Frames and Phaseless Reconstruction

Radu Balan

ABSTRACT. Frame design for phaseless reconstruction is now part of the broader problem of nonlinear reconstruction and is an emerging topic in harmonic analysis. The problem of phaseless reconstruction can be simply stated as follows. Given the magnitudes of the coefficients generated by a linear redundant system (frame), we want to reconstruct the unknown input. This problem first occurred in X-ray crystallography starting in the early 20th century. The same nonlinear reconstruction problem shows up in speech processing, particularly in speech recognition.

In this lecture we shall cover existing analysis results as well as stability bounds for signal recovery including: necessary and sufficient conditions for injectivity, Lipschitz bounds of the nonlinear map and its left inverses, stochastic performance bounds, and algorithms for signal recovery.

1. Introduction

These lecture notes concern the problem of finite dimensional vector reconstruction from magnitudes of frame coefficients.

Variants of this problem appear in several areas of engineering and science. In particular in X-ray crystallography one measures the magnitudes of the Fourier transform of the electron density from which one infers the atomic structure of the crystal [**Fin82**]. In speech processing, automatic speech recognition engines typically use cepstral coefficients, which are absolute values of linear combinations of the short time Fourier transform coefficients [**HLO80, Ba10**]. The phaseless reconstruction problem unifies these and other similar problems [**BCE06**].

While the problem can be stated in the more general context of infinite dimensional Hilbert spaces, in these lectures we focus exclusively on the finite dimensional case. In this case any spanning set is a frame (see [**Ca00**] for a complete definition and list of properties). Specifically let $H = \mathbb{C}^n$ denote the n dimensional complex Hilbert space and let $\mathcal{F} = \{f_1, \ldots, f_m\}$ be a set of $m \geq n$ vectors that span H. Fix a real linear space V, that is also a subset of H, $V \subset H$. Our problem is to study when a vector $x \in V$ can be reconstructed from magnitudes of its frame coefficients $\{|\langle x, f_k \rangle|, 1 \leq k \leq m\}$ and how to do so efficiently. This setup covers both the real case and the complex case as studied before in literature: in the real case

$\mathcal{F} \subset V = \mathbb{R}^n$; in the complex case $V = H = \mathbb{C}^n$. Note we assume V is a real linear space which may not be closed under multiplication with complex scalars. While the analysis both of the real case and complex case is presented in a unified way, the reader should be aware that the two cases are not equally easy. In particular the geometric criterion (4) of Theorem 3.2 does not have a counterpart in the complex case.

Consider the following additional notation. Let

$$(1.1) \qquad\qquad T : H \to \mathbb{C}^m \ , \ (T(x))_k = \langle x, f_k \rangle \ , \ 1 \le k \le m$$

denote the frame *analysis map*. Its adjoint is called the *synthesis map* and is defined by

$$(1.2) \qquad\qquad T^* : \mathbb{C}^m \to H \ , \ T^*(c) = \sum_{k=1}^{m} c_k f_k.$$

We now define the main nonlinear function we discuss in this paper

$$x \mapsto (|\langle x, f_k \rangle|)_{1 \le k \le m}.$$

For two vectors $x, y \in H$, consider the equivalence relation $x \sim y$ if and only if there is a constant c of magnitude 1 so that $x = cy$. Thus $x \sim y$ if and only if $x = e^{i\varphi} y$ for some real φ. Let $\hat{H} = H/\sim$ denote the quotient space. Note the nonlinear map is well defined on $\hat{H}$ since $|\langle cx, f_k \rangle| = |\langle x, f_k \rangle|$ for all scalars c with $|c| = 1$. We let α denote the quotient map

$$(1.3) \qquad\qquad \alpha : \hat{H} \to \mathbb{R}^m \ , \ (\alpha(x))_k = |\langle x, f_k \rangle| \ , \ 1 \le k \le m.$$

For purposes that will become clear later, let us also define the map

$$(1.4) \qquad\qquad \beta : \hat{H} \to \mathbb{R}^m \ , \ (\beta(x))_k = |\langle x, f_k \rangle|^2 \ , \ 1 \le k \le m.$$

For the subspace V denote by $\hat{V}$ the set of equivalence classes $\hat{V} = \{\hat{x} \ , \ x \in V\}$.

DEFINITION 1.1. The frame $\mathcal{F}$ is called a *phase retrievable frame with respect to a set V* if the restriction $\alpha|_{\hat{V}}$ is injective.

In these lecture notes we study the following problems:
  (1) Find necessary and sufficient conditions for $\alpha|_{\hat{V}}$ to be a one-to-one (injective) map;
  (2) Study Lipschitz properties of the maps α, β and their inverses;
  (3) Study robustness guarantees (such as Cramer-Rao Lower Bounds) for any inversion algorithm;
  (4) Recovery using convex algorithms (Linear Tensor recovery, and PhaseLift);
  (5) Recovery using iterative algorithms (Gerchberg-Saxton, Wirtinger flow, regularized least-squares).

2. Geometry of $\hat{H}$ and $\mathcal{S}^{p,q}$ Spaces

2.1. $\hat{H}$. Recall $\hat{H} = \widehat{\mathbb{C}^n} = \mathbb{C}^n/\sim = \mathbb{C}^n/T^1$ where $T^1 = \{z \in \mathbb{C} \ , \ |z| = 1\}$. Algebraically $\widehat{\mathbb{C}^n}$ is a homogeneous space invariant to multiplications by positive real scalars. In particular any $x \in \widehat{\mathbb{C}^n} \setminus \{0\}$ has a unique decomposition $x = rp$, where $r = \|x\| > 0$ and $p \in \mathbb{CP}^{n-1}$ is in the projective space $\mathbb{CP}^{n-1} = \mathbb{P}(\mathbb{C}^n)$. Thus, topologically,

$$\widehat{\mathbb{C}^n} = \{0\} \cup \left((0, \infty) \times \mathbb{CP}^{n-1}\right).$$

The subset

$$\mathring{\widehat{\mathbb{C}^n}} = \widehat{\mathbb{C}^n} \setminus \{0\} = (0, \infty) \times \mathbb{CP}^{n-1}$$

is a real analytic manifold.

Now consider the set $\hat{V}$ of equivalence classes associated to vectors in V. Similar to $\hat{H}$, $\hat{V}$ admits the following decomposition

$$\hat{V} = \{0\} \cup ((0, \infty) \times \mathbb{P}(V)),$$

where $\mathbb{P}(V) = \{ \{zx \,,\, z \in \mathbb{C}\} \,,\, x \in V, x \neq 0 \}$ denotes the *projective space associated to V*. The interior subset

$$\mathring{\hat{V}} = \hat{V} \setminus \{0\} = (0, \infty) \times \mathbb{P}(V)$$

is a real analytic manifold of (real) dimension $1 + \dim_{\mathbb{R}} \mathbb{P}(V)$.

Two important cases are as follows:

- Real case. $V = \mathbb{R}^n$ embedded as $x \in \mathbb{R}^n \mapsto x + i0 \in \mathbb{C}^n = H$. Then two vectors $x, y \in V$ are $\sim$ equivalent if and only if $x = y$ or $x = -y$. Similarly, the projective space $\mathbb{P}(V)$ is diffeomorphically equivalent to the real projective space $\mathbb{RP}^{n-1}$ which is of (real) dimension $n - 1$. Thus

$$\dim_{\mathbb{R}}(\mathring{\hat{V}}) = n.$$

- Complex case. $V = \mathbb{C}^n$ which has real dimension $2n$. Then the projective space $\mathbb{P}(V) = \mathbb{CP}^{n-1}$ has real dimension $2n - 2$ (it is also a Khäler manifold) and thus

$$\dim_{\mathbb{R}}(\mathring{\hat{V}}) = 2n - 1.$$

The significance of the real dimension of $\mathring{\hat{V}}$ is encoded in the following result:

THEOREM 2.1 ([**BCE06**]). *If $m \geq 1 + \dim_{\mathbb{R}}(\mathring{\hat{V}})$ then for a (Zariski) generic frame $\mathcal{F}$ of m elements, the set of vectors $x \in V$ such that $\alpha^{-1}(\alpha(\hat{x}))$ has one point in $\hat{V}$ has dense interior in V.*

The real case of this result is contained in Theorem 2.9, whereas the complex case is contained in Theorem 3.4. Both can be found in [**BCE06**].

2.2. $\mathcal{S}^{p,q}$. Consider now $\mathrm{Sym}(H) = \{T : \mathbb{C}^n \to \mathbb{C}^n \,,\, T = T^*\}$, the real vector space of self-adjoint operators over $H = \mathbb{C}^n$ endowed with the Hilbert-Schmidt scalar product $\langle T, S \rangle_{HS} = \mathrm{trace}(TS)$. We also use the notation $\mathrm{Sym}(W)$ for the real vector space of symmetric operators over a (real or complex) vector space W. In both cases self-adjoint means the operator T satisfies $\langle Tx, y \rangle = \langle x, Ty \rangle$ for every x, y in the underlying vector space W. T^* means the adjoint operator of T, and therefore the transpose conjugate of T, when T is a matrix. When T is a an operator acting on a real vector space, T^T denotes its adjoint. For two vectors $x, y \in \mathbb{C}^n$ we denote

$$(2.1) \qquad [\![x, y]\!] = \frac{1}{2}(xy^* + yx^*) \in \mathrm{Sym}(\mathbb{C}^n),$$

their symmetric outer product. On $\mathrm{Sym}(H)$ and $B(H) = \mathbb{C}^{n \times n}$ we consider the class of p-norms defined by the p-norm of the vector of singular values:

$$(2.2) \qquad \|T\|_p = \begin{cases} \max_{1 \leq k \leq n} \sigma_k(T) & for \quad p = \infty \\ (\sum_{k=1}^n \sigma_k^p)^{1/p} & for \quad 1 \leq p < \infty \end{cases},$$

where $\sigma_k = \sqrt{\lambda_k(T^*T)}$, $1 \le k \le n$, are the singular values of T, with $\lambda_k(S)$, $1 \le k \le n$, denoting the eigenvalues of S.

Fix two integers $p, q \ge 0$ and set

$$
\begin{aligned}
\mathcal{S}^{p,q}(H) \quad = \quad & \{T \in \mathrm{Sym}(H) \ , \ T \text{ has at most } p \text{ strictly positive eigenvalues} \\
& \text{and at most } q \text{ strictly negative eigenvalues}\},
\end{aligned}
\tag{2.3}
$$

$$
\begin{aligned}
\mathring{\mathcal{S}}^{p,q}(H) \quad = \quad & \{T \in \mathrm{Sym}(H) \ , \ T \text{ has exactly } p \text{ strictly positive eigenvalues} \\
& \text{and exactly } q \text{ strictly negative eigenvalues}\}.
\end{aligned}
\tag{2.4}
$$

For instance $\mathring{\mathcal{S}}^{0,0}(H) = \mathcal{S}^{0,0}(H) = \{0\}$ and $\mathring{\mathcal{S}}^{1,0}(H)$ is the set of all non-negative operators of rank exactly one. When there is no confusion we shall drop the underlying vector space $H = \mathbb{C}^n$ from notation.

The following basic properties can be found in [**Ba13**], Lemma 3.6; in fact, the last statement is a special instance of the Witt's decomposition theorem.

LEMMA 2.2.
  (1) *For any $p_1 \le p_2$ and $q_1 \le q_2$, $\mathcal{S}^{p_1,q_1} \subset \mathcal{S}^{p_2,q_2}$;*
  (2) *For any nonnegative integers p, q the following disjoint decomposition holds true*

$$
\mathcal{S}^{p,q} = \cup_{r=0}^p \cup_{s=0}^q \mathring{\mathcal{S}}^{r,s},
\tag{2.5}
$$

  where by convention $\mathring{\mathcal{S}}^{p,q} = \emptyset$ for $p + q > n$.
  (3) *For any $p, q \ge 0$,*

$$
-\mathcal{S}^{p,q} = \mathcal{S}^{q,p}.
\tag{2.6}
$$

  (4) *For any linear operator $T : H \to H$ (symmetric or not, invertible or not) and nonnegative integers p, q,*

$$
T^* \mathcal{S}^{p,q} T \subset \mathcal{S}^{p,q}.
\tag{2.7}
$$

  However if T is invertible then $T^* \mathcal{S}^{p,q} T = \mathcal{S}^{p,q}$.
  (5) *For any nonnegative integers p, q, r, s,*

$$
\mathcal{S}^{p,q} + \mathcal{S}^{r,s} = \mathcal{S}^{p,q} - \mathcal{S}^{s,r} = \mathcal{S}^{p+r,q+s}.
\tag{2.8}
$$

The spaces $\mathcal{S}^{1,0}$ and $\mathcal{S}^{1,1}$ play a special role in the following section. Next we summarize their properties (see Lemmas 3.7 and 3.9 in [**Ba13**], and the comment after Lemma 9 in [**BCMN13**]).

