APPLIED PROBLEMS IN FRAME THEORY

A Dissertation

presented to

the Faculty of the Graduate School

University of Missouri

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

SARA BOTELHO-ANDRADE

Dr. Peter Casazza, Dissertation Supervisor

May 2019

The undersigned, appointed by the Dean of the Graduate School, have examined the dissertation entitled

APPLIED PROBLEMS IN FRAME THEORY

presented by Sara Botelho-Andrade,

a candidate for the degree of Doctor of Philosophy of Mathematics,

and hereby certify that in their opinion it is worthy of acceptance.

Professor Peter Casazza

Professor Stephen Montgomery-Smith

Professor Nakhle Asmar

Professor David Retzloff

ACKNOWLEDGEMENTS

First of all, I wish to express my sincere gratitude to my advisor, Professor Peter Casazza, for his vision and direction. He and his wife, Janet Tremain, have shown unwavering support and are perhaps some of the most generous people I've had the pleasure of knowing.

I would also like to thank my committee members Dr. Nakhle Asmar, Dr. Stephen Montgomery-Smith, and Dr. David Retzloff for their support.

I am sincerely grateful to my friends and family for their support during these years, and to the other graduate students who I've had the pleasure of working with.

Contents

Acknowledgements **ii**

0.1 Preliminaries vii

1 The Phase Retrieval Problem **1**

1.0.1 Lifting 5

1.1 The Phase Property 8

1.1.1 Weak Phase Property 16

1.1.2 Weak Phase Retrieval 22

1.2 Illustrative Examples 25

1.3 Phase Retrieval in ℓ_2 29

1.3.1 Lifting 34

1.3.2 Sets Which do Phase Retrieval in ℓ_2 35

2 Quantum Detection Problem **40**

2.0.1 The Quantum Detection Problem 42

2.1 The Finite Dimensional Case 45

2.1.1 Constructing the Solutions to the Injectivity Problem 56

2.1.2 The Solutions are Open and Dense 60

2.1.3 Solution to the State Estimation Problem 65

2.2 The Infinite Dimensional Case 69

2.2.1 The Solution to the Injectivity Problem 70

2.2.2 Constructing the Solutions to the Injectivity Problem 88

2.2.3 The Solutions are Neither Open nor Dense 91

2.2.4 The Solution to the State Estimation Problem 95

Bibliography **111**

Vita **112**

APPLIED PROBLEMS IN FRAME THEORY

Sara Botelho-Andrade

Dr. Peter Casazza, Dissertation Supervisor

ABSTRACT

This thesis is a study of two applied problems in frame theory: phase retrieval and quantum detection. These problems are inspired by engineering applications in signal processing and information theory.

In signal processing, phase retrieval is the problem of retrieving a signal from a set of intensity measurements. Motivation for this problem comes from engineering applications where phase information is lost, often after passing through a filter or from the measurement process itself. Practical applications include X-ray crystallography, diffraction imaging, optics, speech processing, deep learning, and quantum information theory. In the discrete setting, these measurements correspond to the magnitude of the inner products with the given frame vectors, or $|\langle x, x_k\rangle|$. By generalizing known characterizations of the phase retrieval problem, we arrive at conditions for embedding phase retrievable frames in larger dimensional spaces. We go on to consider a related problem of recovering the *phase* of a vector from given a set of intensity measurements, called the phase property. The study of the phase property motivated a investigation of weakened notions of both problems in Sections 1.1.1 and 1.1.2. The last section in this chapter is aimed at observing the differences between phase retrieval in finite and infinite dimensions. While most characterizations carry over from finite dimensions, there are some surprising differences.

The other problem we will consider, quantum detection, has applications in op-

tical communications, including the detection of coherent light signals such as radio, radar, and laser signals. Quantum detection theory is a reformulation, in quantum-mechanical terms, of statistical decision theory. In this thesis, we consider a Hilbert space frame version of a quantum detection problem. The quantum detection problem can be deconstructed as follows: the injectivity problem and the state estimation problem. We begin by considering the problem in a more general setting and then will show the desired results as particular cases.

The goal of the injectivity problem is to classify frames which are injective with respect to self-adjoint Hilbert-Schmidt operators. By associating vectors $x \in \mathbb{H}^n$ with vectors $\tilde{x}$ in a larger space, we are able to use standard linear algebra and functional analysis techniques to provide characterizations for the injectivity problem in complex and real Hilbert spaces, as well as construct solutions. Given an injective frame, the goal of the state estimation problem is to construct a self-adjoint trace one operator T such that the vector with coordinates $\langle Tx_k, x_k \rangle$ is equal to a predetermined measurement vector. We give equivalent conditions for solvability of the state estimation problem and provide *best* approximate solutions when no exact solution is possible. We also show results about density of both problems.

0.1 Preliminaries

The purpose of this section is to introduce some basic definitions and results from frame theory. For a background on Hilbert space frame theory see [14, 16, 18]. $\mathbb{H}^n$ will be used to denote an n-dimensional real or complex Hilbert space. When applicable, $\mathbb{H}$ will be used to denote a separable Hilbert space. When it is necessary to differentiate the two, the usual notation will be used: $\mathbb{R}^n$ or $\mathbb{C}^n$. As usual ℓ_2 denotes the space of square summable sequences. Throughout this thesis, let $\{e_k\}_{k\in I}$ denote the canonical basis of $\mathbb{H}$. Also, ι will be used to denote the complex unit.

Frames are often considered a generalization of orthogonal basis. By relaxing the orthogonality condition, we gain redundancy. The advantage is that redundant systems can recover lost information, as we will see in the first chapter of this thesis. We start with the definition of a frame in $\mathbb{H}$, which is reminescent of Parseval's identity.

Definition 0.1. A family of vectors $\mathcal{X} = \{x_k\}_{k\in I}$ is a **frame** for (a real or complex) Hilbert space $\mathbb{H}$ if there are constants $0 < A \leq B < \infty$ satisfying:

$$A\|x\|^2 \leq \sum_{k\in I} |\langle x, x_k\rangle|^2 \leq B\|x\|^2, \text{ for all } x \in \mathbb{H}.$$

We have

1. A, B are the **lower and upper frame bounds** of the frame.

2. If $A = B$ this is a **tight frame**. If $A = B = 1$ this is a **Parseval frame**.

3. If we only assume we have $0 < B < \infty$, this is called a **B-Bessel sequence**. Note that $\|x_k\|^2 \leq B$, for all $k \in I$.

In finite dimensions, the definition of a frame is equivalent to a spanning set. However, in infinite dimensions there are examples of spanning sets which do not satisfy the frame inequality.

For a frame vector x_k in $\mathbb{R}^n$ or $\mathbb{C}^n$, we denote its coordinates as

$$x_k = (x_{k1}, x_{k2}, \ldots, x_{kn}),$$

unless otherwise noted. Similarly, we extend this notation for x_k belonging to ℓ_2.

We define the **analysis operator** of the frame as $T : \mathbb{H} \to \ell_2(I)$ by

$$T(x) = (\langle x, x_1\rangle, \langle x, x_2\rangle, \ldots) = \sum_{k\in I} \langle x, x_k\rangle e_k.$$

The **synthesis operator** T^* is given by:

$$T^*\left(\{a_k\}_{k\in I}\right) = \sum_{k\in I} a_k x_k.$$

The **frame operator** is $S = T^*T$. This is a positive, self-adjoint, invertible operator on $\mathbb{H}$ satisfying:

$$Sx = \sum_{k\in I} \langle x, x_k\rangle x_k.$$

With this definition, one reformulation of the frame definition is that the numerical range of S, the set $\langle Sx, x\rangle$ for all $\|x\| = 1$, is an interval in the positive reals. It is known that for any frame $\{x_k\}_{k\in I}$, $\{S^{-1/2}x_k\}_{k\in I}$ is a Parseval frame. It is also known that a frame is Parseval if and only if its frame operator is the identity operator.

Definition 0.2. A frame $\{x_k\}_{k\in I}$ is said to be **bounded** if there is a constant $C > 0$ such that $\|x_k\| \geq C,$ for all $k \in I$.

Chapter 1

The Phase Retrieval Problem

In signal processing, phase retrieval is the problem of retrieving a signal from a set of intensity measurements. The problem has been studied by engineers for many years. Signals passing through linear systems may result in lost or distorted phase information, often after passing through a filter or from the measurement process itself.. This partial loss of phase information occurs in various applications including speech recognition [6, 33, 34], and optics applications such as X-ray crystallography [5, 24, 25]. The concept of *phase retrieval* for Hilbert space frames was introduced in 2006 by Balan, Casazza, and Edidin [3] and since then it has become an active area of research. In the discrete setting, these measurements correspond to the magnitude of the inner products with the given frame vectors, or $|\langle x, x_k\rangle|$. Measurements of this type have an inherent ambiguity, since $|\langle x, x_k\rangle| = |\langle e^{\iota\theta}x, x_k\rangle|$ for all $\theta \in \mathbb{R}$. One question is what is necessary to recover the phase of a signal, given intensity measurements from a redundant linear system? Another questions is given these measurements can we recover the unknown signal itself? By *phase*, we are referring to the unimodular portion of the polar decomposition of x. We will show that these questions are equivalent, but it is not obvious from the definition.

