq 1} (-1)^n (R_A(z)(B - A))^n R_A(z) \ , \tag{41}$$

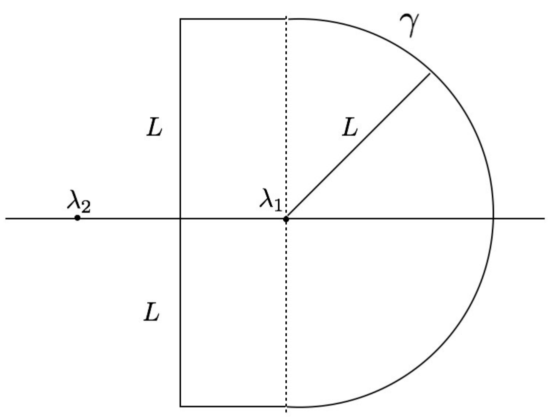

Figure 2: Contour for the integrals

since for large L we have $\|R_A(z)(B-A)\|_\infty \le \|R_A(z)\|_\infty \|B-A\|_p \le \frac{\delta}{\mathrm{dist}(z,\sigma(A))} \le \frac{2\delta}{g} < \frac{2}{2+2^{-\frac{1}{p}}} < 1$, where $\sigma(A)$ denotes the spectrum of A. Therefore we have

$$\begin{aligned}\|(R_A - R_B)(\gamma(t))\|_p &\le \sum_{n\ge 1} \|R_A(\gamma(t))\|_\infty^{n+1} \|A-B\|_p^n \\ &= \frac{\|R_A(\gamma(t))\|_\infty^2 \|A-B\|_p}{1 - \|R_A(\gamma(t))\|_\infty \|A-B\|_p} < \frac{\|A-B\|_p}{\mathrm{dist}^2(\gamma(t),\sigma(A))} \cdot (2^{1+\frac{1}{p}}+1) ,\end{aligned} \tag{42}$$

since $\mathrm{dist}(\gamma(t),\sigma(A)) \ge g/2$ for each t for large L. Here we used the fact that if we order the singular values of any matrix X such that $\sigma_1(X) \ge \sigma_2(X) \ge \cdots$, then for any i we have $\sigma_i(XY) \le \sigma_1(X)\sigma_i(Y)$, and thus for two operators X, $Y \in \mathrm{Sym}(H)$, we have $\|XY\|_p \le \|X\|_\infty \|Y\|_p$.

Hence by (40) and (42) we have

$$\|P_1 - Q_1\|_p \le (2^{\frac{1}{p}} + 2^{-1}) \frac{\|A-B\|_p}{\pi} \int_I \frac{1}{\mathrm{dist}^2(\gamma(t),\sigma(A))} |\gamma'(t)| dt . \tag{43}$$

By evaluating the integral and letting L approach infinity for the contour, we have as in [15]

$$\int_I \frac{1}{\mathrm{dist}^2(\gamma(t),\sigma(A))} |\gamma'(t)| dt = 2\int_0^\infty \frac{1}{t^2 + (\frac{g}{2})^2} dt = \left[\frac{4}{g} \arctan\left(\frac{2t}{g}\right)\right]_0^\infty = \frac{2\pi}{g} . \tag{44}$$

Hence

$$\|P_1 - Q_1\|_p \le (2^{\frac{1}{p}} + 2^{-1}) \frac{\|A-B\|_p}{\pi} \cdot \frac{2\pi}{g} = (2^{1+\frac{1}{p}} + 1)\frac{\delta}{g} . \tag{45}$$

Thus by the first inequality in (39) and (45) we have $\|\pi(A) - \pi(B)\|_p \le (3 + 2^{1+\frac{1}{p}})\delta$.

Now we are left with the case $\lambda_1 - \lambda_2 = 0 < \mu_1 - \mu_2$. Note that in this case we have that $\pi(A) - \pi(B) = -(\mu_1 - \mu_2)Q_1 = ((\lambda_1 - \mu_1) - (\lambda_2 - \mu_2))Q_1$, and therefore

$$\|\pi(A) - \pi(B)\|_p \leq 2\,\|A - B\|_p < (3 + 2^{1+\frac{1}{p}})\,\|A - B\|_p \ .$$

We have proved that $\|\pi(A) - \pi(B)\|_p \leq (3 + 2^{1+\frac{1}{p}})\,\|A - B\|_p$. That is to say, $\pi : (\mathrm{Sym}(H), \|\cdot\|_p) \to (S^{1,0}(H), \|\cdot\|_p)$ is Lipschitz continuous with $\mathrm{Lip}(\pi) \leq 3 + 2^{1+\frac{1}{p}}$.

Now we are ready to prove Theorem 3.3.

4.5. Proof of Theorem 3.3

The proof for α and β are the same in essence. For simplicity we do it for β first.