LEMMA 2.3 (Space $\mathcal{S}^{1,0}$). *The following statements hold true:*
  (1) $\mathring{\mathcal{S}}^{1,0} = \{xx^* \ , \ x \in H, x \ne 0\}$;
  (2) $\mathcal{S}^{1,0} = \{xx^* \ , \ x \in H\} = \{0\} \cup \{xx^* \ , \ x \in H, x \ne 0\}$;
  (3) *The set $\mathring{\mathcal{S}}^{1,0}$ is a real analytic manifold in $\mathrm{Sym}(\mathbb{C}^n)$ of real dimension $2n - 1$. As a real manifold, its tangent space at $X = xx^*$ is given by*

$$
T_X \mathring{\mathcal{S}}^{1,0} = \left\{ [\![x, y]\!] = \frac{1}{2}(xy^* + yx^*) \ , \ y \in \mathbb{C}^n \right\}.
\tag{2.9}
$$

  The $\mathbb{R}$-linear embedding $\mathbb{C}^n \mapsto T_X \mathring{\mathcal{S}}^{1,0}$ given by $y \mapsto [\![x, y]\!]$ has null space $\{iax \ , \ a \in \mathbb{R}\}$.

LEMMA 2.4 (Space $\mathcal{S}^{1,1}$). *The following statements hold true:*
  (1) $\mathcal{S}^{1,1} = \mathcal{S}^{1,0} - \mathcal{S}^{1,0} = \mathcal{S}^{1,0} + \mathcal{S}^{0,1} = \{[\![x, y]\!] \ , \ x, y \in H\}$;

(2) *For any vectors $x, y, u, v \in H$,*

$$(2.10) \qquad xx^* - yy^* \quad = \quad [\![x+y, x-y]\!] = [\![x-y, x+y]\!],$$

$$(2.11) \qquad [\![u, v]\!] \quad = \quad \frac{1}{4}(u+v)(u+v)^* - \frac{1}{4}(u-v)(u-v)^*.$$

Additionally, for any $T \in \mathcal{S}^{1,1}$ let $T = a_1 e_1 e_1^* - a_2 e_2 e_2^*$ be its spectral factorization with $a_1, a_2 \geq 0$ and $\langle e_i, e_j \rangle = \delta_{i,j}$. Then

$$T = [\![\sqrt{a_1}e_1 + \sqrt{a_2}e_2, \sqrt{a_1}e_1 - \sqrt{a_2}e_2]\!].$$

(3) *The set $\mathring{\mathcal{S}}^{1,1}$ is a real analytic manifold in $\mathrm{Sym}(\mathbb{C}^n)$ of real dimension $4n - 4$. Its tangent space at $X = [\![x, y]\!]$ is given by*

$$(2.12) \qquad T_X \mathring{\mathcal{S}}^{1,1} = \{[\![x, u]\!] + [\![y, v]\!] = \frac{1}{2}(xu^* + ux^* + yv^* + vy^*) \ , \ u, v \in \mathbb{C}^n\}.$$

The $\mathbb{R}$-linear embedding $\mathbb{C}^n \times \mathbb{C}^n \mapsto T_X \mathring{\mathcal{S}}^{1,1}$ given by $(u, v) \mapsto [\![x, u]\!] + [\![y, v]\!]$ has null space $\{a(ix, 0) + b(0, iy) + c(y, -x) + d(iy, ix) \ , \ a, b, c, d \in \mathbb{R}\}$.

(4) *Let $T = [\![u, v]\!] \in \mathcal{S}^{1,1}$. Then its eigenvalues and p-norms are:*

$$(2.13) \qquad a_+ \quad = \quad \frac{1}{2}\left(\mathrm{real}(\langle u, v \rangle) + \sqrt{\|u\|^2\|v\|^2 - (\mathrm{imag}(\langle u, v \rangle))^2}\right) \geq 0,$$

$$(2.14) \qquad a_- \quad = \quad \frac{1}{2}\left(\mathrm{real}(\langle u, v \rangle) - \sqrt{\|u\|^2\|v\|^2 - (\mathrm{imag}(\langle u, v \rangle))^2}\right) \leq 0,$$

$$(2.15) \quad \|T\|_1 \quad = \quad \sqrt{\|u\|^2\|v\|^2 - (\mathrm{imag}(\langle u, v \rangle))^2},$$

$$(2.16) \quad \|T\|_2 \quad = \quad \sqrt{\frac{1}{2}\left(\|u\|^2\|v\|^2 + (\mathrm{real}(\langle u, v \rangle))^2 - (\mathrm{imag}(\langle u, v \rangle))^2\right)},$$

$$(2.17) \quad \|T\|_\infty \quad = \quad \frac{1}{2}\left(|\mathrm{real}(\langle u, v \rangle)| + \sqrt{\|u\|^2\|v\|^2 - (\mathrm{imag}(\langle u, v \rangle))^2}\right).$$

(5) *Let $T = xx^* - yy^* \in \mathcal{S}^{1,1}$. Then its eigenvalues and p-norms are:*

$$(2.18) \qquad a_+ \quad = \quad \frac{1}{2}\left(\|x\|^2 - \|y\|^2 + \sqrt{(\|x\|^2 + \|y\|^2)^2 - 4|\langle x, y \rangle|^2}\right),$$

$$(2.19) \qquad a_- \quad = \quad \frac{1}{2}\left(\|x\|^2 - \|y\|^2 - \sqrt{(\|x\|^2 + \|y\|^2)^2 - 4|\langle x, y \rangle|^2}\right),$$

$$(2.20) \quad \|T\|_1 \quad = \quad \sqrt{(\|x\|^2 + \|y\|^2)^2 - 4|\langle x, y \rangle|^2},$$

$$(2.21) \quad \|T\|_2 \quad = \quad \sqrt{\|x\|^4 + \|y\|^4 - 2|\langle x, y \rangle|^2},$$

$$(2.22) \quad \|T\|_\infty \quad = \quad \frac{1}{2}\left(|\|x\|^2 - \|y\|^2| + \sqrt{(\|x\|^2 + \|y\|^2)^2 - 4|\langle x, y \rangle|^2}\right).$$

Note the above results hold true for the case of symmetric operators over the real subspace V. In particular the factorization at Lemma 2.4(1) implies that

$$(2.23) \quad \mathcal{S}^{1,1}(V) = \mathcal{S}^{1,0}(V) - \mathcal{S}^{1,0}(V) = \mathcal{S}^{1,0}(V) + \mathcal{S}^{0,1}(V) = \{[\![u, v]\!] \ , \ u, v \in V\}.$$

More generally this result holds for subsets $V \subset H$ that are closed under addition and subtraction (such as modules over $\mathbb{Z}$).

2.3. Metrics. The space $\hat{H} = \widehat{\mathbb{C}^n}$ admits two classes of distances (metrics). The first class is the "natural metric" induced by the quotient space structure. The second metric is a matrix norm-induced distance.

Fix $1 \le p \le \infty$.

The *natural metric* denoted by $D_p : \hat{H} \times \hat{H} \to \mathbb{R}$ is defined by

$$(2.24) \qquad D_p(\hat{x}, \hat{y}) = \min_{\varphi \in [0,2\pi)} \|x - e^{i\varphi} y\|_p,$$

where $x \in \hat{x}$ and $y \in \hat{y}$. In the case $p = 2$ the distance becomes

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

By abuse of notation we use also $D_p(x, y) = D_p(\hat{x}, \hat{y})$ since the distance does not depend on the choice of representatives.

The *matrix norm-induced distance* denoted by $d_p : \hat{H} \times \hat{H} \to \mathbb{R}$ is defined by

$$(2.25) \qquad d_p(\hat{x}, \hat{y}) = \|xx^* - yy^*\|_p,$$

where again $x \in \hat{x}$ and $y \in \hat{y}$. In the case $p = 2$ we obtain

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

By abuse of notation we use also $d_p(x, y) = d_p(\hat{x}, \hat{y})$ since again the distance does not depend on the choice of representatives.

As analyzed in [**BZ14**], Proposition 2.4, D_p is not Lipschitz equivalent to d_p, however D_p is an equivalent distance to D_q and similarily, d_p is equivalent to d_q, for any $1 \le p, q \le q$ (see also [**BZ15b**] for the last claim below):

LEMMA 2.5.

(1) *For each $1 \le p \le \infty$, D_p and d_p are distances (metrics) on $\hat{H}$;*

(2) *$(D_p)_{1 \le p \le \infty}$ are equivalent distances; that is, each D_p induces the same topology on $\hat{H}$ and, for every $1 \le p, q \le \infty$, the identity map $i : (\hat{H}, D_p) \to (\hat{H}, D_q)$, $i(x) = x$, is Lipschitz continuous with Lipschitz constant*

$$\mathrm{Lip}^D_{p,q,n} = \max(1, n^{\frac{1}{q} - \frac{1}{p}}).$$

(3) *$(d_p)_{1 \le p \le \infty}$ are equivalent distances, that is, each d_p induces the same topology on $\hat{H}$ and, for every $1 \le p, q \le \infty$, the identity map $i : (\hat{H}, d_p) \to (\hat{H}, d_q)$, $i(x) = x$, is Lipschitz continuous with Lipschitz constant*

$$\mathrm{Lip}^d_{p,q,n} = \max(1, 2^{\frac{1}{q} - \frac{1}{p}}).$$

(4) *The identity map $i : (\hat{H}, D_p) \to (\hat{H}, d_p)$, $i(x) = x$ is continuous, but it is not Lipschitz continuous. The identity map $i : (\hat{H}, d_p) \to (\hat{H}, D_p)$, $i(x) = x$ is continuous but it is not Lipschitz continuous. Hence the induced topologies on $(\hat{H}, D_p)$ and $(\hat{H}, d_p)$ are the same, but the corresponding distances are not Lipschitz equivalent.*

(5) *The metric space $(\hat{H}, d_p)$ is isometrically isomorphic to $\mathcal{S}^{1,0}$ endowed with the p-norm. The isomorphism is given by the map*

$$\kappa_\beta : \hat{H} \to \mathcal{S}^{1,0} \ , \ x \mapsto [\![x, x]\!] = xx^*.$$

(6) *The metric space $(\hat{H}, D_2)$ is Lipschitz isomorphic (not isometric) with $\mathcal{S}^{1,0}$ endowed with the 2-norm. The bi-Lipschitz map*

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

has lower Lipschitz constant 1 and upper Lipschitz constant $\sqrt{2}$.

Note the Lipschitz constant $\operatorname{Lip}_{p,q,n}^D$ is equal to the operator norm of the identity map between $(\mathbb{C}^n, \|\cdot\|_p)$ and $(\mathbb{C}^n, \|\cdot\|_q)$: $\operatorname{Lip}_{p,q,n}^D = \|I\|_{l^p(\mathbb{C}^n) \to l^q(\mathbb{C}^n)}$. Note also the equality $\operatorname{Lip}_{p,q,n}^d = \operatorname{Lip}_{p,q,2}^D$. A consequence of the last two claims in the above result is that while the identity map between $(\hat{H}, D_p)$ and $(\hat{H}, d_q)$ is not bi-Lipschitz, the map $x \mapsto \frac{1}{\sqrt{\|x\|}} x$ is bi-Lipschitz.

3. The Injectivity Problem

In this section we summarize existing results on the injectivity of the maps α and β. Our plan is to present the real and the complex case in a unified way.

Recall that V is a real vector space which is also a subset of $H = \mathbb{C}^n$. The special two cases are $V = \mathbb{R}^n$ (the real case) and $V = \mathbb{C}^n$ (the complex case).

First we describe the *realification procedure* of H and V. Consider the $\mathbb{R}$-linear map $\jmath : \mathbb{C}^n \to \mathbb{R}^{2n}$ defined by

$$\jmath(x) = \begin{bmatrix} \operatorname{real}(x) \\ \operatorname{imag}(x) \end{bmatrix}.$$

Let $\mathcal{V} = \jmath(V)$ be the embedding of V into $\mathbb{R}^{2n}$, and let Π denote the orthogonal projection (with respect to the real scalar product on $\mathbb{R}^{2n}$) onto $\mathcal{V}$. Let J denote the folowing orthogonal antisymmetric $2n \times 2n$ matrix

$$(3.1) \qquad J = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix},$$

where I_n denotes the $n \times n$ identity matrix. Note that $J^T = -J$, $J^2 = -I_{2n}$ and $J^{-1} = -J$.

Each vector f_k of the frame set $\mathcal{F} = \{f_1, \ldots, f_m\}$ is mapped by $\jmath$ onto a vector in $\mathbb{R}^{2n}$ denoted by φ_k, and a symmetric operator in $\mathcal{S}^{2,0}(\mathbb{R}^{2n})$ denoted by Φ_k:

$$(3.2) \qquad \varphi_k = \jmath(f_k) = \begin{bmatrix} \operatorname{real}(f_k) \\ \operatorname{imag}(f_k) \end{bmatrix} \ , \ \Phi_k = \varphi_k \varphi_k^T + J\varphi_k \varphi_k^T J^T.$$

Note that when $f_k \neq 0$ the symmetric form Φ_k has rank 2 and belongs to $\mathring{\mathcal{S}}^{2,0}$. Its spectrum has two distinct eigenvalues: $\|\varphi_k\|^2 = \|f_k\|^2$ with multiplicity 2, and 0 with multiplicity $2n - 2$. Furthermore, $\frac{1}{\|\varphi_k\|^2} \Phi_k$ is a rank 2 projection.