Phase retrieval has been defined for vectors as well as for projections. *Phase*

retrieval by projections occur in real life problems, such as crystal twinning [20], where the signal is projected onto some higher dimensional subspaces and has to be recovered from the norms of the projections of the vectors onto the subspaces. We refer the reader to [12] for a detailed study of phase retrieval by projections. At times these projections are identified with their target spaces. Determining when subspaces $\{W_i\}_{i=1}^m$ and $\{W_i^\perp\}_{i=1}^m$ both do phase retrieval has given way to the notion of *norm retrieval* [1], another important area of research.

Next, we give the formal definitions of phase retrieva and norm retrieval.

Definition 1.1. Let $\mathcal{X} = \{x_k\}_{k\in I}$ be a family of vectors in $\mathbb{H}$ (resp. $\{P_k\}_{k\in I}$ is a family of projections on $\mathbb{H}$) satisfying: for every non-zero vectors x and y and

$$|\langle x, x_k\rangle|^2 = |\langle y, x_k\rangle|^2, \text{ for all } k \in I. \tag{1.1}$$

Respectively,

$$\|P_k x\|^2 = \|P_k y\|^2, \text{ for all } k \in I. \tag{1.2}$$

1. If this implies there is a $|\theta| = 1$ so that $x = \theta y$, we say $\mathcal{X}$ does **phase retrieval.** (Respectively, $\{P_k\}_{k\in I}$ does **phase retrieval**.)

2. If this implies $\|x\| = \|y\|$, we say $\mathcal{X}$ does **norm retrieval**. (Respectively, $\{P_k\}_{k\in I}$ does **norm retrieval**.)

Moreover, in the real case, if $\theta = 1$ we say x and y have the same signs and if $\theta = -1$ we say x and y have opposite signs.

We note that tight frames $\mathcal{X} = \{x_k\}_{k\in I}$ for $\mathbb{H}$ do norm retrieval. Indeed, if

$$|\langle x, x_k\rangle| = |\langle y, x_k\rangle|, \text{ for all } k \in I$$

then

$$A\|x\|^2 = \sum_{k\in I} |\langle x, x_k\rangle|^2 = \sum_{k\in I} |\langle y, x_k\rangle|^2 = A\|y\|^2.$$

Phase retrieval in $\mathbb{R}^n$ is classified in terms of a fundamental result called the complement property, which we define below:

Definition 1.2. A frame $\mathcal{X} = \{x_k\}_{k\in I}$ in $\mathbb{H}$ satisfies the **complement property** if for all subsets $S \subset I$, either $\overline{\text{span}}\{x_k\}_{k\in S} = \mathbb{H}$ or $\overline{\text{span}}\{x_k\}_{k\in S^c} = \mathbb{H}$.

Theorem 1.3 ([3] [11])**.** *If $\mathcal{X}$ does phase retrieval in $\mathbb{H}$ then it has complement property. In a real Hilbert space, if $\mathcal{X}$ has complement property then it does phase retrieval.*

In fact this definition fully classifies phase retrievable frames in the real setting. It follows that if $\mathcal{X} = \{x_i\}_{i=1}^m$ does phase retrieval in $\mathbb{R}^n$ then $m \geq 2n-1$. This theorem only states that the complement property is necessary for complex phase retrieval and the minimum number of measurements necessary remains open. It was conjectured that the minimum number of vectors necessary in $\mathbb{C}^n$ was $4n-4$, but a counter example was shown in [36].

For an elementary example of vectors doing phase retrieval we give the definition of full spark.

Definition 1.4. Given a family of vectors $\mathcal{X} = \{x_i\}_{i=1}^m$ in $\mathbb{H}^n$, the **spark** of $\mathcal{X}$ is defined as the cardinality of the smallest linearly dependent subset of $\mathcal{X}$. When $\text{spark}(\mathcal{X}) = n+1$, every subset of size n is linearly independent, and in that case, $\mathcal{X}$ is said to be **full spark**.

Notice that full spark is stronger than the complement property. That is, full spark frames with $m \geq 2n-1$ have the complement property and hence do phase retrieval. Moreover, if $m = 2n-1$ then the complement property clearly implies full spark. To construct a set of vectors with this property begin with any basis $\{e_k\}_{k=1}^n$ in $\mathbb{R}^n$. Let $H_i = \text{ span}_{k\neq i}\, e_k$. Then pick $x_{m+1} \in \left(\cup_{i\in[n]} H_i\right)^c$. This new set of vectors has the property that every n vectors are linearly independent. Repeating this procedure we obtain a set with $(2n-1)$ vectors, this set will do phase retrieval.

One approach for classifying phase retrieval uses the association between vectors and rank one Hermitian matrices (see [4]). The idea is to lift the problem to the higher dimensional space of Hermitian matrices, where the problem becomes linear. In the following theorem, the authors exploit this approach to derive a complement property classification for phase retrieval by projections.

Theorem 1.5 ([12]). *Let P_j be othorgonal projections onto linear subspaces W_j for $1 \leq j \leq m$. Then for every orthonormal basis $(\phi_{j,k})_{k=1}^{n_j}$ of W_j, the collection of vectors $\{\phi_{j,k} \ : \ 1 \leq j \leq m, 1 \leq k \leq n_j\}$ does phase retrieval.*

Next we present a second classification of phase retrieval by projections. The original proof for finite dimensional case was first presented in [21], we extend the proof to all real separable Hilbert spaces.

Theorem 1.6 ([21]). *A family of projections $\{P_i\}_{i\in I}$ on a real Hilbert space $\mathbb{H}$ does phase retrieval if and only if for every $0 \neq x \in \ell_2$, $\overline{\text{span}}\{P_i x\}_{i=1}^{\infty} = \mathbb{H}$.*

Proof. ($\Rightarrow$) We proceed by way of contradiction. So assume that there is an $0 \neq x \in \ell_2$ and $\{P_i x\}_{i=1}^{\infty}$ does not span ℓ_2. Choose $0 \neq y \in \ell_2$ so that $y \perp P_i x$ for all $i = 1, 2, \ldots$.

Let $u = x + y$ and $v = x - y$. Then since $P_i y \perp P_i x$ for all i, we have that

$$\|P_i(x+y)\|^2 = \|P_i x\|^2 + \|P_i y\|^2 = \|P_i(x-y)\|^2.$$

If $\{P_i\}_{i=1}^{\infty}$ does phase retrieval, then $x + y = \pm(x - y)$. This implies $x = 0$ or $y = 0$, which is a contradiction.

($\Leftarrow$) The proof of this theorem in [13] works directly here. ∎

1.0.1 Lifting

In this section we demonstrate an embedding of finite frames in higher dimensions such that the complement property is preserved, which we will refer to as "lifting". We provide necessary and sufficient conditions for when such a construction is possible and an example to demonstrate problems that may arise in infinite dimensions. We begin with a few useful definitions.

Definition 1.7. A frame $\mathcal{X} = \{x_i\}_{i \in I}$ has the overcomplete complement property if for every $S \subset I$, either $\{x_i\}_{i \in S}$ or $\{x_i\}_{i \in S^c}$ spans and is linearly dependent, i.e. it is not a basis.

The overcomplete complement property is a natural generalization of the usual complement property, as will be shown shortly. Next we specify exactly what types of embeddings we are considering.

Definition 1.8. A frame $\{y_i\}_{i=1}^m \subset \mathbb{R}^{n+k}$ is a **k-lifting** of a frame $\{x_i\}_{i=1}^m$ if

$$y_i|_{\mathbb{R}^n} = x_i, \text{ for all } i = 1, 2, \ldots, m. \tag{1.3}$$

The next theorem classifies when 1-lifts are possible and provides a construction for the choice of coordinates to adjoin.

Theorem 1.9. *A phase retrievable frame $\mathcal{X} = \{x_i\}_{i=1}^m \subset \mathbb{R}^n$ can be 1-lifted to a phase retrievable frame if and only if $\mathcal{X}$ has the overcomplete complement property.*

Proof. For the sufficiency we shall provide a constructive proof. The idea of the proof will be to produce a vector $v \in \mathbb{R}^m$ such that the i^{th} coordinate of v will be the $(n+1)^{th}$ coordinate of $\hat{x}_i$. Given a subset $S \subset [m]$, by assumption either $\mathcal{X}_S = \{x_i\}_{i\in S}$ or $\chi_{S^c} = \{x_i\}_{i\in S^c}$ spans $\mathbb{R}^n$ and is linearly dependent. We begin by demonstrating an embedding of vectors from the spanning set that still span in $\mathbb{R}^{n+1}$. Without loss of generality, in our notation we shall assume χ_S is always the overcomplete spanning set of vectors. Then for some choice of coefficients we have $\sum_{i\in S} \alpha_i x_i = 0$ where α_i are not all zero. Denote $\alpha_S = (\alpha_1, \alpha_2, ..., \alpha_{|S|}) \in \mathbb{R}^{|S|}$ and pick $\beta_S \in \mathbb{R}^{|S|}$ such that $\langle \alpha_S, \beta_S \rangle \neq 0$. Define the embedded vectors $\hat{\mathcal{X}}_S = \{\hat{x}_i\}_{i\in S} \in \mathbb{R}^{n+1}$ as follows

$$\hat{x}_i(j) = \begin{cases} x_i(j) & j \in [n] \\ \beta_S(i) & j = n+1. \end{cases} \tag{1.4}$$