We need to construct a map $\psi : (\mathbb{R}^m, \|\cdot\|_p) \to (\hat{H}, d_q)$ so that $\psi(\beta(x)) = x$ for all $x \in \hat{H}$, and ψ is Lipschitz continuous. We prove the Lipschitz bound (15), which implies (14) for $p = 2$ and $q = 1$.

Set $M = \beta(\hat{H}) \subset \mathbb{R}^m$. By the result in Section 2.3, there is a map $\tilde{\psi}_1 : M \to \hat{H}$ that is Lipschitz continuous and satisfies $\tilde{\psi}_1(\beta(x)) = x$ for all $x \in \hat{H}$. Additionally, the Lipschitz bound between $(M, \|\cdot\|_2)$ (that is, M with Euclidean distance) and $(\hat{H}, d_1)$ is given by $1/\sqrt{a_0}$.

First we change the metric on $\hat{H}$ from d_1 to d_2 and embed isometrically $\hat{H}$ into $\mathrm{Sym}(H)$ with Frobenius norm (i.e. the Euclidean metric):

$$(M, \|\cdot\|_2) \xrightarrow{\tilde{\psi}_1} (\hat{H}, d_1) \xrightarrow{i_{1,2}} (\hat{H}, d_2) \xrightarrow{\kappa_\beta} (\mathrm{Sym}(H), \|\cdot\|_{Fr}) \ , \tag{46}$$

where $i_{1,2}(x) = x$ is the identity of $\hat{H}$ and κ_β is the isometry (10) . We obtain a map $\tilde{\psi}_2 : (M, \|\cdot\|_2) \to (\mathrm{Sym}(H), \|\cdot\|_{Fr})$ of Lipschitz constant

$$\mathrm{Lip}(\tilde{\psi}_2) \leq \mathrm{Lip}(\tilde{\psi}_1)\mathrm{Lip}(i_{1,2})\mathrm{Lip}(\kappa_\beta) = \frac{1}{\sqrt{a_0}} \ ,$$

where we used $\mathrm{Lip}(i_{1,2}) = L^d_{1,2,n} = 1$ by (8).

Kirszbraun Theorem [14] extends isometrically $\tilde{\psi}_2$ from M to the entire $\mathbb{R}^m$ with Euclidean metric $\|\cdot\|$. Thus we obtain a Lipschitz map $\psi_2 : (\mathbb{R}^m, \|\cdot\|) \to (\mathrm{Sym}(H), \|\cdot\|_{Fr})$ of Lipschitz constant $\mathrm{Lip}(\psi_2) = \mathrm{Lip}(\tilde{\psi}_2) \leq \frac{1}{\sqrt{a_0}}$ so that $\psi_2(\beta(x)) = xx^*$ for all $x \in \hat{H}$.

The third step is to piece together ψ_2 with norm changing identities. For $q \leq 2$ we consider the following maps:

$$\begin{aligned}(\mathbb{R}^m, \|\cdot\|_p) &\xrightarrow{j_{p,2}} (\mathbb{R}^m, \|\cdot\|_2) \xrightarrow{\psi_2} (\mathrm{Sym}(H), \|\cdot\|_{Fr}) \\ &\xrightarrow{\pi} (S^{1,0}(H), \|\cdot\|_{Fr}) \xrightarrow{\kappa_\beta^{-1}} (\hat{H}, d_2) \xrightarrow{i_{2,q}} (\hat{H}, d_q) \ ,\end{aligned} \tag{47}$$

where $j_{p,2}$ and $i_{2,q}$ are identity maps on the respective spaces that change the metric and π is the map defined in Eq. (23). The map ψ claimed by Theorem 3.3 is obtained by composing:

$$\psi : (\mathbb{R}^m, \|\cdot\|_p) \to (\hat{H}, d_q) \ \ , \ \ \psi = i_{2,q} \cdot \kappa_\beta^{-1} \cdot \pi \cdot \psi_2 \cdot j_{p,2} \ .$$

Its Lipschitz constant is bounded by

$$\begin{aligned}\mathrm{Lip}(\psi)_{p,q} &\leq \mathrm{Lip}(j_{p,2})\mathrm{Lip}(\psi_2)\mathrm{Lip}(\pi)\mathrm{Lip}(\kappa_\beta^{-1})\mathrm{Lip}(i_{2,q}) \\ &\leq \max(1, m^{\frac{1}{2}-\frac{1}{p}})\frac{1}{\sqrt{a_0}} \cdot (3+2\sqrt{2}) \cdot 1 \cdot 2^{\frac{1}{q}-\frac{1}{2}} \ .\end{aligned}$$

Hence we obtained (20). The other equation (14) follows for $p=2$ and $q=1$.