Let $\xi = \jmath(x)$ and $\eta = \jmath(y)$ denote the realifications of vectors $x, y \in \mathbb{C}^n$. Then a bit of algebra shows that

$$(3.3) \qquad \langle x, f_k \rangle \ = \ \langle \xi, \varphi_k \rangle + i \langle \xi, J\varphi_k \rangle,$$

$$\langle F_k, xx^* \rangle_{HS} = \operatorname{trace}(F_k xx^*) = |\langle x, f_k \rangle|^2 = \langle \Phi_k \xi, \xi \rangle \ = \ \langle \Phi_k, \xi \xi^T \rangle_{HS},$$

$$\langle F_k, [\![x, y]\!] \rangle_{HS} = \operatorname{trace}(F_k [\![x, y]\!]) = \operatorname{real}(\langle x, f_k \rangle \langle f_k, y \rangle) \ = \ \langle \Phi_k \xi, \eta \rangle = \operatorname{trace}(\Phi_k [\![\xi, \eta]\!])$$

$$= \ \langle \Phi_k, [\![\xi, \eta]\!] \rangle_{HS},$$

where $F_k = [\![f_k, f_k]\!] = f_k f_k^* \in \mathcal{S}^{1,0}(H)$.

The following objects play an important role in the subsequent theory:

$$(3.4) \qquad R : \mathbb{C}^n \to \mathrm{Sym}(\mathbb{C}^n) \ , \ R(x) = \sum_{k=1}^{m} |\langle x, f_k \rangle|^2 f_k f_k^* \ , \ x \in \mathbb{C}^n,$$

$$(3.5) \qquad \mathcal{R} : \mathbb{R}^{2n} \to \mathrm{Sym}(\mathbb{R}^{2n}) \ , \ \mathcal{R}(\xi) = \sum_{k=1}^{m} \Phi_k \xi \xi^T \Phi_k \ , \ \xi \in \mathbb{R}^{2n},$$

$$(3.6) \qquad \mathcal{S} : \mathbb{R}^{2n} \to \mathrm{Sym}(\mathbb{R}^{2n}) \ , \ \mathcal{S}(\xi) = \sum_{k : \Phi_k \xi \neq 0} \frac{1}{\langle \Phi_k \xi, \xi \rangle} \Phi_k \xi \xi^T \Phi_k \ , \ \xi \in \mathbb{R}^{2n},$$

$$(3.7) \qquad \mathcal{Z} : \mathbb{R}^{2n} \to \mathbb{R}^{2n \times m} \ , \ \mathcal{Z}(\xi) = \begin{bmatrix} \Phi_1 \xi & | & \dots & | & \Phi_m \xi \end{bmatrix} \ , \ \xi \in \mathbb{R}^{2n}.$$

Note $\mathcal{R} = \mathcal{Z}\mathcal{Z}^T$.

Following [**BBCE07**] we note that $|\langle x, f_k \rangle|^2$ is the Hilbert-Schmidt scalar product between two rank 1 symmetric forms:

$$|\langle x, f_k \rangle|^2 = \mathrm{trace}\,(F_k X) = \langle F_k, X \rangle_{HS},$$

where $X = xx^*$. Thus the nonlinear map β induces a linear map on the real vector space $\mathrm{Sym}(\mathbb{C}^n)$ of symmetric forms over $\mathbb{C}^n$:

$$(3.8) \qquad \mathbb{A} : \mathrm{Sym}(\mathbb{C}^n) \to \mathbb{R}^m \ , \ (\mathbb{A}(T))_k = \langle T, F_k \rangle_{HS} = \langle T f_k, f_k \rangle \ , \ 1 \le k \le m$$

Similarly it induces a linear map on $\mathrm{Sym}(\mathbb{R}^{2n})$, the space of symmetric forms over $\mathbb{R}^{2n} = \jmath(\mathbb{C}^n)$, that is denoted by $\mathcal{A}$:
$$(3.9)$$
$$\mathcal{A} : \mathrm{Sym}(\mathbb{R}^{2n}) \to \mathbb{R}^m \ , \ (\mathcal{A}(T))_k = \langle T, \Phi_k \rangle_{HS} = \langle T\varphi_k, \varphi_k \rangle + \langle TJ\varphi_k, J\varphi_k \rangle \ , \ 1 \le k \le m.$$

Now we are ready to state a necessary and sufficient condition for injectivity that works in both the real and the complex case:

THEOREM 3.1 ([**HMW11, BCMN13, Ba13**]). *Let $H = \mathbb{C}^n$ and let V be a real vector space that is also a subset of H, $V \subset H$. Denote by $\mathcal{V} = \jmath(V)$ the realification of V. Assume $\mathcal{F}$ is a frame for V. The following statements are equivalent:*

(1) *The frame $\mathcal{F}$ is phase retrievable with respect to V;*
(2) $\ker \mathbb{A} \cap \left(\mathcal{S}^{1,0}(V) - \mathcal{S}^{1,0}(V) \right) = \{0\}$;
(3) $\ker \mathbb{A} \cap \mathcal{S}^{1,1}(V) = \{0\}$;
(4) $\ker \mathbb{A} \cap \left(\mathcal{S}^{2,0}(V) \cup \mathcal{S}^{1,1}(V) \cup \mathcal{S}^{0,2} \right) = \{0\}$;
(5) *There do not exist vectors $u, v \in V$ with $[\![u, v]\!] \neq 0$ so that*

$$\mathrm{real}\,(\langle u, f_k \rangle \langle f_k, v \rangle) = 0 \ , \ \forall 1 \le k \le m;$$

(6) $\ker \mathcal{A} \cap \left(\mathcal{S}^{1,0}(\mathcal{V}) - \mathcal{S}^{1,0}(\mathcal{V}) \right) = \{0\}$;
(7) $\ker \mathcal{A} \cap \mathcal{S}^{1,1}(\mathcal{V}) = \{0\}$;
(8) *There do not exist vectors $\xi, \eta \in \mathcal{V}$, with $[\![\xi, \eta]\!] \neq 0$ so that*

$$\langle \Phi_k \xi, \eta \rangle = 0 \ , \ \forall 1 \le k \le m.$$

Proof.
$(1) \Leftrightarrow (2)$ It is immediate once we notice that any element in the null space of $\mathbb{A}$ of the form $xx^* - yy^*$ implies $\mathbb{A}(xx^*) = \mathbb{A}(yy^*)$ for some $x, y \in V$ with $\hat{x} \neq \hat{y}$.

$(2) \Leftrightarrow (3)$ and $(3) \Leftrightarrow (5)$ are consequences of (2.23).

For (4) first note that $\ker \mathbb{A} \cap \mathcal{S}^{2,0}(V) = \{0\} = \ker \mathbb{A} \cap \mathcal{S}^{0,2}(V)$ since $\mathcal{F}$ is a frame for V. Thus $(3) \Leftrightarrow (4)$.

(6),(7) and (8) are simply restatements of (2),(3) and (4) using the realification procedure. □

In the case (4) above, note that $\mathcal{S}^{2,0}(V) \cup \mathcal{S}^{1,1}(V) \cup \mathcal{S}^{0,2}(V)$ is the set of all rank less than or equal to 2 symmetric operators in $\mathrm{Sym}(V)$ (This statement has been proposed in [**BCMN13**]).

The above general injectivity result is next made more explicit in the cases $V = \mathbb{C}^n$ and $V = \mathbb{R}^n$.

THEOREM 3.2 ([**BCE06, Ba12**]). *(The real case) Assume $\mathcal{F} \subset \mathbb{R}^n$. The following are equivalent:*

(1) *$\mathcal{F}$ is phase retrievable for $V = \mathbb{R}^n$;*
(2) *$R(x)$ is invertible for every $x \in \mathbb{R}^n$, $x \neq 0$;*
(3) *There do not exist vectors $u, v \in \mathbb{R}^n$ with $u \neq 0$ and $v \neq 0$ so that*

$$\langle u, f_k \rangle \langle f_k, v \rangle = 0 \ , \ \forall 1 \leq k \leq m;$$

(4) *For any disjoint partition of the frame set $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$, either $\mathcal{F}_1$ spans $\mathbb{R}^n$ or $\mathcal{F}_2$ spans $\mathbb{R}^n$.*

Recall a set $\mathcal{F} \subset \mathbb{C}^n$ is called *full spark* if any subset of n vectors is linearly independent. Then an immediate corollary of the above result is the following

COROLLARY 3.3 ([**BCE06**]). *Assume $\mathcal{F} \subset \mathbb{R}^n$. Then*

(1) *If $\mathcal{F}$ is phase retrievable for $\mathbb{R}^n$ then $m \geq 2n - 1$;*
(2) *If $m = 2n - 1$, then $\mathcal{F}$ is phase retrievable if and only if $\mathcal{F}$ is full spark;*

Proof.

Indeed, the first claim follows from Theorem 3.2(4): If $m \leq 2n - 2$ then there is a partition of $\mathcal{F}$ into two subsets each of cardinal less than or equal to $n - 1$. Thus neither set can span $\mathbb{R}^n$. Contradiction.

The second claim is immediate from the same statement as above. □

A more careful analysis of Theorem 3.2(4) gives a recipe of constructing two non-similar vectors $x, y \in \mathbb{R}^n$ so that $\alpha(x) = \alpha(y)$. Indeed, if $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$ so that $\dim \mathrm{span}(\mathcal{F}_1) < n$ and $\dim \mathrm{span}(\mathcal{F}_2) < n$ then there are non-zero vectors $u, v \in \mathbb{R}^n$ with $\langle u, f_k \rangle = 0$ for all $k \in I$, and $\langle v, f_k \rangle = 0$ for all $k \in I^c$. Here I is the index set of frame vectors in $\mathcal{F}_1$ and I^c denotes its complement in $\{1, \ldots, m\}$. Set $x = u + v$ and $y = u - v$. Then $|\langle x, f_k \rangle| = |\langle v, f_k \rangle| = |\langle y, f_k \rangle|$ for all $k \in I$, and $|\langle x, f_k \rangle| = |\langle u, f_k \rangle| = |\langle y, f_k \rangle|$ for all $k \in I^c$. Thus $\alpha(x) = \alpha(y)$, but $x \neq y$ and $x \neq -y$.

THEOREM 3.4 ([**BCMN13, Ba13**]). *(The complex case) The following are equivalent:*

(1) *$\mathcal{F}$ is phase retrievable for $H = \mathbb{C}^n$;*
(2) *$\mathrm{rank}(\mathcal{Z}(\xi)) = 2n - 1$ for all $\xi \in \mathbb{R}^{2n}$, $\xi \neq 0$;*
(3) *$\dim \ker \mathcal{R}(\xi) = 1$ for all $\xi \in \mathbb{R}^{2n}$, $\xi \neq 0$;*
(4) *There do not exist $\xi, \eta \in \mathbb{R}^{2n}$, $\xi \neq 0$ and $\eta \neq 0$ so that $\langle J\xi, \eta \rangle = 0$ and*

(3.10) $$\langle \Phi_k \xi, \eta \rangle = 0 \ , \ \forall 1 \leq k \leq m.$$

In terms of cardinality, here is what we know:

THEOREM 3.5 ([**Mi67, HMW11, BH13, Ba15, MV13, CEHV13, Viz15**]).

(1) [**HMW11**] *If $\mathcal{F}$ is a phase retrievable frame for $\mathbb{C}^n$ then*

(3.11) $$m \geq 4n - 2 - 2b + \begin{cases} 2 & \text{if } n \text{ odd and } b = 3 \bmod 4 \\ 1 & \text{if } n \text{ odd and } b = 2 \bmod 4 \\ 0 & \text{otherwise} \end{cases} ,$$

where $b = b(n)$ denotes the number of 1's in the binary expansion of $n-1$.

(2) [**BH13**] *For any positive integer n there is a frame with $m = 4n-4$ vectors so that $\mathcal{F}$ is phase retrievable for $\mathbb{C}^n$;*

(3) [**CEHV13**] *If $m \geq 4n-4$ then a (Zariski) generic frame is phase retrievable for $\mathbb{C}^n$;*

(4) [**Ba15**] *The set of phase retrievable frames is open in $\mathbb{C}^n \times \cdots \times \mathbb{C}^n$. In particular phase retrievable property is stable under small perturbations.*

(5) [**CEHV13**] *If $n = 2^k + 1$ and $m \leq 4n - 5$ then $\mathcal{F}$ cannot be phase retrievable for $\mathbb{C}^n$.*

(6) [**Viz15**] *For $n = 4$ there is a frame with $m = 11 < 4n - 4 = 12$ vectors that is phase retrievable for $\mathbb{C}^n$.*

4. Robustness of Reconstruction

In this section we analyze stability bounds for reconstruction. Specifically we analyze two types of margins:

- Deterministic, worst-case type bounds: These bounds are given by lower Lipschitz constants of the forward nonliner analysis maps;
- Stochastic, average type bounds: Cramer-Rao Lower Bounds (CRLB).