Where $x_i(j) = x_{ij}$ denotes the j^{th} coordinate of x_i. To show that $\hat{\mathcal{X}}_S$ spans $\mathbb{R}^{n+1}$, observe $\frac{1}{\langle \alpha_S, \beta_S \rangle} \sum_{i\in S} \alpha_i \hat{x}_i = e_{n+1}$. Since $\mathcal{X}_S$ spans $\mathbb{R}^n$, it follows that $\hat{\mathcal{X}}_S$ spans $\mathbb{R}^{n+1}$. This construction gives a procedure for an embedding which spans the larger space $\mathbb{R}^{n+1}$, but is dependent on the subset S. Also observe we haven't posed any conditions on how to extend the vectors in S^c. For each choice of S, we have the associated vectors $\alpha_S, \beta_S \in \mathbb{R}^{|S|}$. Let $H_S \subset \mathbb{R}^{|S|}$ denote the hyperplane perpendicular to α_S. Then our construction depends on being able to choose a vector in the complement of H_S for all subsets S. But the cardinality of S is changing as we range over all possibilities. To overcome this we will work with the larger space $\mathbb{R}^m = \mathbb{R}^{|S|} \times \mathbb{R}^{|S^c|}$. There are finitely many choices of S therefore $\bigcup_{S\subset[m]} H_S \times \mathbb{R}^{|S^c|} \neq \mathbb{R}^m$. Then for

$v \in \left(\bigcup_{S\subset[m]} H_S \times \mathbb{R}^{|S^c|}\right)^c$ we defined

$$\hat{x}_i(j) = \begin{cases} x_i(j) & j \in [n] \\ v(i) & j = n+1. \end{cases} \tag{1.5}$$

Then it follows that $\hat{\mathcal{X}} = \{\hat{x}_i\}_{i=1}^m$ has the complement property in $\mathbb{R}^{n+1}$.

For necessity assume $\mathcal{X}$ does phase retrieval but does not have the overcomplete complement property. Any spanning set that is a basis cannot be 1-lifted since by our assumption there exists S such that vectors in S nor S^c are linearly dependant.

■

The result above may be generalized for a k-lift with minimal effort. Naturally the overcompleteness of each subset S is critical in determining what integers k are plausible. More specifically, we define the lifting number of a phase retrievable frame as follows:

Definition 1.10. Given a frame $\mathcal{X} = \{x_i\}_{i\in[m]} \subset \mathbb{R}^n$, let

$$L_{\mathcal{X}} = \min\{|S| - n : \text{span}\{x_i\}_{i\in S} = \mathbb{R}^n \text{ and } |S| \geq |S^c|\}, \tag{1.6}$$

then $L_{\mathcal{X}}$ is the lifting number for the frame $\mathcal{X}$.

From the previous theorem we see immediately that if $\mathcal{X}$ has the overcomplete complement property then $L_{\mathcal{X}} \geq 1$. The lifting number tells us how many dimensions higher we can lift $\mathcal{X}$. If $L_{\mathcal{X}} \geq 1$ then when we 1-lift, each overcomplete spanning subset will be lifted to a spanning set in $\mathbb{R}^{n+1}$ with the same cardinality. If $L_{\mathcal{X}} > 1$ that means each spanning subset with higher cardinality (S or S^c) will be lifted to a spanning set which is still not a basis in $\mathbb{R}^{n+1}$, hence can be lifted again. The idea is

that after each lift, the lifting number of the subsequent lifted frame $\hat{\mathcal{X}}$ is one smaller than $L_{\mathcal{X}}$. That is, if $\hat{\mathcal{X}}$ is a 1-lift of $\mathcal{X}$ then $L_{\hat{\mathcal{X}}} = L_{\mathcal{X}} - 1$.

Corollary 1.11. *$\mathcal{X} \subset \mathbb{R}^n$ can be k-lifted if and only if $k \leq L_{\mathcal{X}}$.*

Theorem 1.12. *If a frame $\mathcal{X} \subset \mathbb{R}^n$ contains $2n + 2m + 1$ vectors with a $2n + 2m$ full spark subset. $\mathcal{X}$ can be $(m + 1)$-lifted.*

Proof. Clearly if a frame contains a $2n + 2m$ full spark subset it does phase retrieval as it contains a $2n - 1$ full spark subset which already does phase retrieval. Let $\mathcal{X} = \{x_i\}_{i\in[2n+2m+1]}$ and $H = \{x_i\}_{i\in[2n+2m]}$ be a full spark subset. Given any $S \subset [2n + 2m]$, if S contains more than half the elements in H then it will be a spanning set with more than $n + m$ vectors hence its cardinality minus n will be greater than m hence at least $m + 1$. If it contains less than half of the elements of H then the same holds for S^c. If it contained exactly half then both S and S^c will contain $n + m$ elements of H hence they both span. Whichever set that contains $x_{2n+2m+1}$ will be a spanning set of cardinality $n + m + 1$. Hence $\mathcal{X}$ will have lifting number $m + 1$. ∎

The set of $2n + 2m + 1$ full spark vectors is open, dense, and contains a subset of $2n + 2m$ full spark vectors. Then the previous theorem shows the set of $2n + 2m + 1$ vectors in $\mathbb{R}^n$ that can be $m+1$-lifted contains an open dense set. Hence almost every set of $2n + 2m + 1$ or $2n + 2m + 2$ vectors can be $m + 1$-lifted.

1.1 The Phase Property

In this section we investigate the relationship between phase retrieval and recovery of the phase of signal in the finite dimensional case. Let $x = (a_1, a_2, \ldots, a_n)$ and

$y = (b_1, b_2, \ldots, b_n)$ be vectors in $\mathbb{H}^n$. We say that x, y have the *same phase* if there exists $|\theta| = 1$

$$\arg a_i = \theta \arg b_i, \text{ for all } i = 1, 2, \ldots, n.$$

We begin this section with a simple definition.

Definition 1.13. Let $\mathcal{X} = \{x_k\}_{k=1}^m$ be a family of vectors in $\mathbb{H}^n$ such that for every non-zero x and y

$$|\langle x, x_k\rangle|^2 = |\langle y, x_k\rangle|^2 \quad \text{for all } k = 1, 2, \ldots m.$$

If this implies x and y have the same phases, we say $\mathcal{X}$ has the **phase property**.

Definition 1.14. A family of orthogonal projections $\{P_k\}_{k=1}^m$ on $\mathbb{H}^n$ satisfies the **phase property** if for every non-zero x and y

$$\|P_k x\|^2 = \|P_k y\|^2, \text{ for all } k \in [m]$$

imples x and y have the same phase.

We will prove that phase retrieval implies the complement property for both the real and complex cases and even in a more general setting. First, observe a few consequences of this definiton. If $x = (a_1, a_2, \ldots, a_n)$ and $y = (b_1, b_2, \ldots, b_n)$ have the same phases, then $a_i = 0$ if and only if $b_i = 0$ (i.e. zero has no phase). Note that if $\{x_k\}_{k=1}^m$ has the phase property in $\mathbb{H}^n$, then $\text{span}\{x_k\}_{k=1}^m = \mathbb{H}^n$. For otherwise, there would exist $0 \neq x \in \mathbb{H}^n$ so that

$$\langle x, x_k\rangle = \langle 0, x_k\rangle = 0, \text{ for all } i = 1, 2, \ldots, m,$$

while x and 0 do not have the same phase.

Over the years, the term phase retrieval has been used to replace the original nomenclature of *phaseless reconstruction*. At a meeting in 2012, Casazza asked if this shift in terminology was accurate. That is, are the phase property and phase retrieval really the same? In this section we will answer this question in the affirmative.

The problem occurred here because of the way we translated the engineering version of *phase retrieval* into the language of frame theory. The engineers are working with the modulus of the Fourier transform and want to recover the phases so they can invert the Fourier transform to discover the signal. So all they need to do is to recover the phase. But in the frame theory version of this, for $x = (a_1, a_2, \ldots, a_n)$ we are really trying to recover two things:

1. Recover the phases of the a_i.

2. Recover $|a_i|$ (which in the engineering case, is already known).

Theorem 1.15. *Let $\{P_i\}_{i=1}^m$ be projections onto the subspaces $\{W_i\}_{i=1}^m$ of $\mathbb{H}^n$ which have the phase property, then for every orthonormal basis $\{\phi_{i,j}\}_{j=1}^{D_i}$ of W_i, the set $\{\phi_{i,j}\}_{i=1,j=1}^{m,\ D_i}$ has complement property.*

Proof. Suppose $\{W_i\}_{i=1}^m$ satisfy the phase property, but fail phase retrieval. By Theorem 1.5, there exist an orthonormal basis $\{\phi_{i,j}\}_{j=1}^{D_i}$ of each W_i such that the set $\{\phi_{i,j}\}_{i=1,j=1}^{m,\ D_i}$ fails the complement property. In other words, there exists $I \subset \{(i,j) : 1 \leq i \leq m \text{ and} 1 \leq j \leq D_i\}$ so that $\{\phi_{i,j}\}_{(i,j)\in I}$ and $\{\phi_{i,j}\}_{(i,j)\in I^c}$ do not span $\mathbb{H}^n$. Choose vectors $x, y \in \mathbb{H}^n$ with $\|x\| = 1 = \|y\|$, and $x = (a_1, a_2, \ldots, a_n)$ and $y = (b_1, b_2, \ldots, b_n)$ such that $x \perp \phi_{i,j}$ for all $(i,j) \in I$ and $y \perp \phi_{i,j}$ for all $(i,j) \in I^c$. Note this choice of vectors forces that for each (i,j) either $\langle x, \phi_{i,j}\rangle = 0$ or $\langle y, \phi_{i,j}\rangle = 0$.