For $q>2$ we use:

$$\begin{aligned}(\mathbb{R}^m, \|\cdot\|_p) &\xrightarrow{j_{p,2}} (\mathbb{R}^m, \|\cdot\|_2) \xrightarrow{\psi_2} (\mathrm{Sym}(H), \|\cdot\|_{Fr}) \\ &\xrightarrow{I_{2,q}} (\mathrm{Sym}(H), \|\cdot\|_q) \xrightarrow{\pi} (S^{1,0}(H), \|\cdot\|_q) \xrightarrow{\kappa_\beta^{-1}} (\hat{H}, d_q) \ ,\end{aligned} \tag{48}$$

where $j_{p,2}$ and $I_{2,q}$ are identity maps on the respective spaces that change the metric. The map ψ claimed by Theorem 3.3 is obtained by composing:

$$\psi : (\mathbb{R}^m, \|\cdot\|_p) \to (\hat{H}, d_q) \quad , \quad \psi = \kappa_\beta^{-1} \cdot \pi \cdot I_{2,q} \cdot \psi_2 \cdot j_{p,2} \ .$$

Its Lipschitz constant is bounded by

$$\mathrm{Lip}(\psi)_{p,q} \leq \mathrm{Lip}(j_{p,2})\mathrm{Lip}(\psi_2)\mathrm{Lip}(I_{2,q})\mathrm{Lip}(\pi)\mathrm{Lip}(\kappa_\beta^{-1}) \leq \max(1, m^{\frac{1}{2}-\frac{1}{p}})\frac{1}{\sqrt{a_0}} \cdot 1 \cdot (3+2^{1+\frac{1}{q}}) \cdot 1 \ .$$

Hence we obtained (21).

Replace β by α, ψ by ω, and κ_β by κ_α in the proof above, using the Lipschitz constants for κ_α in Proposition 3.1, we obtain (16) and (17).

Acknowledgements

The authors were supported in part by NSF grant DMS-1109498 and DMS-1413249. The first author acknowledges fruitful discussions with Krzysztof Nowak and Hugo Woerdeman (both from Drexel University) who pointed out several references, with Stanislav Minsker (Duke University) for pointing out [15] and [12], and Vern Paulsen (University of Houston), Marcin Bownick (University of Oregon) and Friedrich Philipp (University of Berlin). We also thank the reviewers for their constructive comments and suggestions.

References

[1] R. Balan, *On Signal Reconstruction from Its Spectrogram*, Proceedings of the CISS Conference, Princeton NJ, May 2010.

[2] R. Balan, *Reconstruction of Signals from Magnitudes of Redundant Representations*, available online arXiv:1207.1134v1 [math.FA] 4 July 2012.

[3] R. Balan, *Reconstruction of Signals from Magnitudes of Redundant Representations: the Complex Case*, to appear in Foundations of Computational Mathematics 2015.

[4] R. Balan, P. Casazza, D. Edidin, *On signal reconstruction without phase*, Appl.Comput.Harmon.Anal. **20** (2006), 345–356.

[5] R. Balan, B. Bodmann, P. Casazza, D. Edidin, *Painless reconstruction from Magnitudes of Frame Coefficients*, J.Fourier Anal.Applic., **15** (4) (2009), 488–501.

[6] R. Balan, Y. Wang, *Invertibility and Robustness of Phaseless Reconstruction*, Appl. Comp. Harm. Anal., **38**(3) (2015), 469–488.

[7] R. Balan, D. Zou, *On Lipschitz Inversion of Nonlinear Redundant Representations*, to appear in Contemporary Mathematics 2015.

[8] A.S. Bandeira, J. Cahill, D.G. Mixon, A.A. Nelson, *Saving phase: Injectivity and stability for phase retrieval*, Appl. Comp. Harm. Anal. **37** (1) (2014), 106–125.

[9] Y. Benyamini, J. Lindenstrauss, **Geometric Nonlinear Functional Analysis**, vol. 1, AMS Colloquium Publications, vol. 48, 2000.

[10] R. Bhatia, **Matrix Analysis**, Graduate Texts in Mathematics 169, Springer-Verlag 1997.

[11] E. Candés, T. Strohmer, V. Voroninski, *PhaseLift: Exact and Stable Signal Recovery from Magnitude Measurements via Convex Programming*, Communications in Pure and Applied Mathematics **66** (8) (2013), 1241–1274.

[12] C. Davis, W.M. Kahan, *Some new bounds on perturbation of subspaces*, Bull. Amer. Math. Soc. **75** (4) (1969), 863–868.

[13] Y. C. Eldar, S. Mendelson, *Phase retrieval: Stability and recovery guarantees*, Appl. Comp. Harm. Anal. **36** (3) (2014), 473-494.

[14] J.H. Wells, L.R. Williams, **Embeddings and Extensions in Analysis**, Ergebnisse der Mathematik und ihrer Grenzgebiete Band 84, Springer-Verlag 1975.

[15] L. Zwald, G. Blanchard, *On the convergence of eigenspaces in kernel Principal Component Analysis*, Proc. NIPS 05, vol. 18, 1649–1656, MIT Press, 2006.