4.1. Bi-Lipschitzianity of the Nonlinear Analysis Maps. In Section 2 we introduced two distances on $\hat{H}$. As the following theorem shows, the nonlinear maps α and β are bi-Lipschitz with respect to the corresponding distance:

THEOREM 4.1. [**Ba12, EM12, BCMN13, Ba13, BW13, BZ14, BZ15a, BZ15b**] *Let $\mathcal{F}$ be a phase retrievable frame for V, a real linear space, subset of $H = \mathbb{C}^n$. Then:*

(1) *The nonlinear map $\alpha : (\hat{V}, D_2) \to (\mathbb{R}^m, \|\|_2)$ is bi-Lipschitz. Specifically there are positive constants $0 < A_0 \leq B_0 < \infty$ so that*

(4.1) $$\sqrt{A_0} D_2(x,y) \leq \|\alpha(x) - \alpha(y)\|_2 \leq \sqrt{B_0} D_2(x,y) \ , \ \forall x,y \in V.$$

(2) *The nonlinear map $\beta : (\hat{V}, d_1) \to (\mathbb{R}^m, \|\|_2)$ is bi-Lipschitz. Specifically there are positive constants $0 < a_0 \leq b_0 < \infty$ so that*

(4.2) $$\sqrt{a_0} d_1(x,y) \leq \|\beta(x) - \beta(y)\|_2 \leq \sqrt{b_0} d_1(x,y) \ , \ \forall x,y \in V.$$

The converse is also true: If either (4.1) or (4.2) holds true for all $x,y \in V$ then $\mathcal{F}$ is phase retrievable for V.

The choice of distance D_2 and d_1 in the statement of this theorem is only for reasons of convenience since these specific constants will appear later in the text. Any other distance D_p instead of D_2, and d_q instead of d_1 would work. The Lipschitz constants would be different, of course. This result was first obtained for the real case in [**EM12**] for the map α and in [**Ba12**] for the map β. The complex case for map β was shown independently in [**BCMN13**] and [**Ba13**]. The complex case for the more challenging map α was proved in [**BZ15b**]. The paper [**BW13**]

computes the optimal bound A_0 in the real case. The statement presented here (Theorem 4.1) unifies these two cases.

On the other hand the condition that $\mathcal{F}$ is phase retrievable for V is equivalent to the existence of a lower bound for a family of quadratic forms. We state this condition now:

THEOREM 4.2. *Let $\mathcal{F} \subset H = \mathbb{C}^n$ and let V be a real vector space, subset of H. Denote by $\mathcal{V} = \jmath(V) \subset \mathbb{R}^{2n}$ the realification of V, and let Π denote the projection onto $\mathcal{V}$. Then the following statements are equivalent:*

(1) *$\mathcal{F}$ is phase retrievable for V;*
(2) *There is a constant $a_0 > 0$ so that*

$$(4.3) \qquad \Pi \mathcal{R}(\xi) \Pi \geq a_0 \Pi P_{J\xi}^{\perp} \Pi \ , \ \forall \xi \in \mathcal{V}, \|\xi\| = 1,$$

where $P_{J\xi}^{\perp} = I_{2n} - P_{J\xi} = I_{2n} - J\xi\xi^T J^T$ is the orthogonal projection onto the orthogonal complement to $J\xi$;
(3) *There is $a_0 > 0$ so that for all $\xi, \eta \in \mathbb{R}^{2n}$,*

$$(4.4) \qquad \sum_{k=1}^{m} |\langle \Pi \Phi_k \Pi \xi, \eta \rangle|^2 \geq a_0 \left(\|\Pi\xi\|^2 \|\Pi\eta\|^2 - |\langle J\Pi\xi, \Pi\eta \rangle|^2 \right).$$

Note the same constant a_0 can be chosen in (4.2) and (4.3) and (4.4). This result was shown separately for the real and complex case. Here we state these conditions in a unified way.

Proof.

(1) $\Leftrightarrow$ (2) If $\mathcal{F}$ is a phase retrievable frame for V then, by Theorem 3.1(8), for all vectors $\xi, \eta \in \mathcal{V}$, with $[\![\xi, \eta]\!] \neq 0$ we have $\langle \Phi_k \xi, \eta \rangle \neq 0$, for some $1 \leq k \leq m$. Take $\mu \in \mathbb{R}^{2n}$ and set $\eta = \Pi\mu$. Normalize ξ to $\|\xi\| = 1$. Then

$$\sum_{k=1}^{m} |\langle \Phi_k \xi, \eta \rangle|^2 = \langle \mathcal{R}(\xi)\Pi\mu, \Pi\mu \rangle,$$

and by (2.15),

$$\|[\![\xi, \eta]\!]\|_1^2 = \|\xi\|^2 \|\eta\|^2 - |\langle \xi, J\eta \rangle|^2 = \|\Pi\mu\|^2 - |\langle J\xi, \Pi\mu \rangle|^2 = \langle (I_{2n} - J\xi\xi^T J^T)\Pi\mu, \Pi\mu \rangle.$$

Thus if μ satisfies $[\![\xi, \Pi\mu]\!] = 0$ then it must also satisfy $\Pi\mu = tJ\xi$ for some real t. In this case $\Pi\mu$ lies in the null space of $\mathcal{R}(\xi)$. In particular this proves that the following quotient of quadratic forms

$$\frac{\langle \Pi \mathcal{R}(\xi)\Pi\mu, \mu \rangle}{\langle \Pi(I_{2n} - J\xi\xi^T J^T)\Pi\mu, \mu \rangle}$$

is bounded above and below away from zero. This proves that (4.3) must hold for some $a_0 > 0$. Conversely, if (4.3) holds true, then for every $\xi, \eta \in \mathcal{V}$ with $[\![\xi, \eta]\!] \neq 0$, $\langle P_{J\xi}^{\perp}\eta, \eta \rangle \neq 0$ and thus $\langle \Phi_k \xi, \eta \rangle \neq 0$ for some k. This shows that $\mathcal{F}$ is a phase retrievable frame for V.

(2) $\Leftrightarrow$ (3) This follows by writing out (4.3) explicitly.
$\square$

REMARK 4.3. Condition (2) of this theorem expressed by Equation (4.3) can be used to check if a given frame is phase retrievable as we explain next.

In the real case, $\Pi = I_n \oplus 0$, and this condition reduces to

$$R(x) = \sum_{k=1}^{m} |\langle x, f_k \rangle|^2 f_k f_k^T \geq a_0 \|x\|^2 I_H \ , \ \forall x \in H = \mathbb{R}^n.$$

In turn this is equivalent to any of the conditions of Theorem 3.2.

In the complex case the condition (4.3) turns into

$$(4.5) \qquad \lambda_{2n-1}(\mathcal{R}(\xi)) \geq a_0 \ , \ \forall \xi \in \mathbb{R}^{2n}, \|\xi\| = 1,$$

where $\lambda_{2n-1}(\mathcal{R}(\xi))$ denotes the next to the smallest eigenvalue of $\mathcal{R}(\xi)$. The algorithm requires an upper bound for $b_0 = \max_{\|\xi\|=1} \lambda_1(\mathcal{R}(\xi))$. For instance $b_0 \leq B \max_k \|f_k\|^2$, where B is the frame upper bound [**Ba15**]. The condition (4.5) can be checked using an ε-net of the unit sphere in $\mathbb{R}^{2n}$. Specifically let $\{\xi_j^\varepsilon\}$ be such an ε-net, that is $\|\xi_j^\varepsilon\| = 1$ and $\|\xi_j^\varepsilon - \xi_k^\varepsilon\| < \varepsilon$ for all $j \neq k$. Set $a_0 = \frac{1}{2} \min_j \lambda_{2n-1}(\mathcal{R}(\xi_j^\varepsilon))$. If $2b_0\varepsilon \leq a_0$ then stop, otherwise set $\varepsilon = \frac{1}{2}\varepsilon$ and construct a new ε-net.

The condition $2b_0\varepsilon \leq a_0$ guarantees that for every $\xi \in \mathbb{R}^{2n}$ with $\|\xi\| = 1$, $\lambda_{2n-1}(\mathcal{R}(\xi)) \geq a_0$ since (see also [**Ba15**] for a similar derivation)

$$\|\mathcal{R}(\xi) - \mathcal{R}(\xi_j^\varepsilon)\| \leq \sqrt{b_0^2 \|\xi - \xi_j^\varepsilon\| \|\xi + \xi_j^\varepsilon\|} \leq 2b_0 \|\xi - \xi_j^\varepsilon\| \leq 2b_0\varepsilon$$

and by Weyl's perturbation theorem (see III.2.6 in [**Bh97**])

$$\lambda_{2n-1}(\mathcal{R}(x)) \geq \lambda_{2n-1}(\mathcal{R}(\xi_j^\varepsilon)) - \|\mathcal{R}(\xi) - \mathcal{R}(\xi_j^\varepsilon)\| \geq 2a_0 - 2b\varepsilon \geq a_0.$$

Unfortunately such an approach has at least an NP computational cost since the cardinality of an ε-net is of the order $\left(\frac{1}{\varepsilon}\right)^n$.

The computations of lower bounds is fairly subtle. In fact there is a distinction between local bounds and global bounds. Specifically for every $z \in V$ we define the following bounds:

The *type I local lower Lipschitz bounds* are defined by:

$$(4.6) \qquad A(z) \ = \ \lim_{r \to 0} \inf_{x,y \in V, D_2(x,z)<r, D_2(y,z)<r} \frac{\|\alpha(x) - \alpha(y)\|_2^2}{D_2(x,y)^2},$$

$$(4.7) \qquad a(z) \ = \ \lim_{r \to 0} \inf_{x,y \in V, d_1(x,z)<r, d_1(y,z)<r} \frac{\|\beta(x) - \beta(y)\|_2^2}{d_1(x,y)^2}.$$

The *type II local lower Lipschitz bounds* are defined by:

$$(4.8) \qquad \tilde{A}(z) \ = \ \lim_{r \to 0} \inf_{y \in V, D_2(y,z)<r} \frac{\|\alpha(z) - \alpha(y)\|_2^2}{D_2(z,y)^2},$$

$$(4.9) \qquad \tilde{a}(z) \ = \ \lim_{r \to 0} \inf_{y \in V, d_1(y,z)<r} \frac{\|\beta(z) - \beta(y)\|_2^2}{d_1(z,y)^2}.$$

Similarly the *type I local upper Lipschitz bounds* are defined by:

$$(4.10) \qquad B(z) \ = \ \lim_{r \to 0} \sup_{x,y \in V, D_2(x,z)<r, D_2(y,z)<r} \frac{\|\alpha(x) - \alpha(y)\|_2^2}{D_2(x,y)^2},$$

$$(4.11) \qquad b(z) \ = \ \lim_{r \to 0} \sup_{x,y \in V, d_1(x,z)<r, d_1(y,z)<r} \frac{\|\beta(x) - \beta(y)\|_2^2}{d_1(x,y)^2}$$

and the *type II local upper Lipschitz bounds* are defined by:

$$(4.12) \qquad \tilde{B}(z) \quad = \quad \lim_{r \to 0} \sup_{y \in V, D_2(y,z) < r} \frac{\|\alpha(z) - \alpha(y)\|_2^2}{D_2(z,y)^2},$$

$$(4.13) \qquad \tilde{b}(z) \quad = \quad \lim_{r \to 0} \sup_{y \in V, d_1(y,z) < r} \frac{\|\beta(z) - \beta(y)\|_2^2}{d_1(z,y)^2}.$$

The *global lower bounds* are defined by

$$(4.14) \qquad A_0 \quad = \quad \inf_{x,y \in V, D_2(x,y) > 0} \frac{\|\alpha(x) - \alpha(y)\|_2^2}{D_2(x,y)^2},$$

$$(4.15) \qquad a_0 \quad = \quad \inf_{x,y \in V, d_1(x,y) > 0} \frac{\|\beta(x) - \beta(y)\|_2^2}{d_1(x,y)^2},$$

whereas the *global upper bounds* are defined by

$$(4.16) \qquad B_0 \quad = \quad \sup_{x,y \in V, D_2(x,y) > 0} \frac{\|\alpha(x) - \alpha(y)\|_2^2}{D_2(x,y)^2},$$

$$(4.17) \qquad b_0 \quad = \quad \sup_{x,y \in V, d_1(x,y) > 0} \frac{\|\beta(x) - \beta(y)\|_2^2}{d_1(x,y)^2}$$

and represent the square of the corresponding Lipschitz constants .

Due to homogeneity $A_0 = A(0)$, $B_0 = B(0)$, $a_0 = a(0)$, and $b_0 = b(0)$. On the other hand, for $z \neq 0$, $A(z) = A(\frac{z}{\|z\|})$, $B(z) = B(\frac{z}{\|z\|})$, $a(z) = a(\frac{z}{\|z\|})$, and $b(z) = b(\frac{z}{\|z\|})$. Note that $A(z)$ stands for the local lower Lipschitz bound of type I at z, whereas A denotes the optimal lower frame bound of $\mathcal{F}$.