Fix $0 \neq c$. Then for each $1 \leq i \leq m$

$$|\langle x + cy, \phi_{ij} \rangle| = |\langle x - cy, \phi_{ij} \rangle|, \text{ for all } i, j.$$

Hence,

$$\|P_i(x + cy)\|^2 = \sum_{j=1}^{D_i} |\langle x + cy, \phi_{i,j} \rangle|^2 = \sum_{j=1}^{D_i} |\langle x - cy, \phi_{i,j} \rangle|^2 = \|P_i(x - cy)\|^2.$$

By assumption that $\{P_i\}_{i=1}^m$ does phase retrieval, this implies there is a $|\theta| = 1$ so that $x + cy$ and $\theta(x - cy)$ have the same phases. Assume there exists some $1 \leq i_0 \leq n$ so that $a_{i_0} \neq 0 \neq b_{i_0}$ and let $c = \frac{-a_{i_0}}{b_{i_0}}$. Then

$$(x + cy)_{i_0} = a_{i_0} + cb_{i_0} = a_{i_0} + \frac{-a_{i_0}}{b_{i_0}} b_{i_0} = 0,$$

while

$$(x - cy)_{i_0} = a_{i_0} - \frac{-a_{i_0}}{b_{i_0}} = 2a_{i_0} \neq 0.$$

But this contradicts the observation that zero has a unique phase. It follows that for every $1 \leq i \leq n$, either $a_i = 0$ or $b_i = 0$. Let $\{e_i\}_{i=1}^n$ be an orthonormal basis for $\mathbb{H}^n$ and let $I = \{1 \leq i \leq n : b_i = 0\}$. Then

$$x + y = \sum_{i \in I} a_i e_i + \sum_{i \in I^c} b_i e_i, \text{ and } x - y = \sum_{i \in I} a_i e_i + \sum_{i \in I^c} (-b_i) e_i.$$

By the above argument, $x + y$ and $(x - y)$ have this same phase, but this is a contradiction. ∎

We have a number of consequences of Theorem 1.15. Letting the subspaces W_i be one dimensional, this becomes a theorem about vectors.

Corollary 1.16. *If $\mathcal{X} = \{x_i\}_{i=1}^m$ does phase retrieval in $\mathbb{H}^n$, then $\mathcal{X}$ has the complement property. Hence, in the real case, the phase property and phase retrieval are equivalent properties.*

Combining Theorems 1.15, 1.5 we get the following corollary.

Corollary 1.17. *In $\mathbb{R}^n$, a family of projections $\{P_i\}_{i=1}^m$ has the phase property if and only if it does phase retrieval.*

In the complex case, the complement property is not equivalent to phase retrieval. However, we can show that the phase property and phase retrieval are equivalent in the complex case by using the following criteria:

Theorem 1.18 ([4])**.** *Consider $\mathcal{X} = \{x_k\}_{k=1}^m \subseteq \mathbb{C}^n$ and the mapping $\mathcal{A} : \mathbb{C}^n/\mathbb{T} \to \mathbb{R}^m$ defined by $(\mathcal{A}(x))(k) := |\langle x, x_k\rangle|^2$. Viewing $\{x_k x_k^* u\}_{k=1}^m$ as vectors in $\mathbb{R}^{2n}$, denote $S(u) := \operatorname{span}_{\mathbb{R}}\{x_k x_k^* u\}_{k=1}^m$. Then the following are equivalent:*

(a) $\mathcal{A}$ is injective.

(b) $\dim S(u) \geq 2n-1$ for every $u \in \mathbb{C}^n \setminus \{0\}$.

(c) $S(u) = \operatorname{span}_{\mathbb{R}}\{\iota u\}^{\perp}$ for every $u \in \mathbb{C}^n \setminus \{0\}$.

To prove the desired result we first need a few lemmas. For this section we adopt the notation $\langle a, b\rangle_{\mathbb{R}}$ to denote $\operatorname{Re}\langle a, b\rangle$.

Lemma 1.19. *Given $\{x_k\}_{k=1}^m \subseteq \mathbb{C}^n$ and any $u \in \mathbb{C}^n$ then $\langle x_k x_k^* u, \iota u\rangle_{\mathbb{R}} = 0$*

Proof. The following calculation gives the result almost immediately:

$$\begin{aligned}\langle x_k x_k^* u, \iota u\rangle_{\mathbb{R}} =& \langle \langle u, x_k\rangle x_k, \iota u\rangle_{\mathbb{R}} = \operatorname{Re}(-\iota\langle u, x_k\rangle\langle x_k, u\rangle) \\ =& -\operatorname{Re}(\iota|\langle u, x_k\rangle|^2) = 0.\end{aligned}$$

∎

Lemma 1.20. *Given $\{x_k\}_{k=1}^m \subseteq \mathbb{C}^n$ and any $u, v \in \mathbb{C}^n$ then for each x_k,*

$$|\langle u+v, x_k\rangle|^2 - |\langle u-v, x_k\rangle|^2 = 4\langle x_k x_k^* u, v\rangle_{\mathbb{R}}.$$

Proof. Consider the following

$$|\langle u+v, x_k\rangle|^2 = |\langle u, x_k\rangle|^2 + 2\,\mathrm{Re}(\langle u, x_k\rangle\overline{\langle v, x_k\rangle}) + |\langle v, x_k\rangle|^2 \tag{1.7}$$

and

$$|\langle u-v, x_k\rangle|^2 = |\langle u, x_k\rangle|^2 - 2\,\mathrm{Re}(\langle u, x_k\rangle\overline{\langle v, x_k\rangle}) + |\langle v, x_k\rangle|^2. \tag{1.8}$$

Then subtracting (1.8) from (1.7) we obtain

$$|\langle u+v, x_k\rangle|^2 - |\langle u-v, x_k\rangle|^2 = 4\,\mathrm{Re}(\langle u, x_k\rangle\overline{\langle v, x_k\rangle}) = 4\langle x_k x_k^* u, v\rangle_{\mathbb{R}}$$

∎

Corollary 1.21. *If $\{x_k\}_{k=1}^m$ does phase retrieval and $\langle x_k x_k^* u, v\rangle_{\mathbb{R}} = 0$ for each k then $u+v = \omega(u-v)$ for $|\omega| = 1$ and thus $v = \frac{2\,\mathrm{Im}(\omega)}{|1+\omega|^2}u$.*

Proof. If $u+v = \omega u - \omega v$ then $v = \frac{\omega-1}{\omega+1}u = -\frac{(1-\omega)(1+\overline{\omega})}{|1+\omega|^2}u = \frac{2\,\mathrm{Im}(\omega)}{|1+\omega|^2}u$. ∎

Lemma 1.22. *Given any u, let $v = \alpha\iota u$ for $\alpha \in \mathbb{R}$ and let $\omega = \frac{1+\alpha\iota}{1-\alpha\iota}$ then $|\omega| = 1$ and $u+v = u(1+\alpha\iota) = \frac{1+\alpha\iota}{1-\alpha\iota}(u-\alpha\iota u) = \omega(u-v)$.*

Lemma 1.23. *If $x - y \neq 0$ then $\langle \phi\phi^*(x-y), x+y\rangle_{\mathbb{R}} = 0$.*

Proof. Consider the following calculation,

$$\begin{aligned}\langle \phi\phi^*(x-y), x+y\rangle_{\mathbb{R}} &= \mathrm{Re}((x+y)^*\phi\phi^*(x-y)) \\ &= \mathrm{Re}(|\phi^* x|^2 - x^*\phi\phi^* y + y^*\phi\phi^* x - |\phi y|^2) \\ &= \mathrm{Re}(-x^*\phi\phi^* y + x^*\phi\phi^* y) = 0.\end{aligned}$$

∎

Lemma 1.24. *Let $a, b \in \mathbb{C}$ such that $|a| + |b| > 0$. If*

$$\arg(a+b) = \arg(e^{\iota\theta}(a-b)),$$

then

$$\tan\theta = \frac{2\,\mathrm{Im}(\bar{a}b)}{|a|^2 - |b|^2}$$

for $|a| \neq |b|$ and $\theta = \pi/2$ otherwise.

Theorem 1.25. *The phase property implies phase retrieval in the complex case.*

Proof. Suppose $\mathcal{X} = \{x_k\}_{k=1}^m \subseteq \mathbb{C}^n$ does phase retrieval. Let u, v be non-zero vectors in $\mathbb{C}^n$ such that $\langle x_k x_k^* u, v\rangle_{\mathbb{R}} = 0$ for all k. Note that Lemma 1.20 ensures that $|\langle u+v, x_k\rangle|^2 = |\langle u-v, x_k\rangle|^2$ for each k. To apply the results in Theorem 1.18, we must show $v = \lambda\iota u$ for some $\lambda \in \mathbb{R}$. For simplicity, denote $u = (u_1, u_2, ...)$ and $v = (v_1, v_2, ...)$. Consider the following cases:

Case 1: $u_j v_j = 0$ for all $1 \leq j \leq n$.