The exact expressions of these bounds are summarized by the following results. For any $I \subset \{1, 2, \ldots, m\}$ let $\mathcal{F}[I] = \{f_k , k \in I\}$ denote the subset indexed by I. Also let $\sigma_1^2[I]$ and $\sigma_n^2[I]$ denote the upper and the lower frame bound of set $\mathcal{F}[I]$, respectively. Thus:

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

As usual, I^c denotes the complement of the index set I, that is $I^c = \{1, \ldots, m\} \setminus I$.

THEOREM 4.4 ([**BW13, BCMN13**]). *(The real case) Assume $\mathcal{F} \subset \mathbb{R}^n$ is a phase retrievable frame for $\mathbb{R}^n$. Let A and B denote its optimal lower and upper frame bound, respectively. Then:*

(1) *For every $0 \neq x \in \mathbb{R}^n$, $A(x) = \sigma_n^2[\text{supp}(\alpha(x))]$, where $\text{supp}(\alpha(x)) = \{k , \langle x, f_k \rangle \neq 0\}$;*
(2) *For every $x \in \mathbb{R}^n$, $\tilde{A}(x) = A$;*
(3) $A_0 = A(0) = \min_I(\sigma_n^2[I] + \sigma_n^2[I^c])$;
(4) *For every $x \in \mathbb{R}^n$, $B(x) = \tilde{B}(x) = B$;*
(5) $B_0 = B(0) = \tilde{B}(0) = B$, *the optimal upper frame bound;*
(6) *For every $0 \neq x \in \mathbb{R}^n$, $a(x) = \tilde{a}(x) = \lambda_{\min}(R(x))/\|x\|^2$;*
(7) $a_0 = a(0) = \tilde{a}(0) = \min_{\|x\|=1} \lambda_{\min}(R(x))$;
(8) *For every $0 \neq x \in \mathbb{R}^n$, $b(x) = \tilde{b}(x) = \lambda_{\max}(R(x))/\|x\|^2$;*
(9) $b_0 = b(0) = \tilde{b}(0) = \max_{\|x\|=1} \lambda_{\max}(R(x))$;

(10) a_0 is the largest constant so that

$$R(x) \geq a_0 \|x\|^2 I_n \ , \ \forall \, x \in \mathbb{R}^n,$$

or, equivalently,

$$\sum_{k=1}^{m} |\langle x, f_k \rangle|^2 |\langle y, f_k \rangle|^2 \geq a_0 \|x\|^2 \|y\|^2 \ , \ \forall \, x, y \in \mathbb{R}^n;$$

(11) b_0 is the 4th power of the frame analysis operator norm $T : (\mathbb{R}^n, \|\cdot\|_2) \to (\mathbb{R}^m, \|\cdot\|_4)$,

$$b_0 = \|T\|_{B(l^2, l^4)}^4 = \max_{\|x\|_2 = 1} \sum_{k=1}^{m} |\langle x, f_k \rangle|^4.$$

The complex case is subtler. The following result presents some of the local and global Lipschitz bounds.

THEOREM 4.5 ([**BZ15b**]). *(The complex case) Assume $\mathcal{F}$ is phase retrievable for $H = \mathbb{C}^n$ and A, B are its optimal frame bounds. Then:*

(1) *For every $0 \neq z \in \mathbb{C}^n$, $A(z) = \lambda_{2n-1}(\mathcal{S}(\mathsf{J}(z)))$ (the next to the smallest eigenvalue);*

(2) $A_0 = A(0) > 0$;

(3) *For every $z \in \mathbb{C}^n$, $\tilde{A}(z) = \lambda_{2n-1}\left(\mathcal{S}(\mathsf{J}(z)) + \sum_{k:\langle z, f_k \rangle = 0} \Phi_k\right)$ (the next to the smallest eigenvalue);*

(4) $\tilde{A}(0) = A$, *the optimal lower frame bound;*

(5) *For every $z \in \mathbb{C}^n$, $B(z) = \tilde{B}(z) = \lambda_1\left(\mathcal{S}(\mathsf{J}(z)) + \sum_{k:\langle z, f_k \rangle = 0} \Phi_k\right)$ (the largest eigenvalue);*

(6) $B_0 = B(0) = \tilde{B}(0) = B$, *the optimal upper frame bound;*

(7) *For every $0 \neq z \in \mathbb{C}^n$, $a(z) = \tilde{a}(z) = \lambda_{2n-1}(\mathcal{R}(\mathsf{J}(z)))/\|z\|^2$ (the next to the smallest eigenvalue);*

(8) *For every $0 \neq z \in \mathbb{C}^n$, $b(z) = \tilde{b}(z) = \lambda_1(\mathcal{R}(\mathsf{J}(z)))/\|z\|^2$ (the largest eigenvalue);*

(9) a_0 *is the largest constant so that*

$$\mathcal{R}(\xi) \geq a_0 (I - J\xi\xi^T J^T) \ , \ \forall \, \xi \in \mathbb{R}^{2n}, \|\xi\| = 1,$$

or, equivalently,

$$\sum_{k=1}^{m} |\langle \Phi_k \xi, \eta \rangle|^2 \geq a_0 \left(\|\xi\|^2 \|\eta\|^2 - |\langle J\xi, \eta \rangle|^2\right) \ , \ \forall \, \xi, \eta \in \mathbb{R}^{2n};$$

(10) $b(0) = \tilde{b}(0) = b_0$ *is the 4th power of the frame analysis operator norm $T : (\mathbb{C}^n, \|\cdot\|_2) \to (\mathbb{R}^m, \|\cdot\|_4)$,*

$$b_0 = \|T\|_{B(l^2, l^4)}^4 = \max_{\|x\|_2 = 1} \sum_{k=1}^{m} |\langle x, f_k \rangle|^4;$$

(11) $\tilde{a}(0)$ *is given by*

$$\tilde{a}(0) = \min_{\|z\|=1} \sum_{k=1}^{m} |\langle z, f_k \rangle|^4.$$

The results presented so far show that both α and β admit left inverses that are Lipschitz continuous on their domains of definition. One remaining problem is to know whether these left inverses can be extended to Lipschitz maps over the entire $\mathbb{R}^m$. The following two results provide a positive answer (see [**BZ14, BZ15b**] for details).

THEOREM 4.6 ([**BZ15b**]). *Assume $\mathcal{F} \subset H = \mathbb{C}^n$ is a phase retrievable frame for $\mathbb{C}^n$. Let $\sqrt{A_0}$ be the lower Lipschitz constant of the map $\alpha : (\hat{H}, D_2) \rightarrow (\mathbb{R}^m, \|\cdot\|_2)$. Then there is a Lipschitz map $\omega : (\mathbb{R}^m, \|\cdot\|_2) \rightarrow (\hat{H}, D_2)$ so that: (i) $\omega(\alpha(x)) = x$ for all $x \in \hat{H}$, and (ii) its Lipschitz constant is $\mathrm{Lip}(\omega) \leq \frac{4+3\sqrt{2}}{\sqrt{A_0}}$.*

THEOREM 4.7 ([**BZ14, BZ15a**]). *Assume $\mathcal{F} \subset H = \mathbb{C}^n$ is a phase retrievable frame for $\mathbb{C}^n$. Let $\sqrt{a_0}$ be the lower Lipschitz constant of the map $\beta : (\hat{H}, d_1) \rightarrow (\mathbb{R}^m, \|\cdot\|_2)$. Then there is a Lipschitz map $\psi : (\mathbb{R}^m, \|\cdot\|_2) \rightarrow (\hat{H}, d_1)$ so that: (i) $\psi(\beta(x)) = x$ for all $x \in \hat{H}$, and (ii) its Lipschitz constant is $\mathrm{Lip}(\psi) \leq \frac{4+3\sqrt{2}}{\sqrt{a_0}}$.*

Sketch of Proof

Proofs of both results follow a similar strategy. First both metric spaces $(\hat{H}, d_1)$ and $(\hat{H}, D_2)$ are bi-Lipschitz isomorphic with $\mathcal{S}^{1,0}$ via Lemma 2.5. Then one uses Kirszbraun's Theorem (see, e.g., [**BL00, HG13, WW75**]) to obtain an isometric Lipschitz extension of the left inverse of α (or β) from its range to the entire $(\mathbb{R}^m, \|\cdot\|_2)$ into $(\mathrm{Sym}(H), \|\cdot\|_2)$. The final step is to construct a Lipschitz map $\pi : \mathrm{Sym}(H) \rightarrow \mathcal{S}^{1,0}(H)$ so that $\pi(x^*) = xx^*$ for every $x \in H$. This map is realized as $\pi(A) = (\lambda_1 - \lambda_2)P_1$, where $\lambda_1 \geq \lambda_2$ are the two largest eigenvalues of A, and P_1 is the principal eigenprojector. Using the integration contour from [**ZB06**] and Weyl's inequalities (see III.2 in [**Bh97**]) the authors of [**BZ15b**] obtained that π is Lipschitz with $\mathrm{Lip}(\pi) \leq 3 + 2\sqrt{2}$ for $\pi : (\mathrm{Sym}(H), \|\cdot\|_2) \rightarrow (\mathcal{S}^{1,0}, \|\cdot\|_2)$.

4.2. Fisher Information Matrices and Cramer-Rao Lower Bounds. Throughout this section assume $\mathcal{F} = \{f_1, \ldots, f_m\} \subset H = \mathbb{C}^n$ is a phase retrievable frame for V, where $V \subseteq H$ is a real linear space, and $x \in V$.

Consider two measurement processes. The first is the Additive White Gaussian Noise (AWGN) model

$$(4.18) \qquad y_k = |\langle x, f_k \rangle|^2 + \nu_k \ , \ 1 \leq k \leq m,$$

where $(\nu_k)_{1 \leq k \leq m}$ are independent and identically distributed (i.i.d.) realizations of a normal random variable of zero mean and variance σ^2. The second process is a non-Additive White Gaussian Noise (nonAWGN) model where the noise is added prior to taking the absolute value:

$$(4.19) \qquad y_k = |\langle x, f_k \rangle + \mu_k|^2 \ , \ 1 \leq k \leq m,$$

where $(\mu_k)_{1 \leq k \leq m}$ are i.i.d. realizations of a Gaussian complex process with zero mean and variance ρ^2.

First we present the *Fisher Information matrices* $\mathbb{I}$ for these two processes. The general definition of the Fisher Information Matrix is (see [**Ky10**])

$$\mathbb{I}(x) = \mathbb{E}[(\nabla_x \log p(y; x))(\nabla_x \log p(y; x))^T].$$

Following [**BCMN13**] and [**Ba13**] for the AWGN model (4.18) we obtain:

$$(4.20) \qquad \mathbb{I}^{\mathrm{AWGN}}(x) = \frac{4}{\sigma^2} \mathcal{R}(\xi) = \frac{4}{\sigma^2} \sum_{k=1}^m \Phi_k \xi \xi^T \Phi_k,$$

where $\xi = \mathrm{J}(x) \in \mathbb{R}^{2n}$. In general $\mathbb{I}(x)$ has rank at most $2n - 1$ because $J\xi$ is always in its kernel.

In [**Bal15**] the Fisher information matrix for the nonAWGN model (4.19) is shown to have the following form:

$$
\mathbb{I}^{\mathrm{nonAWGN}}(x) = \frac{4}{\rho^4} \sum_{k=1}^{m} \left(G_1 \left(\frac{\langle \Phi_k \xi, \xi \rangle}{\rho^2} \right) - 1 \right) \Phi_k \xi \xi^* \Phi_k
$$

$$
\text{(4.21)} \qquad = \frac{4}{\rho^2} \sum_{k=1}^{m} G_2 \left(\frac{\langle \Phi_k \xi, \xi \rangle}{\rho^2} \right) \frac{1}{\langle \Phi_k \xi, \xi \rangle} \Phi_k \xi \xi^* \Phi_k,
$$

where the two universal scalar functions $G_1, G_2 : \mathbb{R}^+ \to \mathbb{R}^+$ are given by

$$
\text{(4.22)} \quad G_1(a) = \frac{e^{-a}}{a} \int_0^\infty \frac{I_1^2(2\sqrt{at})}{I_0(2\sqrt{at})} t e^{-t} dt = \frac{e^{-a}}{8a^3} \int_0^\infty \frac{I_1^2(t)}{I_0(t)} t^3 e^{-\frac{t^2}{4a}} dt
$$

$$
G_2(a) = a(G_1(a) - 1),
$$

where I_0 and I_1 are the modifed Bessel functions of the first kind and order 0 and 1, respectively. Both Fisher information matrices have the same null space spanned by $J\xi$.

Next we present a lower bound on the variance of any unbiased estimator for x. Let $z_0 \in V$ be a fixed vector. Define

$$
\text{(4.23)} \qquad V_{z_0} = \{ x \in V \ , \ \langle x, z_0 \rangle > 0, \}
$$

where $\langle \cdot, \cdot \rangle$ is the complex scalar product in H. Set $E_{z_0} = \mathrm{span}_{\mathbb{R}}(V_{z_0})$ the real vector space spanned by V_{z_0}. Note $E_{z_0} = \{ x \in V \ , \ \mathrm{imag}(\langle x, z_0 \rangle) = 0 \}$.

To make (4.18) identifiable we select the representative $x \in V_{z_0}$ of the class $\hat{x}$. This is a mild condition since it only asks for the class $\hat{x}$ not to be orthogonal to z_0 with respect to the scalar product of H. An estimator $\omega : \mathbb{R}^m \to E_{z_0}$ is unbiased if $\mathbb{E}[\omega(\beta(x) + \nu)] = x$ for all $x \in V_{z_0}$. Here the expectation is taken with respect to the noise random variable.

A careful analysis (see [**Ba13**]) of the estimation process shows that the *Cramer-Rao Lower Bound (CRLB)* for either measurement process (4.18) and (4.19) is given by $(\Pi_{z_0} \mathbb{I}(x) \Pi_{z_0})^\dagger$, where Π_{z_0} is the orthogonal projection onto $\mathcal{V}_{z_0} = \mathrm{J}(E_{z_0})$ in $\mathbb{R}^{2n}$, and upper script $\dagger$ denotes the Moore-Penrose pseudo-inverse. Here $\mathbb{I}(x)$ stands for the Fisher information matrix $\mathbb{I}^{\mathrm{AWGN}}(x)$ or $\mathbb{I}^{\mathrm{nonAWGN}}(x)$. Then the covariance of any unbiased estimator $\omega : \mathbb{R}^m \to E_{z_0}$ is bounded as follows:

$$
\text{(4.24)} \qquad \mathrm{Cov}[\omega] \geq (\Pi_{z_0} \mathbb{I}(x) \Pi_{z_0})^\dagger .
$$

In particular, if one choses $z_0 = x$ then Π_{z_0} becomes the orthogonal projection onto the range of $\mathbb{I}(x)$ and $(\Pi_{z_0} \mathbb{I}(x) \Pi_{z_0})^\dagger = \mathbb{I}(x)^\dagger$ (see (4.25) below).

In the real case, $\mathcal{F} \subset V = \mathbb{R}^n \subset \mathbb{C}^n$, the Fisher information matrices of the AWGN model (4.18) and of the nonAWGN model (4.19) take the form

$$
\mathbb{I}^{\mathrm{AWGN}}(x) = \frac{4}{\sigma^2} \begin{bmatrix} R(x) & 0 \\ 0 & 0 \end{bmatrix} , \ \mathbb{I}^{\mathrm{nonAWGN}}(x) = \frac{4}{\sigma^2} \sum_{k=1}^{m} G_2(\frac{|\langle x, f_k \rangle|^2}{\rho^2}) \begin{bmatrix} f_k f_k^T & 0 \\ 0 & 0 \end{bmatrix} .
$$

Restricting to the real component of the estimator, the CRLB for the AWGN model (4.18) becomes

$$
\mathrm{Cov}[\omega^{\mathrm{AWGN}}] \geq \frac{\sigma^2}{4} R(x)^{-1}
$$

whereas for the nonAWGN model (4.19), the CRLB becomes

$$\mathrm{Cov}[\omega^{\mathrm{nonAWGN}}] \geq \frac{\rho^2}{4} \left(\sum_{k=1}^{m} G_2(\frac{|\langle x, f_k \rangle|^2}{\rho^2}) f_k f_k^T \right)^{-1}.$$

In the complex case $\mathcal{F} \subset V = H = \mathbb{C}^n$, $\Pi_{z_0} = I_{2n} - J\psi_0\psi_0^T J^T$ with $\psi_0 = \mathrm{J}(z_0)$ and the CRLB for AWGN becomes

$$\mathrm{Cov}[\omega] \geq \frac{\sigma^2}{4}((I_{2n} - J\psi_0\psi_0^T J^T)\mathcal{R}(\xi)(I_{2n} - J\psi_0\psi_0^T J^T))^\dagger.$$

Since $\mathcal{F}$ is phase retrievable for H, by Theorem 4.5(9) $\mathcal{R}(\xi)$ satisfies the lower bound $\mathcal{R}(\xi) \geq a_0(\|\xi\|^2 I_{2n} - J\xi\xi^T J^T)$. A little bit of algebra shows

$$
\begin{aligned}
a_0|\langle x, z_0 \rangle|^2 \Pi_{z_0} &= a_0|\langle \xi, \psi_0 \rangle|^2 \Pi_{z_0} \\
&\leq a_0|\langle \xi, \psi_0 \rangle|^2 (I_{2n} - J\psi_0\psi_0^T J^T) + a_0 \Big[\|\xi - \langle \xi, \psi_0 \rangle \psi_0\|^2 I_{2n} \\
&\quad -J(\xi - \langle \xi, \psi_0 \rangle \psi_0)(\xi - \langle \xi, \psi_0 \rangle \psi_0)^T J^T \Big] \\
&= (I_{2n} - J\psi_0\psi_0^T J^T)\mathcal{R}(\xi)(I_{2n} - J\psi_0\psi_0^T J^T).
\end{aligned}
$$

In particular this inequality shows that, if an unbiased estimator $\omega^0 : \mathbb{R}^m \to E_{z_0}$ for the AWGN model (4.18) achieves the CRLB then its covariance matrix is upper bounded by

$$\mathrm{Cov}[\omega^0] \leq \frac{\sigma^2}{4a_0|\langle x, z_0 \rangle|^2} \Pi_{z_0}.$$

This result was derived in [**Ba13**].

Finally, if the global phase is provided by an oracle by correlating the estimated signal with the original signal x, then we can choose $z_0 = x$ and $\Pi_x = I_{2n} - J\xi\xi^T J^T$. But then

$$\Pi_x \mathbb{I}^{\mathrm{AWGN}}(x)\Pi_x = \mathbb{I}^{\mathrm{AWGN}}(x) , \ \Pi_x \mathbb{I}^{\mathrm{nonAWGN}}(x)\Pi_x = \mathbb{I}^{\mathrm{nonAWGN}}(x)$$

which implies the CRLBs:

$$(4.25) \qquad \mathrm{Cov}[\omega^{\mathrm{AWGN}}] \geq \left(\mathbb{I}^{\mathrm{AWGN}}(x)\right)^\dagger , \ \mathrm{Cov}[\omega^{\mathrm{nonAWGN}}] \geq \left(\mathbb{I}^{\mathrm{nonAWGN}}(x)\right)^\dagger.$$

5. Reconstruction Algorithms

We present two types of reconstruction algorithms:

- Rank 1 tensor recovery: Linear Reconstruction, PhaseLift;
- Iterative algorithms: Gerchberg-Saxton, Mean-Squares Optimization: Wirtinger flow and IRLS.

The literature contains more algorithms than those presented here, see e.g. [**WAM12, ABFM12, Ba10, FMNW13, Fin82**].

Throughout this section we assume $\mathcal{F}$ is a phase retrievable frame for $H = \mathbb{C}^n$. We let $y = (y_k)_{1 \leq k \leq m}$ denote the vector of measurements. We analyze two cases: The noiseless case, when $y = \beta(x)$, and the additive noise case, when $y = \beta(x) + \nu$, where $\nu \in \mathbb{R}^m$ denotes the noise.

5.1. Rank 1 Tensor Recovery. The matrix recovery algorithms attempt to estimate the rank 1 matrix $X = xx^*$ from the measurements $y = (y_k)_{1 \le k \le m}$. We present two such algorithms: the linear reconstruction algorithm using lifting , and PhaseLift. An extension of the linear reconstruction algorithm from a matrix to a higher order tensor setting is also included.

5.1.1. *Linear Reconstruction Using Lifting.*

(i) Order 2 **Tensor Embedding.**

Linear reconstruction works well when the frame has high redundancy. Specifically if $m \ge \dim_{\mathbb{R}}(\mathrm{Sym}(H)) = n^2$ then, generically, the set of rank 1 operators $\{F_k = f_k f_k^* , \ 1 \le k \le m\}$ is a frame for $\mathrm{Sym}(H)$. In this case the measurements are linear on the space of matrices:

$$y_k = \langle F_k, X \rangle_{HS} + \nu_k , \ 1 \le k \le m.$$

Let $\{\tilde{F}_k , \ 1 \le k \le m\}$ denote the canonical dual frame to $\{F_k , \ 1 \le k \le m\}$. Then the minimum Frobenius norm estimate of X is given by the linear formula $X_{\mathrm{est}} = \sum_{k=1}^m y_k \tilde{F}_k$. The class $\hat{x}$ is recovered using the spectral decomposition of X_{est},

$$X_{\mathrm{est}} = \sum_{j=1}^d \lambda_{r(j)} P_j,$$

where d denotes the number of distinct eigenvalues of X_{est}, $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$ are the eigenvalues of X_{est}, P_j is the orthogonal projection onto the eigenspace associated to eigenvalue $\lambda_{r(j)}$, and $r(j+1) = r(j) + \mathrm{rank}(P_j)$, with $r(1) = 1$. The least squares estimator x^{LS} of $\hat{x}$ from X_{est} minimizes $\|X_{\mathrm{est}} - xx^*\|^2$. The solution is unique when the top eigenvalue of X_{est} is simple, $\lambda_1 > \lambda_2$. In this case let $P = ee^*$ for some unit norm vector e. The least squares estimator x^{LS} is given explicitly by

$$(5.1) \qquad\qquad x^{LS} = \left\{ \begin{array}{ll} \sqrt{\lambda_1}e & if \quad \lambda_1 \ge 0 \\ 0 & otherwise \end{array} \right. .$$

In the case $\lambda_1 = \lambda_2 > 0$ there are infinitely many possible top eigenvectors. Unfortunately any such choice produces an estimator that is discontinuous as a function of y. On the other hand the following estimator

$$(5.2) \qquad\qquad x^{\mathrm{Lip}} = \left\{ \begin{array}{ll} \sqrt{\lambda_1 - \lambda_2}e & if \quad \lambda_1 > \lambda_2 \\ 0 & otherwise \end{array} \right.$$

is an exact reconstruction scheme in the absence of noise and it is a Lipschitz continuous map with respect to the measurement vector y. Its Lipschitz constant and its performance with respect to additive noise are described in [**BZ15b**].

(ii) Higher Order Tensor Embeddings.

The idea of lifting into the space $\mathcal{S}^{1,0}$ of rank 1 matrices can be extended to spaces of higher order tensors (see [**Ba09**]). Fix an integer $d \ge 1$ and denote by $O_{n,d} = \{\gamma \in \mathbb{N}^d , \ 1 \le \gamma(1) \le \cdots \le \gamma(d) \le n\}$ the set of ordered d-tuples of positive integers up to n. Let $\mathcal{P}_{d,d}(Z_1, \ldots, Z_n)$ denote the real linear space of symmetric homogeneous polynomials in n variables $Z_1, \ldots, Z_n$ of degree (d, d), meaning that each monomial has degree d in variables $Z_1, \ldots, Z_n$ and degree d in conjugate

variables $\overline{Z_1}, \ldots, \overline{Z_n}$,

$$P = \sum_{\gamma, \delta \in O_{n,d}} c_{\gamma, \delta} Z_{\gamma(1)} \cdots Z_{\gamma(d)} \overline{Z_{\delta(1)} \cdots Z_{\delta(d)}} \ , \ c_{\gamma, \delta} = \overline{c_{\delta, \gamma}} \in \mathbb{C}.$$

In the case $d = 1$, $\mathcal{P}_{1,1}(Z_1, \ldots, Z_n)$ is a linear $\mathbb{R}$-space isomorphic to $\mathrm{Sym}(\mathbb{C}^n)$. In general, $\mathcal{P}_{d,d}(Z_1, \ldots, Z_n)$ is isomorphic to the $\mathbb{R}$-linear space of (d,d)-sesquilinear functionals over $\mathbb{C}^n$ (denoted by $\Lambda_{d,d}(\mathbb{C}^n)$ in [**Ba09**]). For a given ordered d-tuple $\gamma \in O_{n,d}$ we denote by $\Pi(\gamma)$ the collection of all permutations of d elements that produce distinct d-tuples when applied to γ. Let $d_1, d_2, \ldots, d_n$ denote, respectively, the number of repetitions of $1, 2, \ldots, n$ in γ. Then the cardinal of $\Pi(\gamma)$ is given by the multinomial formula $\mathrm{Card}(\Pi(\gamma)) = \frac{d!}{d_1! \cdots d_n!}$. On $\mathcal{P}_{d,d}(Z_1, \ldots, Z_n)$ consider the sesquilinear scalar product $\langle\langle \cdot, \cdot \rangle\rangle$ so that

$$\{ Z^{(\gamma, \delta)} := (\mathrm{Card}(\Pi(\gamma)) \, \mathrm{Card}(\Pi(\delta)))^{1/2} \, Z_{\gamma(1)} \cdots Z_{\gamma(d)} \overline{Z_{\delta(1)} \cdots Z_{\delta(d)}} \ ; \ \gamma, \delta \in O_{n,d} \}$$

is an orthonormal basis. Let

$$\kappa_{d,d} : \widehat{\mathbb{C}^n} \times \cdots \times \widehat{\mathbb{C}^n} \to \mathcal{P}_{d,d}(Z_1, \ldots, Z_n) \ , \ \kappa_{d,d}(x^1, \ldots, x^d) = \prod_{k=1}^{d} |x_1^k Z_1 + \cdots + x_n^k Z_n|^2.$$