Without loss of generality, suppose $u = (e^{\iota\alpha_1}, 0, ...)$ and $v = (0, e^{\iota\beta_2}, ...)$ for some $\alpha_1, \beta_1 \in \mathbb{R}$. Since χ has the phase property, we have that $u+v$ has the same phase as $e^{\iota\gamma}(u-v)$, with some real constant γ. In particular $\arg(u_1+v_1) = \arg(e^{\iota\gamma}(u_1-v_1))$, i.e. $\arg(e^{\iota\alpha_1}) = \arg(e^{\iota\gamma}e^{\iota\alpha_1})$. Similarly $\arg(u_2+v_2) = \arg(e^{\iota\gamma}(u_2-v_2))$, i.e. $\arg(e^{\iota\beta_2}) = \arg(-e^{\iota\gamma}e^{\iota\beta_2})$. However the first condition implies $\gamma = 0$ and the second gives $\gamma = \pi$, a contradiction.

Case 1: $u_j v_j \neq 0$ for some $1 \leq j \leq n$.

Without loss of generality, we can assume $u_1 v_1 \neq 0$ and by multiplying by the appropriate constants we may also assume $|u_1| = |v_1| = r_1 > 0$. Then by Lemma

1.24, for each $1 \leq j \leq n$ we have that

$$\tan(\gamma) = \frac{2\,\mathrm{Im}(\overline{u_j}v_j)}{|u_j|^2 - |v_j|^2}.$$

By assumption $|u_1| = |v_1|$, therefore $\gamma = \pi/2$ and hence $|u_j| = |v_j|$ for all $1 \leq j \leq n$. So we have shown that

$$u = (r_1 e^{\iota\alpha_1}, r_2 e^{\iota\alpha_2}, \ldots, r_n e^{\iota\alpha_n}) \text{ and } v = (r_1 e^{\iota\beta_1}, r_2 e^{\iota\beta_2}, \ldots, r_n e^{\iota\beta_n}).$$

Now we claim that $\sin(\beta_j - \alpha_j) = c$ for all j. To see this note that since $\arg(2u_j + v_j) = \arg(e^{\iota\theta}(2u_j - v_j))$ for all j and fixed θ, then by Lemma 1.24 we see that

$$c = \tan\theta = \frac{4\,\mathrm{Im}(\overline{u_j}v_j)}{3r_j^2} = \frac{4}{3}\sin(\beta_j - \alpha_j) \quad \forall 1 \leq j \leq n.$$

For each j, set $a_j = \cos(\beta_j - \alpha_j) = \pm\sqrt{1 - c^2}$. We can express $v = w + c\iota u$ where

$$w = (a_1 r_1 e^{\iota\alpha_1}, a_2 r_2 e^{\iota\alpha_2}, \ldots, a_n r_n e^{\iota\alpha_n}).$$

Now we rewrite

$$v = \left(r_1 e^{\iota\alpha_1} e^{\iota(\beta_1 - \alpha_1)}, r_2 e^{\iota\alpha_2} e^{\iota(\beta_2 - \alpha_2)}, \ldots, r_n e^{\iota\alpha_n} e^{\iota(\beta_n - \alpha_n)}\right)$$

and each $e^{\iota(\beta_j - \alpha_j)} = \cos(\beta_j - \alpha_j) + \iota\sin(\beta_j - \alpha_j) = a_j + \iota c$. We must show $w = 0$. Recall that for every k we have

$$0 = \langle x_k x_k^* u, w + c\iota u\rangle_{\mathbb{R}} = \langle x_k x_k^* u, w\rangle_{\mathbb{R}} + \langle x_k x_k^* u, c\iota u\rangle_{\mathbb{R}}.$$

By Lemma 1.19 we see that $\langle x_k x_k^* u, w\rangle_{\mathbb{R}} = 0$ for all k. Note that $w = 0$ if and only if $a_j = 0$ for all j. This is clear since that if $a_1 \neq 0$ then the first component of $a_1 u + w$ is non-zero but the the first component of $a_1 u - w$ is 0 (assuming $u_1 \neq 0$) which contradicts to $w = 0$.

■

1.1.1 Weak Phase Property

While investigating the relationship between phase retrieval and the phase property, it was noted that if two vectors have the same phase then they will be zero in the same coordinates. This gave way to a weakening of phase retrieval, known as weak phase retrieval. In this work, we study the weakened notions of the phase property and phase retrieval. One limitation of current methods used for retrieving the phase of a signal is computing power. Recall that a generic family of $(2m-1)$-vectors in $\mathbb{R}^n$ does phase retrieval, however no set of $(2n-2)$-vectors can (See [3] for details). By generic we are referring to an open dense set in the set of $(2n-1)$-element frames in $\mathbb{H}^m$. We started with the motivation that weak phase retrieval could be done with $n+1$ vectors in $\mathbb{R}^n$. However, it will be shown that the cardinality condition can only be relaxed to $2n-2$. Nevertheless, the results we obtain in this work are interesting in their own right and contribute to the overall understanding of phase retrieval. We provide illustrative examples in the real and complex cases for weak phase retrieval and weak phase property.

In this section we will make a detailed study of the weakened phase property. First we define the weak phase property and obtain the minimum number of vectors required.

Definition 1.26. Two vectors in $\mathbb{H}^n$, $x = (a_1, a_2, \ldots, a_n)$ and $y = (b_1, b_2, \ldots, b_n)$ **weakly have the same phase** if there is a $|\theta| = 1$ so that

$$\arg(a_i) = \theta \arg(b_i), \text{ for all } i = 1, 2, \ldots, n, \text{ for which } a_i b_i \neq 0.$$

In the real case, if $\theta = 1$ we say x, y **weakly have the same signs** and if $\theta = -1$ they **weakly have opposite signs.**

In the definition above note that we are only comparing the phase of x and y for entries where both are nonzero. Hence, two vectors may *weakly* have the same phase but not have the same phase in the usual sense. We define weak phase retrieval formally as follows:

Definition 1.27. A family of vectors $\{x_k\}_{k=1}^m$ in $\mathbb{H}^n$ have the **weak phase property** if for any $x = (a_1, a_2, \ldots, a_n)$ and $y = (b_1, b_2, \ldots, b_n)$ in $\mathbb{H}^n$, with

$$|\langle x, x_k\rangle|^2 = |\langle y, x_k\rangle|^2, \text{ for all } k = 1, 2, \ldots, m,$$

then x, y weakly have the same phase.

Observe that the difference between the phase property and the weak phase phase property is that in the later it is possible for $a_i = 0$ but $b_i \neq 0$. Now we begin our study of the weak phase property in $\mathbb{R}^n$. The following proposition provides a useful criteria for determining when two vectors have weakly the same or opposite phases.

Proposition 1.28. *Let $x = (a_1, a_2, \ldots, a_n)$ and $y = (b_1, b_2, \ldots, b_n)$ in $\mathbb{R}^n$. The following are equivalent:*

1. *We have*

$$sgn\ (a_i a_j) = sgn\ (b_i b_j),\ \textit{for all } a_i a_j \neq 0 \neq b_i b_j.$$

2. *Either x, y have weakly the same signs or they have weakly opposite signs.*

Proof. (1) $\Rightarrow$ (2): Let

$$I = \{1 \leq i \leq n : a_i = 0\} \text{ and } J = \{1 \leq i \leq n : b_i = 0\}.$$

Let $K = [n] \setminus (I \cup J)$. So $i \in K$ if and only if $a_i \neq 0 \neq b_i$. Let $i_0 = min\ K$. We examine two cases:

Case 1: $sgn\ a_{i_0} = sgn\ b_{i_0}$.

For any $i_0 \neq k \in K$, $sgn\ (a_{i_0}a_k) = sgn\ (b_{i_0}b_k)$, implies $sgn\ a_k = sgn\ b_k$. Since all other coordinates of either x or y are zero, it follows that x, y weakly have the same signs.

Case 2: $sgn\ a_{i_0} = -sgn\ b_{i_0}$.

For any $i_o \neq k \in K$, $a_{i_0}a_k = b_{i_0}b_k$ implies $sgn\ a_k = -sgn\ b_k$. Again, since all other coordinates of either x or y are zero, it follows that x, y weakly have opposite signs.

(2) $\Rightarrow$ (1): This is immediate. ∎

The next lemma will be useful in the following proofs as it gives a criteria for showing when vectors do not weakly have the same phase.

Lemma 1.29. *Let $x = (a_1, a_2, ..., a_n)$ and $y = (b_1, b_2, ..., b_n)$ be vectors in $\mathbb{R}^n$. If there exists $i \in [n]$ such that $a_i b_i \neq 0$ and $\langle x, y\rangle = 0$, then x and y do not have weakly the same or opposite signs.*

Proof. We proceed by way of contradiction. If x and y weakly have the same phase then $a_j b_j \geq 0$ for all $j \in [n]$ and in particular we arrive at the following contradiction

$$\langle x, y\rangle = \sum_{j=1}^{n} a_j b_j \geq a_i b_i > 0.$$

If x and y weakly have opposite phases then $a_j b_j \leq 0$ for all $j \in [n]$ and by reversing the inequalities in the expression above we get the desired result. ∎

The following result relates the weak phase property and the original phase property. Recall that in the real case, it is known that the phase property, phase retrieval and the complement property are equivalent [3, 10].

Corollary 1.30. *Suppose $\mathcal{X} = \{x_k\}_{k=1}^m \subset \mathbb{R}^n$ has the weak phase property, but fails complement property. Then there exists two vectors $v, w \in \mathbb{R}^n$ such that $v \perp w$ and*

$$|\langle v, x_k\rangle| = |\langle w, x_k\rangle| \text{ for all } k. \tag{1.9}$$

Further, v and w are disjointly supported.

Proof. By the assumption, $\mathcal{X} = \{x_k\}_{k=1}^m$ fails complement property so there exists $S \subset [m]$, s.t. $A = \text{span}\{x_k\}_{k\in S} \neq \mathbb{R}^n$ and $B = \text{span}\{x_k\}_{k\in S^c} \neq \mathbb{R}^n$. Choose $\|x\| = \|y\| = 1$ such that $x \perp A$ and $y \perp B$. Then

$$|\langle x + y, x_k\rangle| = |\langle x - y, x_k\rangle| \text{ for all i=1, 2, } \ldots, \text{ m.}$$

Let $w = x + y$ and $v = x - y$. Then $u \perp v$. Observe

$$\langle w, v\rangle = \langle x + y, x - y\rangle = \|x\|^2 + \langle y, x\rangle - \langle x, y\rangle - \|y\|^2 = 0.$$

Moreover, the assumption that $\mathcal{X}$ has the weak phase property implies u and w have weakly the same or opposite phases. Then it follows from Lemma 1.29 that $u_i w_i = 0$ for all $i = 1, 2, \ldots, n$ and so u and w are disjointly supported. ∎

Example 1.31. In $\mathbb{R}^2$, let $x_1 = (1, 1)$ and and $x_2 = (1, -1)$. These vectors clearly fail complement property. But if $x = (a_1, a_2)$, $b = (b_1, b_2)$ and we have,

$$|\langle x, x_k\rangle| = |\langle y, x_k\rangle|, \text{ for } k = 1, 2,$$

then

$$|a_1 + a_2|^2 = |b_1 + b_2|^2 \text{ and } |a_1 - a_2|^2 = |b_1 - b_2|^2.$$

By squaring these out and subtracting the result we get $4a_1a_2 = 4b_1b_2$. Hence, either x, y have the same signs or opposite signs. These vectors have the weak phase property.

With some particular assumptions, the following proposition determines a constraint on vectors which have the weak phase property, but not phase retrieval.

Proposition 1.32. *Let $\mathcal{X} = \{x_k\}_{i=1}^m \in \mathbb{R}^n$ such that $\mathcal{X}$ has the weak phase property, but fails complement property. Let $x = (a_1, a_2, \ldots, a_n)$, $y = (b_1, b_2, \ldots b_n) \in \mathbb{R}^n$ such that $x + y \perp x - y$ and satisfy equation 1.9. If $a_i b_i \neq 0$, $a_j b_j \neq 0$ for some i, j and all other co-ordinates of x and y are zero, then*

$$|x_{kj}| = c\,|x_{ki}|\,;\ \textit{for some } c > 0.$$

where $x_k = (x_{k1}, x_{k2}, \ldots, x_{kn})$.

Proof. Without loss of generality, take $x = (a_1, a_2, 0, \ldots, 0)$ and $y = (b_1, b_2,$ $0, \ldots, 0)$. Observe that both $x + y$ and $x - y$ either weakly have the same phase or weakly have the opposite phase. Thus, by Lemma 1.29, $x + y$ and $x - y$ have disjoint support as these vectors are orthogonal. Thus it reduces to the cases where either $a_1 = b_1,\ a_2 = -b_2$ or $a_1 = -b_1,\ a_2 = b_2$. In both cases, it follows from equation 1.9 that $|x_{k2}| = c\,|x_{k1}|$ where $c = \left|\frac{-b_1}{b_2}\right| > 0$. ∎

The next theorem gives the minimum number of vectors necessary to satisfy the weak phase property in $\mathbb{R}^n$. Recall that phase retrieval (or the phase property) requires $m \geq 2n - 1$ vectors.

Theorem 1.33. *If $\{x_k\}_{i=1}^m$ in $\mathbb{R}^n$ satisfy the weak phase property, then $m \geq 2n - 2$.*

Proof. For a contradiction assume $m \leq 2n - 3$ and choose $S \subset [m]$ with $S = [n - 2]$. Then $|S| = n - 2$ and $|S^c| \leq n - 1$. For this partition of $[m]$, let $x + y$ and $x - y$ be as in Corollary 1.30. Then $x + y$ and $x - y$ must be disjointly supported and therefore

for each i we have $a_i = \epsilon_i b_i$, where $\epsilon_i = \pm 1$ for each i. Observe the conclusion holds for a fixed x and any $y \in (\text{span}\{x_k\}_{k\in S})^\perp$ and dim $(\text{span}\{x_k\}_{k\in S})^\perp \geq 2$. However this poses a contradiction since there are infinitely many distinct choices of y in this space, while our argument shows that there are at most 2^n choices of y. ∎

Contrary to the initial hopes, the previous result shows that the minimal number of vectors needed to satisfy the weak phase property is only one less than the number of vectors doing phase retrieval. However, it is interesting to note that a minimal set of vectors satisfying the weak phase property is necessarily full spark, as is true for the minimal number of vectors doing phase retrieval.

Theorem 1.34. *If $\mathcal{X} = \{x_k\}_{k=1}^{2n-2}$ satisfies the weak phase property in $\mathbb{R}^n$, then $\mathcal{X}$ is full spark.*

Proof. We proceed by way of contradiction. Assume $\mathcal{X}$ is not full spark. Then there exists $S \subset \{1, 2, ..., 2n-2\}$ with $|S| = n$ such that $\dim \text{span}\{x_k\}_{k\in S} \leq n-1$. Observe that the choice of S above implies $|S^c| = n - 2$. Now we arrive at a contradiction by applying the same argument used in (the proof of) Theorem 1.33. ∎

It is important to note that the converse of Theorem 1.34 does not hold. For example, the canonical basis in $\mathbb{R}^2$ is trivially full spark but does not have the weak phase property.

If χ is as in Theorem 1.34, then it is possible to add a vector to this set and obtain a collection which does phase retrieval. However, the following corollary provides a slightly stronger result.

Corollary 1.35. *If $\mathcal{X}$ is as in Theorem* 1.34*, then by the construction of full spark*

there exists a dense set of vectors F *in* $\mathbb{R}^n$ *such that* $\{\psi\} \cup \mathcal{X}$ *does phase retrieval for any* $\psi \in \mathsf{F}$.

Proof. We observe that the set of $\psi \in \mathbb{R}^n$ such that $\mathcal{X} \cup \{\psi\}$ is full spark is dense in $\mathbb{R}^n$. To see this let $\mathsf{G} = \bigcup_{\substack{I \subset [2n-2] \\ |I|=n-1}} \operatorname{span}\{x_k\}_{k\in I}$. Then G is the finite union of hyperplanes so G^c is dense and $\{\psi\} \cup \mathcal{X}$ is full spark for any $\psi \in \mathsf{G}^c$. To verify that this collection of vectors is full spark. Note that either a sub-collection of m-vectors is contained in $\mathcal{X}$, then it spans $\mathbb{R}^n$, or the subcollection contains the vector ψ. In this case, denote $I \subset [2n-2]$ with $|I| = n-1$ and suppose $\sum_{k\in I} a_k x_k + a\psi = 0$. Therefore $a\psi = -\sum_{k\in I} a_k x_k$ and if $a \neq 0$ then $a\psi \in \operatorname{span}\{x_k\}_{k\in I}$, a contradiction. It follows $a = 0$ and since $\mathcal{X}$ is full spark (see Theorem 1.34), in particular $\{x_k\}_{k\in I}$ are linearly independent, it follows that $a_k = 0$ for all $k \in I$. ∎

1.1.2 Weak Phase Retrieval

In this section, we define weak phase retrieval and show that in the real case it is equivalent to phase retrieval. A formal definition is given below:

Definition 1.36. A family of vectors $\mathcal{X} = \{x_k\}_{k=1}^m$ in $\mathbb{H}^n$ does **weak phase retrieval** if for any $x = (a_1, a_2, \ldots, a_n)$ and $y = (b_1, b_2, \ldots, b_n)$ in $\mathbb{H}^n$, with

$$|\langle x, x_k\rangle|^2 = |\langle y, x_k\rangle|^2, \text{ for all } k = 1, 2, \ldots, m, \tag{1.10}$$

there is a $|\theta| = 1$ so that

$$a_i = \theta b_i, \text{ for all } i = 1, 2, \ldots, n, \text{ for which } a_i \neq 0 \neq b_i.$$

In particular, $\mathcal{X}$ does phase retrieval for vectors having all non-zero coordinates. Clearly if $\mathcal{X} \subset \mathbb{R}^n$ does weak phase retrieval, then it has the weak phase property.

The converse is not true in general. Let $x = (a_1, a_1, ..., a_m)$ and $y = (b_1, b_2, ..., b_n)$. If $\mathcal{X} = \{x_k\}_{k=1}^m \subset \mathbb{R}^n$ has the weak phase property and $|\{i|a_i b_i \neq 0\}| = 2$ then $\mathcal{X}$ may not do weak phase retrieval. If $a_i a_j = b_i b_j$ where $a_i b_i \neq 0$ then we certainly cannot conclude in general that $|a_i| = |b_i|$ (see Example 1.40). The following theorems provide conditions under which the weak phase property is equivalent to weak phase retrieval.