Note $x \mapsto \kappa_{d,d}(x, x, \ldots, x)$ is an embedding of $\widehat{\mathbb{C}^n}$ into $\mathcal{P}_{d,d}(Z_1, \ldots, Z_n)$. Let $P = \kappa_{d,d}(x, x, \ldots, x)$ and $Q_{k_1, \ldots, k_d} = \kappa_{d,d}(f_{k_1}, \ldots, f_{k_d})$. Then a little algebra shows that

$$(5.3) \qquad\qquad \langle\langle P, Q_{k_1, \ldots, k_d} \rangle\rangle = |\langle x, f_{k_1} \rangle|^2 \cdots |\langle x, f_{k_d} \rangle|^2.$$

Now the phase retrieval problem can be restated as the problem of finding a homogeneous polynomial P of rank 1 (that is, of the form $\kappa_{d,d}(x, \ldots, x)$) so that

$$(5.4) \qquad\qquad \langle\langle P, Q_{k_1, \ldots, k_d} \rangle\rangle = y_{k_1} \cdots y_{k_d} \ , \ \forall (k_1, \ldots, k_d) \in O_{m,d}.$$

The number of equations in (5.4) is

$$M_{m,d} = \mathrm{Card}(O_{m,d}) = \binom{m + d - 1}{d},$$

whereas the dimension of the real linear space $\mathcal{P}_{d,d}(Z_1, \ldots, Z_n)$ is

$$N_{n,d} = \dim_{\mathbb{R}} \mathbb{P}_{d,d}(Z_1, \ldots, Z_n) = (\mathrm{Card}(O_{n,d}))^2 = \binom{n + d - 1}{d}^2.$$

If d is so that $M_{m,d} \geq N_{n,d}$ and the set of polynomials $\mathcal{Q} = \{Q_\varepsilon \ ; \ \varepsilon \in O_{m,d}\}$ forms a frame for $\mathcal{P}_{d,d}(Z_1, \ldots, Z_n)$ then P can be obtained by solving a linear system (albeit of dimension growing exponentially with n and m). In particular if the set $\mathcal{Q}$ forms a d-design the reconstruction is particularly simple. The case $d = 2$ has been also explored in [**BE15**]. In the absence of noise $P \in \kappa_{d,d}(\widehat{\mathbb{C}^n})$ and thus $\hat{x}$ is found by solving a factorization problem. In the presence of noise, P is no longer of rank 1 and a different estimation procedure should be used. For instance one can find the "closest" rank 1 homogeneous polynomial in $\mathcal{P}_{d,d}(Z_1, \ldots, Z_n)$ and invert $\kappa_{d,d}$ to estimate x.

5.1.2. *PhaseLift.* Consider the noiseless case $y = \beta(x)$.   The main idea is embodied in the following feasibility problem:

$$\text{find}_{\text{subject to:} \mathbb{A}(X)=y, X=X^* \geq 0, \text{rank}(X)=1} X.$$

Except for the condition $\text{rank}(X) = 1$, the optimization problem would be convex. However the rank constraint destroys the convexity property. Once a solution X is found, the vector x can be obtained by solving the factorization problem $X = xx^*$.

The feasibility problem admits at most a unique solution and so does the following optimization problem:

$$(5.5) \qquad \min_{\mathbb{A}(X)=y, X=X^* \geq 0} \text{rank}(X),$$

which is still non-convex. The insight provided by matrix completion theory and exploited in [**CSV12, CESV12**] is to replace $\text{rank}(X)$ by $\text{trace}(X)$ which is convex. Thus one obtains

$$(5.6) \qquad (\text{PhaseLift}) \quad \min_{\mathbb{A}(X)=y, X=X^* \geq 0} \text{trace}(X),$$

which is a convex optimization problem (a semi-definite program). In [**CL12**] the authors proved that for random frames, with high probability, the problem (5.6) has the same solution as the problem (5.5):

THEOREM 5.1. *Assume each vector f_k is drawn independently from $\mathcal{N}(0, I_n/2) + i\mathcal{N}(0, I_n/2)$, or each vector is drawn independently from the uniform distribution on the complex sphere of radius $\sqrt{n}$. Then there are universal constants $c_0, c_1, \gamma > 0$ so that for $m \geq c_0 n$, for every $x \in \mathbb{C}^n$ the problem (5.6) has the same solution as (5.5) with probability at least $1 - c_1 e^{-\gamma n}$.*

As explained in [**DH14**] and [**CL12**], the minimization of trace is not necessary; in the absence of noise it reduces to a feasibility problem.

The PhaseLift algorithm is also robust to noise. Consider the measurement process

$$y = \beta(x) + \nu,$$

for some $\nu \in \mathbb{R}^m$ noise vector. Consider the following modified optimization problem:

$$(5.7) \qquad \min_{X=X^* \geq 0} \|\mathbb{A}(X) - y\|_1.$$

In [**CL12**] the following result has been shown:

THEOREM 5.2. *Consider the same stochastic process for a random frame $\mathcal{F}$. There is a universal constant $C_0 > 0$ so that for all $x \in C^n$ the solution to (5.7) obeys*

$$\|X - xx^*\|_2 \leq C_0 \frac{\|\nu\|_1}{m}.$$

For the Gaussian model this holds with the same probability as in the noiseless case, whereas the probability of failure is exponentially small in n in the uniform model. The principal eigenvector x^0 of X (normalized by the square root of the principal eigenvalue) obeys

$$D_2(x^0, x) \leq C_0 \min(\|x\|_2, \frac{\|\nu\|_1}{m\|x\|_2}).$$

5.2. Iterative Algorithms. We present two classes of iterative algorithms: the Gerchberg-Saxton algorithm and mean-squares minimization algorithms.

5.2.1. *The Gerchberg-Saxton Algorithm.* Let $c = (c_k)_{1 \leq k \leq m} \in \mathbb{C}^m$ denote a sequence of the frame coefficients $c_k = \langle x, f_k \rangle$. Let $E = \{((\langle x, f_k \rangle)_{1 \leq k \leq m}, x \in \mathbb{C}^n\}$ denote the range of frame coefficients. Assume the measurements are all nonnegative, $y_k \geq 0$ for all k (otherwise rectify at 0). Denote by $\{\tilde{f}_k, 1 \leq k \leq m\}$ the canonical dual frame of $\{f_1, \ldots, f_m\}$. The Gerchberg-Saxton algorithm first introduced in [**GS72**] iterates between two sets of constraints: $|c_k| = \sqrt{y_k}$, and $c \in E$. Let $x^0 \in H$ be an initialization and set $t = 0$. The algorithm repeats the following steps:

(1) Linear Analysis: $c_k = \langle x^t, f_k \rangle$, $1 \leq k \leq m$;
(2) Magnitude Adjustment: $d_k = \sqrt{y_k} c_k / |c_k|$, $1 \leq k \leq m$;
(3) Linear Synthesis: $x^{t+1} = \sum_{k=1}^{m} d_k \tilde{f}_k$;
(4) Increment $t = t + 1$;

until a stopping criterion is achieved. The main advantage of this algorithm is its simplicity. It can easily incorporate additional constraints on x or its transform c. Unfortunately it suffers of a couple of disadvantages. Namely, the convergence is not guaranteed, and furthermore, when it converges it only converges to a local minimum. However, despite these shortcomings, the algorithm performs relatively well when x is highly constrained, for instance when all of its entries are nonnegative (see e.g. [**Fin82**]).

5.2.2. *Mean-Squares Minimization Algorithms.* Consider again the measurement process

$$y_k = |\langle x, f_k \rangle|^2 + \nu_k, \ 1 \leq k \leq m.$$

The Least-Squares criterion

$$\min_{x \in \mathbb{C}^n} \sum_{k=1}^{m} ||\langle x, f_k \rangle|^2 - y_k|^2$$

can be understood as the Maximum Likelihood Estimator (MLE) when the additive noise vector $\nu \in \mathbb{R}^m$ is normal distributed with zero mean and covariance $\sigma^2 I_m$. However the optimization problem is not convex and has many local minima.

We present two algorithms that minimize the mean-squares error: the Wirtinger flow and the Iterative Regularized Least-Squares.

(i) Gradient Descent Using the Wirtinger Flow

This algorithm has been introduced in [**CLS14**]. The idea is to follow the gradient descent for the criterion

(5.8) $$f(x) = \sum_{k=1}^{m} |y_k - |\langle x, f_k \rangle|^2|^2.$$

The initialization is performed using the spectral method. Specifically:

Step 1. Initialization. Compute the principal eigenvector of $R_y = \sum_{k=1}^{m} y_k f_k f_k^*$ using, e.g., the power method. Let (e_1, a_1) be the eigen-pair with $e_1 \in \mathbb{C}^n$, $\|e_1\| = 1$, and $a_1 \in \mathbb{R}$. Initialize:

(5.9) $$x^0 = \sqrt{n \frac{\sum_{k=1}^{m} y_k}{\sum_{k=1}^{m} \|f_k\|^2}} e_1, \ t = 0.$$

Step 2. Iteration. Repeat:

2.1 Gradient descent:

$$(5.10) \qquad x^{t+1} = x^t - \frac{\mu_{t+1}}{\|x^0\|^2} \left(\frac{1}{m} \sum_{k=1}^{m} (|\langle x^t, f_k \rangle|^2 - y_k) \langle x^t, f_k \rangle f_k \right).$$

2.2 Update

$$(5.11) \qquad \mu^{t+1} = \min(\mu_{\max}, 1 - e^{-\tau/\tau_0}).$$

Step 3. Stopping. Stop after a fixed number of iterations or an error criterion is achieved.

The authors of [**CLS14**] showed that this algorithm converges with high probability to the exact solution in the absence of noise:

THEOREM 5.3 ([**CLS14**]). *Let $x \in \mathbb{C}^n$ and $y = \beta(x)$ with $m \geq c_0 n \log n$, where c_0 is a sufficiently large constant. Then the Wirtinger flow initial estimate x^0, normalized to have squared Euclidean norm equal to $\frac{1}{m} \sum_{k=1}^{m} y_k$, obeys $D_2(x^0, x) \leq \frac{1}{8}\|x\|$ with probability at least $1 - 10e^{-\gamma n} - \frac{8}{n^2}$, where γ is a fixed positive numerical constant. Further, take a constant learning parameter sequence, $\mu^t = \mu$ for all $t \geq 1$ and assume $\mu \leq \frac{c_1}{n}$ for some fixed numerical constant c_1. Then there is an event of probability at least $1 - 13e^{-\gamma n} - me^{-1.5m} - \frac{8}{n^2}$, such that on this event, starting from any initial solution x^0 obeying $D_2(x^0, x) \leq \frac{1}{8}\|x\|$, we have*

$$(5.12) \qquad D_2(x^t, x) \leq \frac{1}{8} \left(1 - \frac{\mu}{4} \right)^{t/2} \|x\|.$$

(ii) The Iterative Regularized Least-Squares Algorithm

The iterative algorithm described next tries to find the global minimum using a regularization term. Consider the following optimization criterion:

$$(5.13) \qquad J(u, v; \lambda, \mu) = \sum_{k=1}^{m} \left| \frac{1}{2} (\langle u, f_k \rangle \langle f_k, v \rangle + \langle v, f_k \rangle \langle f_k, u \rangle) - y_k \right|^2$$
$$+ \lambda \|u\|_2^2 + \mu \|u - v\|_2^2 + \lambda \|v\|_2^2.$$

The *Iterative Regularized Least-Squares (IRLS)* algorithm presented in [**Ba13**] works as follows.

Fix a stopping criterion, such as a tolerance ε, a desired level of signal-to-noise-ratio snr, or a maximum number of steps T. Fix an initialization parameter $\rho \in (0, 1)$, a learning rate $\gamma \in (0, 1)$, and a saturation parameter $\mu_{\min} > 0$.