Proposition 1.37. *Suppose $\mathcal{X} = \{x_k\}_{k=1}^m$ has the weak phase property on vectors $x = (a_1, a_2, ..., a_n)$ and $y = (b_1, b_2, ..., b_n)$ in $\mathbb{H}^n$. If $|I| = |\{i : a_i b_i \neq 0\}| \geq 3$ and $a_i a_j = b_i b_j$ for all $i, j \in I$, then $\mathcal{X}$ does weak phase retrieval on these vectors.*

Proof. If i, j, k are members of I such that $a_i a_j = b_i b_j$, $a_i a_k = b_i b_k$ and $a_k a_j = b_k b_j$, then a short calculation gives $a_i^2 a_j a_k = b_i^2 b_j b_k$ and hence $|a_i| = |b_i|$. This computation holds for each $i \in I$ and since χ does has the weak phase property, there is a $|\theta| = 1$ so that $\arg a_i = \theta \arg b_i$ for all i. It follows that $a_i = \theta\ b_i$ for all i such that $a_i b_i \neq 0$. ∎

It turns out that whenever a frame contains the unit basis vectors, then the weak phase property and phase retrieval are the equivalent, as shown in the following theorem.

Proposition 1.38. *Suppose $\mathcal{X} = \{x_k\}_{k=1}^m \subset \mathbb{R}^n$ has the weak phase property. If $\mathcal{X}$ contains the standard basis vectors, then χ does phase retrieval.*

Proof. Let $x = (a_1, a_2, \ldots, a_n)$, $y = (b_1, b_2, \ldots, b_n) \in \mathbb{R}^m$. If $\mathcal{X}$ satisfies the equation 1.10, then (for basis vectors) the equation 1.10 implies that $|a_i| = |b_i|$, $\forall i = 1, 2, \ldots, n$. ∎

We conclude this section by showing the surprising result that weak phase retrieval is equivalent to phase retrieval in $\mathbb{R}^n$. In other words, it was never really weak.

Theorem 1.39. *Frames in $\mathbb{R}^n$ which do weak phase retrieval do phase retrieval.*

Proof. For a contradiction assume $\mathcal{X} = \{x_k\}_{k=1}^m \subset \mathbb{R}^n$ does weak phase rectrieval but fails the complement property. Then there exists $S \subset [m]$ such that $\text{span}_{k\in S}\ x_k \neq \mathbb{R}^n$ and $\text{span}_{k\in S^c}\ x_k \neq \mathbb{R}^n$. Pick non-zero vectors $x, y \in \mathbb{R}^n$ such that $x \perp \text{span}_{k\in S}\ x_k \neq \mathbb{R}^n$ and $\text{span}_{k\in S^c}\ x_k \neq \mathbb{R}^n$. Then for any $c \neq 0$ we have

$$|\langle x + cy, x_k\rangle| = |\langle x - cy, x_k\rangle| \quad \text{for all } k \in [m].$$

Now we consider the following cases where a_i and b_i denotes the i^{th} coordinate of the vectors x and y.

Case 1: $\{i : a_i \neq 0\} \cap \{i : b_i \neq 0\} = \emptyset$

Set $c = 1$ and observe since $x \neq 0$ there exists some $i \in [n]$ such that $a_i \neq 0$ and $b_i = 0$ and similarly there exists $j \in [n]$ such that $b_j \neq 0$ but $a_j = 0$. Then $x+y$ and $x - y$ have the same sign in the i^{th}-coordinate but opposite signs in the j^{th} coordinate, this contradicts the assumption that $\mathcal{X}$ does weak phase retrieval.

Case 2: There exists $i, j \subset [n]$ such that $a_i b_i \neq 0$ and $a_j = 0$, $b_j \neq 0$.

Without loss of generality, we may assume $a_i b_i > 0$ otherwise consider $-x$ or $-y$. If $0 < c \leq \frac{a_i}{b_i}$, then the i^{th} coordinate of $x + cy$ and $x - cy$ have the same sign whereas the j^{th} coordinates have opposite signs which contradicts the assumption. By considering $y + cx$ and $y - cx$ this argument holds in the case that $b_j = 0$ and $a_j \neq 0$.

Case 3: $a_i = 0$ if and only if $b_i = 0$.

By choosing c small enough, we have that $a_i + cb_i \neq 0$ if and only if $a_i - cb_i \neq 0$. By weak phase retrieval, we have $a_i + cb_i = \pm(x_i - cy_i)$. This forces either $a_i \neq 0$ or $b_i \neq 0$ (but not both), which contradicts the assumption for case 3.

■

It is known [1] that if $\mathcal{X} = \{x_k\}_{k=1}^m \subset \mathbb{R}^n$ has the phase property or does phase retrieval in $\mathbb{H}^n$ and T is an invertible operator on $\mathbb{H}^n$ then $\{Tx_k\}_{k=1}^m$ does phase retrieval. It follows that the same result holds for weak phase retrieval. However, this result does not hold for the weak phase property. Indeed, if $x_1 = (1, 1)$ and $x_2 = (1, -1)$, then we have seen that this frame does weak phase retrieval in $\mathbb{R}^2$. But the invertible operator $T(x_1) = (1, 0),\ T(x_2) = (0, 1)$ maps this frame to a frame which fails weak phase retrieval.

1.2 Illustrative Examples

In this section, we provide examples of frames that have the weak phase property in $\mathbb{R}^3$ and $\mathbb{R}^4$.

Our first example is a frame which has the weak phase property, but fails weak phase retrieval (as we have seen in $\mathbb{R}^2$).

Example 1.40. We work with the row vectors of

$$\mathcal{X} = \left[\begin{array}{c|ccc} x_1 & 1 & 1 & 1 \\ x_2 & -1 & 1 & 1 \\ x_3 & 1 & -1 & 1 \\ x_4 & 1 & 1 & -1 \end{array}\right]$$

Observe that the rows of this matrix form an equal norm tight frame $\mathcal{X}$. Recall that a tight frame must do norm retrieval since the frame operator is a constant multiple of the identity. That is, $A\|x\|^2 = \sum_{k=1}^{4} |\langle x, x_k\rangle|^2$. If $x = (a_1, a_2, a_3)$ the following is the coefficient matrix where the row E_i represents the coefficients obtained from the expansion $|\langle x, x_k\rangle|^2$

$$1/2 \left[\begin{array}{c|cccc} & a_1a_2 & a_1a_3 & a_2a_3 & \sum_{i=1}^{3} a_i^2 \\ E_1 & 1 & 1 & 1 & 1/2 \\ E_2 & -1 & -1 & 1 & 1/2 \\ E_3 & -1 & 1 & -1 & 1/2 \\ E_4 & 1 & -1 & -1 & 1/2 \end{array}\right]$$

Then the following row operations give

$$1/2 \left[\begin{array}{c|cccc} & a_1a_2 & a_1a_3 & a_2a_3 & \sum_{i=1}^{3} a_i^2 \\ F_1 = E_1 - E_2 & 1 & 1 & 0 & 0 \\ F_2 = E_3 - E_4 & -1 & 1 & 0 & 0 \\ F_3 = E_1 - E_3 & 1 & 0 & 1 & 0 \\ F_4 = E_2 - E_4 & -1 & 0 & 1 & 0 \\ F_4 = E_1 - E_4 & 0 & 1 & 1 & 0 \\ F_5 = E_2 - E_3 & 0 & -1 & 1 & 0 \end{array}\right]$$

$$1/2 \left[\begin{array}{c|cccc} & a_1a_2 & a_1a_3 & a_2a_3 & \sum_{i=1}^{3} a_i^2 \\ F_1 - F_2 & 1 & 0 & 0 & 0 \\ F_3 + F_4 & 0 & 0 & 1 & 0 \\ F_5 - F_6 & 0 & 1 & 0 & 0 \end{array}\right]$$

Therefore we have demonstrated a procedure to identify a_ia_j for all $1 \leq i \neq j \leq 3$. This shows that given $y = (b_1, b_2, b_3)$ satisfying $|\langle x, x_k\rangle|^2 = |\langle y, x_k\rangle|^2$ then by the procedure outlined above we obtain

$$a_ia_j = b_ib_j, \text{ for all } 1 \leq i \neq j \leq 3.$$

By Proposition 2.7, these four vectors do weak sign retrieval in $\mathbb{R}^3$. However this family fails to do weak phaseless reconstruction. Observe the vectors $x = (1, 2, 0)$

and $y = (2, 1, 0)$ satisfy $|\langle x, x_k\rangle| = |\langle y, x_k\rangle|$ however do not have the same absolute value in each coordinate.

Our next example is a frame which has the weak phase property, but fails phase retrieval.