Step 1. Initialization. Compute the principal eigenvector of $R_y = \sum_{k=1}^{m} y_k f_k f_k^*$ using e.g. the power method. Let (e_1, a_1) be the eigen-pair with $e_1 \in \mathbb{C}^n$, $\|e_1\| = 1$, and $a_1 \in \mathbb{R}$. If $a_1 \leq 0$ then set $x = 0$ and exit. Otherwise initialize:

$$(5.14) \qquad x^0 = \sqrt{\frac{(1 - \rho)a_1}{\sum_{k=1}^{m} |\langle e_1, f_k \rangle|^4}} \, e_1 \ , \ \lambda_0 = \rho a_1 \ , \ \mu_0 = \rho a_1 \ , \ t = 0.$$

Step 2. Iteration. Perform:
2.1 Solve the least-squares problem:

$$x^{t+1} = \operatorname{argmin}_u J(u, x^t; \lambda_t, \mu_t)$$

using the conjugate gradient method.
2.2 Update:

$$\lambda_{t+1} = \gamma \lambda_t \ , \ \mu_t = \max(\gamma \mu_t, \mu_{\min}) \ , \ t = t + 1.$$

Step 3. Stopping. Repeat Step 2 until

- The error criterion is achieved: $J(x^t, x^t; 0, 0) < \varepsilon$;
- The desired signal-to-noise-ratio is reached: $\frac{\|x^t\|^2}{J(x^t, x^t; 0, 0)} > \text{snr}$; or
- The maximum number of iterations is reached: $t > T$.

The final estimate can be x^T or the best estimate obtained in the iteration path: $x^{\text{est}} = x^{t_0}$ where $t_0 = \operatorname{argmin}_t J(x^t, x^t; 0, 0)$.

The initialization (5.14) is performed for the following reason. Consider the modified criterion:

$$
\begin{aligned}
H(x; \lambda) &= J(x, x; \lambda, 0) = \|\beta(x) - y\|_2^2 + 2\lambda \|x\|_2^2 \\
&= \sum_{k=1}^m |\langle x, f_k \rangle|^4 + 2\langle (\lambda I_n - R_y)x, x \rangle + \|y\|_2^2.
\end{aligned}
$$

In general this function is not convex in x, except for large values of λ. Specifically for $\lambda > a_1$, the largest eigenvalue of R_y, $x \mapsto H(x; \lambda)$ is convex and has a unique global minimum at $x = 0$. For $a_1 - \varepsilon < \lambda < a_1$ the criterion is no longer convex, but the global minimum stays in a neighborhood of the origin. Neglecting the 4th order terms, the critical points are given by the eigenvectors of R_y. Choosing $\lambda = \rho a_1$ and $x = s e_1$, the optimal value of s for $s \mapsto H(s e_1; \rho a_1)$ is given in (5.14).

The path of iterates $(x^t)_{t \geq 0}$ can be thought of as trying to approximate the measured vector y with $\mathbb{A}([\![x^{t-1}, x^t]\!])$, where $\mathbb{A}$ is defined in (3.8). The parameter μ penalizes the unique negative eigenvalue of $[\![x^{t-1}, x^t]\!]$; the larger the value of μ_t the smaller the iteration step $\|x^{t+1} - x^t\|$ and the smaller the deviation of the matrix $[\![x^{t+1}, x^t]\!]$ from a rank 1 matrix; the smaller the parameter μ_t the larger in magnitude the negative eigenvalue of $[\![x^{t+1}, x^t]\!]$. This fact explains why in the noisy case the iterates first decrease the matching error $\|\mathbb{A}(x^t (x^t)^*) - y\|_2$ to some t_0, and then they may start to increase this error; instead the rank 2 self-adjoint operator $T = [\![x^{t+1}, x^t]\!]$ always decreases the matching error $\|\mathbb{A}(T) - y\|_2$.

At any point on the path, if the value of criterion J is smaller than the value reached at the target vector x, then the algorithm is guaranteed to converge near x. Specifically in [**Ba13**] the following result has been proved:

THEOREM 5.4 ([**Ba13**] Theorem 5.6). *Fix $0 \neq z_0 \in \mathbb{C}^n$. Assume the frame $\mathcal{F}$ is so that $\ker \mathbb{A} \cap \mathcal{S}^{2,1} = \{0\}$. Then there is a constant $A_3 > 0$ that depends on $\mathcal{F}$ so that for every $x \in \mathbb{C}^n$ with $\langle x, z_0 \rangle > 0$ and $\nu \in \mathbb{C}^n$ that produce $y = \beta(x) + \nu$ if there are $u, v \in \mathbb{C}^n$ so that $J(u, v; \lambda, \mu) < J(x, x; \lambda, \mu)$ then*

$$
(5.15) \qquad \| [\![u, v]\!] - xx^* \|_1 \leq \frac{4\lambda}{A_3} + \frac{2\|\nu\|_2}{\sqrt{A_3}}.
$$

Moreover, let $[\![u, v]\!] = a_1 e_1 e_1^* + a_2 e_2 e_2^*$ be its spectral factorization with $a_1 \geq 0 \geq a_2$ and $\|e_1\| = \|e_2\| = 1$. Set $\tilde{x} = \sqrt{a_1} e_1$. Then

$$
(5.16) \qquad D_2(x, \tilde{x})^2 \leq \frac{4\lambda}{A_3} + \frac{2\|\nu\|_2}{\sqrt{A_3}} + \frac{\|\nu\|_2^2}{4\mu} + \frac{\lambda \|x\|_2^2}{2\mu}.
$$

The kernel requirement on $\mathbb{A}$ is satisfied for generic frames when $m \geq 6n$. In particular this condition requires the frame $\mathcal{F}$ is phase retrievable for $\mathbb{C}^n$.

References

ABFM12. B. Alexeev, A. S. Bandeira, M. Fickus, D. G. Mixon, *Phase Retrieval with Polarization*, SIAM J. Imaging Sci., **7** (1) (2014), 35–66.

BE15. C. Bachoc, M. Ehler: *Signal reconstruction from the magnitude of subspace components*, available online arXiv:1209.5986, IEEE Trans. Inform. Theory, vol. 61 (2015), no. 7, 1-13.

Ba09. R. Balan, *A Nonlinear Reconstruction Algorithm from Absolute Value of Frame Coefficients for Low Redundancy Frames*, Proceedings of SampTA Conference, Marseille, France May 2009.

Ba10. R. Balan, *On Signal Reconstruction from Its Spectrogram*, Proceedings of the CISS Conference, Princeton NJ, May 2010.

Ba12. R. Balan, *Reconstruction of Signals from Magnitudes of Redundant Representations*, available online arXiv:1207.1134v1 [math.FA] 4 July 2012.

Ba13. R. Balan, *Reconstruction of Signals from Magnitudes of Redundant Representations: The Complex Case*, available online arXiv:1304.1839v1 [math.FA] 6 April 2013, to appear in Foundations of Computational Mathematics (2015).

Ba15. R. Balan, *Stability of frames which give phase retrieval*, to appear in the Houston Journal of Mathematics 2015.

Bal15. R. Balan, *The Fisher Information Matrix and the CRLB in a Non-AWGN Model for the Phase Retrieval Problem*, proceedings of SampTA 2015.

BCE06. R. Balan, P. Casazza, D. Edidin, *On signal reconstruction without phase*, Appl. Comput. Harmon. Anal. **20** (2006), 345–356.

BBCE07. R. Balan, B. Bodmann, P. Casazza, D. Edidin, *Painless reconstruction from Magnitudes of Frame Coefficients*, J.Fourier Anal.Applic., **15** (4) (2009), 488–501.

BW13. R. Balan and Y. Wang, *Invertibility and Robustness of Phaseless Reconstruction*, available online arXiv:1308.4718v1, Appl. Comp. Harm. Anal., **38** (2015), 469–488.

BZ14. R. Balan and D. Zou, *Phase Retrieval using Lipschitz Continuous Maps*, available online arXiv:1403.2304v1.

BZ15a. R. Balan and D. Zou, *On Lipschitz Inversion of Nonlinear Redundant Representations*, Contemporary Mathematics 650, "Trends in Harmonic Analysis and Its Applications", 15–22, 2015.

BZ15b. R. Balan and D. Zou, *On Lipschitz Analysis and Lipschitz Synthesis for the Phase Retrieval Problem*, available online arXiv:1506.02092

BCMN13. A.S. Bandeira, J. Cahill, D.G. Mixon, A.A. Nelson, *Saving phase: Injectivity and stability for phase retrieval*, available online arXiv:1302.4618v2. Appl. Comp. Harm. Anal. **37** (1) (2014), 106–125.

BL00. Y. Benyamini, J. Lindenstrauss, **Geometric Nonlinear Functional Analysis**, vol. 1, AMS Colloquium Publications, vol. 48, 2000.

Bh97. R. Bhatia, **Matrix Analysis**, Graduate Texts in Mathematics 169, Springer-Verlag 1997.

BH13. B. G. Bodmann and N. Hammen, *Stable Phase Retrieval with Low-Redundancy Frames*, available online arXiv:1302.5487v1. Adv. Comput. Math., accepted 10 April 2014.

CCPW13. J. Cahill, P.G. Casazza, J. Peterson, L. Woodland, *Phase retrieval by projections*, available online arXiv: 1305.6226v3

CSV12. E. Candés, T. Strohmer, V. Voroninski, *PhaseLift: Exact and Stable Signal Recovery from Magnitude Measurements via Convex Programming*, Communications in Pure and Applied Mathematics vol. 66, 1241–1274 (2013).

CESV12. E. Candés, Y. Eldar, T. Strohmer, V. Voroninski, *Phase Retrieval via Matrix Completion Problem*, SIAM J. Imaging Sci., **6**(1) (2013), 199-225.

CL12. E. Candès, X. Li, *Solving Quadratic Equations Via PhaseLift When There Are As Many Equations As Unknowns*, available online arXiv:1208.6247

CLS14. E. J. Cands, X. Li and M. Soltanolkotabi, *Phase retrieval via Wirtinger flow: theory and algorithms*, IEEE Transactions on Information Theory 61(4), (2014) 1985–2007.

Ca00. P. Casazza, *The art of frame theory*, Taiwanese J. Math., (2) **4** (2000), 129–202.

CEHV13. A. Conca, D. Edidin, M. Hering, V. Vinzant, *An algebraic characterization of injectivity in phase retrieval*, available online arXiv:1312:0158, to appear in ACHA 2014.

DH14. L. Demanet, P. Hand, *Stable optimizationless recovery from phaseless linear measurements*, available online arXiv:1208.1803v1, J. Fourier Anal. Appl., 20(1):199-221, 2014.

EM12. Y. C. Eldar, S. Mendelson, *Phase retrieval: Stability and recovery guarantees*, Available online: arXiv:1211.0872.

FMNW13. M. Fickus, D.G. Mixon, A.A. Nelson, Y. Wang, *Phase retrieval from very few measurements*, available online arXiv:1307.7176v1. Linear Algebra and its Applications **449** (2014), 475-499

Fin82. J.R. Fienup. Phase retrieval algorithms: A comparison. *Applied Optics*, 21(15):2758–2768, 1982.

GS72. R. W. Gerchberg and W. O. Saxton, *A practical algorithm for the determination of the phase from image and diffraction plane pictures*, Optik 35, 237 (1972).

HLO80. M. H. Hayes, J. S. Lim, and A. V. Oppenheim, *Signal Reconstruction from Phase and Magnitude*, IEEE Trans. ASSP **28**, no.6 (1980), 672–680.

HMW11. T. Heinosaari, L. Mazzarella, M. M. Wolf, *Quantum Tomography under Prior Information*, available online arXiv:1109.5478v1 [quant-ph], 26 Sept 2011. Commun. Math. Phys. **318** (2) (2013), 355–374.

HG13. M.J. Hirn and E. Le Gruyer, *A general theorem of existence of quasi absolutely minimal Lipschitz extensions*, arXiv:1211.5700v2 [math.FA], 8 Aug 2013.

Ky10. S. M. Kay, **Fundamentals of Statistical Signal Processing. I. Estimation Theory**, Prentice Hall PTR, 18th Printing, 2010.

MV13. D. Mondragon and V. Voroninski, *Determination of all pure quantum states from a minimal number of observables*, online arXiv:1306.1214v1 [math-ph] 5 June 2013.

Mi67. R. J. Milgram, *Immersing Projective Spaces*, Annals of Mathematics, vol. **85**, no. 3 (1967), 473–482.

NQL82. H. Nawab, T. F. Quatieri, and J. S. Lim, *Signal Reconstruction from the Short-Time Fourier Transform Magnitude*, in Proceedings of ICASSP 1984.

Viz15. C. Vinzant, *A small frame and a certificate of its injectivity*, available online arXiv:1502.0465v1 [math.FA] 16 Feb. 2015.

WAM12. I. Waldspurger, A. d'Aspremont, S. Mallat, *Phase recovery, MaxCut and complex semidefinite programming*, Available online: arXiv:1206.0102

WW75. J.H. Wells and L.R. Williams, **Embeddings and Extensions in Analysis**, Ergebnisse der Mathematik und ihrer Grenzgebiete Band 84, Springer-Verlag 1975.

ZB06. L. Zwald and G. Blanchard, *On the convergence of eigenspaces in kernel Principal Component Analysis*, Proc. NIPS 05, vol. 18, 1649-1656, MIT Press, 2006.

Index

Department of Mathematics, Center for Scientific Computation and Mathematical Modeling, Norbert Wiener Center, University of Maryland, College Park MD 20742
E-mail address: rvbalan@math.umd.edu