Example 1.41. The set of vectors below satisfy the weak phase property in $\mathbb{R}^4$. In this case, our vectors are the rows of the matrix:

$$\mathcal{X} = \left[\begin{array}{c|cccc} x_1 & 1 & 1 & 1 & -1 \\ x_2 & -1 & 1 & 1 & 1 \\ x_3 & 1 & -1 & 1 & 1 \\ x_4 & 1 & 1 & -1 & -1 \\ x_5 & 1 & -1 & 1 & -1 \\ x_6 & 1 & -1 & -1 & 1 \end{array}\right]$$

Note that $\mathcal{X}$ fails to do phase retrieval as it requires seven vectors in $\mathbb{R}^4$ to do phase retrieval in $\mathbb{R}^4$. Given $x = (a_1, a_2, a_3, a_4),\ y = (b_1, b_2, b_3, b_4)$ we assume

$$|\langle x, x_k\rangle|^2 = |\langle y, x_k\rangle|^2, \text{ for all } k = 1, 2, 3, 4, 5, 6. \tag{1.11}$$

Step 1: The following is the coefficient matrix obtained after expanding $|\langle x, x_k\rangle|^2$ for $k = 1, 2, \ldots, 6$.

$$\frac{1}{2}\left[\begin{array}{c|cccccc|c} & a_1a_2 & a_1a_3 & a_1a_4 & a_2a_3 & a_2a_4 & a_3a_4 & \sum_{i=1}^4 a_i^2 \\ E_1 & 1 & 1 & -1 & 1 & -1 & -1 & \frac{1}{2} \\ E_2 & -1 & -1 & -1 & 1 & 1 & 1 & \frac{1}{2} \\ E_3 & -1 & 1 & 1 & -1 & -1 & 1 & \frac{1}{2} \\ E_4 & 1 & -1 & -1 & -1 & -1 & 1 & \frac{1}{2} \\ E_5 & -1 & 1 & -1 & -1 & 1 & -1 & \frac{1}{2} \\ E_6 & -1 & -1 & 1 & 1 & -1 & -1 & \frac{1}{2} \end{array}\right]$$

Step 2: Consider the following row operations, the last column becomes all zeroes so we drop it and we get:

$$\left[\begin{array}{l|cccccc} F_1 = \frac{1}{2}(E_1 - E_4) & 0 & 1 & 0 & 1 & 0 & -1 \\ F_2 = \frac{1}{2}(E_2 - E_5) & 0 & -1 & 0 & 1 & 0 & 1 \\ F_3 = \frac{1}{2}(E_3 - E_6) & 0 & 1 & 0 & -1 & 0 & 1 \\ A_1 = \frac{1}{2}(F_1 + F_2) & 0 & 0 & 0 & 1 & 0 & 0 \\ A_2 = \frac{1}{2}(F_1 + F_3) & 0 & 1 & 0 & 0 & 0 & 0 \\ A_3 = \frac{1}{2}(F_2 + F_3) & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

Step 3: Subtracting out A_1, A_2 and A_3 from E_1, E_2, E_3 and E_4, we get:

$$\left[\begin{array}{l|cccccc} E_1' = & 1 & 0 & -1 & 0 & -1 & 0 \\ E_2' = & -1 & 0 & -1 & 0 & 1 & 0 \\ E_3' = & -1 & 0 & 1 & 0 & -1 & 0 \\ E_4' = & 1 & 0 & -1 & 0 & -1 & 0 \end{array}\right]$$

Step 4: We will show that $a_i a_j = b_i b_j$ for all $i \neq j$.

Performing the given operations we get:

$$\left[\begin{array}{l|cccccc} D_1 = \frac{-1}{2}(E_2' + E_3') & 1 & 0 & 0 & 0 & 0 & 0 \\ A_2 & 0 & 1 & 0 & 0 & 0 & 0 \\ D_2 = \frac{-1}{2}(E_1' + E_2') & 0 & 0 & 1 & 0 & 0 & 0 \\ A_1 & 0 & 0 & 0 & 1 & 0 & 0 \\ D_3 = \frac{-1}{2}(E_3' + E_4') & 0 & 0 & 0 & 0 & 1 & 0 \\ A_3 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

Doing the same operations with $y = (b_1, b_2, b_3, b_4)$ we get:

$$a_i a_j = b_i b_j, \text{ for all } 1 \leq i \neq j \leq 4.$$

It should be noted that weak phase retrieval does not imply norm retrieval. We may use the previous example to illustrate this. Let $\mathcal{X} = \{x_i\}_{i=1}^6$ be as in Example 1.41. Suppose $\mathcal{X}$ does norm retrieval. Since there are only 6 vectors $\mathcal{X}$ fails the complement property. Now, take $x = (1, 1, -1, 1) \perp \{x_1, x_2, x_3\}$ and $y = (1, 1, 1, 1) \perp \{x_4, x_5, x_6\}$. Then, we have $|\langle x + y, x_i| = |\langle x - y, x_i\rangle|$ for all $i = 1, 2, \ldots 6$. The definition implies $\|x + y\| = \|x - y\|$. Since $\|x\| = \|y\|$, this implies that $x \perp y$, which is a contradiction.

1.3 Phase Retrieval in ℓ_2

The purpose of this section is to highlight some differences between finite and infinite dimensional phase retrieval. While most of the known classifications hold in ℓ_2, there are many surprising differences between the finite and infinite dimensional cases.

It is known [12] that the families of vectors $\{x_i\}_{i=1}^m$ which do phase retrieval in $\mathbb{R}^n$ are dense in the family of $m \geq (2n-1)$-element sets of vectors in $\mathbb{R}^n$. This follows from the fact that full spark families of $m \geq 2n-1$ vectors are dense and do phase retrieval. The corresponding result fails in infinite dimensions.

Definition 1.42. We say a family of sequences of vectors $\mathcal{F}$ is **dense** in ℓ_2 if given any sequence of vectors $\mathcal{Y} = \{y_i\}_{i=1}^\infty \subset \ell_2$ and any $\epsilon > 0$ there an $\mathcal{X} = \{x_i\}_{i=1}^\infty \in \mathcal{F}$ so that

$$d(\mathcal{X}, \mathcal{Y})^2 = \sum_{i=1}^\infty \|x_i - y_i\|^2 < \epsilon.$$

First we need the definition of a Riesz basis.

Definition 1.43. A collection of vectors $\{x_i\}_{i\in I}$ in a Hilbert space $\mathbb{H}$ is a Riesz basis for $\mathbb{H}$ if it is the image of an orthonormal basis for $\mathbb{H}$ under an invertible linear transformation. In other words, if there is an orthonormal basis $\{e_i\}$ for $\mathbb{H}$ and an invertible transformation T such that $Te_i = x_i$ for all i.

Note that a Riesz basis cannot do phase retrieval since it fails complement property.

Proposition 1.44. *Let $\mathcal{X} = \{x_i\}_{i=1}^\infty \subset \ell_2$ be such that $\sum_{i=1}^\infty \|x_i - e_i\|^2 \leq 1 - \epsilon$. Then $\mathcal{X}$ is a Riesz basis for ℓ_2.*

Proof. Define an operator $T : \ell_2 \to \ell_2$ by $Te_i = x_i$, for all $i = 1, 2, \ldots$. Given $a = \sum_{i=1}^{\infty} a_i e_i \in \ell_2$ we have

$$\|(I - T)a\|^2 = \|\sum_{i=1}^{\infty} a_i(e_i - x_i)\|^2 \tag{1.12}$$

$$\leq \sum_{i=1}^{\infty} |a_i| \|e_i - x_i\| \tag{1.13}$$

$$\leq \left(\sum_{i=1}^{\infty} |a_i|^2\right)\left(\sum_{i=1}^{\infty} \|x_i - e_i\|^2\right) \tag{1.14}$$

$$\leq (1 - \epsilon)\|a\|^2. \tag{1.15}$$

It follows that T is an invertible operator and so $\{Te_i\}_{i=1}^{\infty}$ is a Riesz basis for ℓ_2. ∎

Proposition 1.45. *The families of vectors which do phase retrieval in ℓ_2 are not dense in the infinite families of vectors in ℓ_2.*

Proof. Let $0 < \epsilon < 1$. If $\mathcal{X} = \{x_i\}_{i=1}^{\infty}$ is any family of unit vectors with

$$\sum_{i=1}^{\infty} \|x_i - e_i\|^2 < 1 - \epsilon, \tag{1.16}$$

then $\mathcal{X}$ is a Riesz basis and hence cannot do phase retrieval. ∎

It is known in finite dimensions [2, 12] that if $\mathcal{X} = \{x_i\}_{i=1}^{m}$ does phase retrieval, there is an $\epsilon > 0$ so that whenever $\mathcal{Y} = \{y_i\}_{i=1}^{m}$ satisfies:

$$\sum_{i=1}^{m} \|x_i - y_i\|^2 < \epsilon, \tag{1.17}$$

then $\mathcal{Y}$ does phase retrieval. The above is called a **ϵ-perturbation** of $\mathcal{X}$. The corresponding result fails in ℓ_2 as was shown in [11].

Theorem 1.46 ([11])**.** *Given a frame $\{x_i\}_{i=1}^{\infty}$ doing phase retrieval in ℓ_2 and an $\epsilon > 0$, there is a frame $\{y_i\}_{i=1}^{\infty}$ which fails phase retrieval in ℓ_2 and satisfies:*

$$\sum_{i=1}^{\infty} \|x_i - y_i\|^2 < \epsilon. \tag{1.18}$$

Next we try to generalize the definition of full spark, however we will see that such families may not do phase retrieval.

Definition 1.47. A set of vectors $\{x_i\}_{i=1}^{\infty}$ in ℓ_2 is **finitely full spark** if for every $I \subset \mathbb{N}$ with $|I| = n$, $\{P_I x_i\}_{i=1}^{\infty}$ is full spark (i.e. spark n+1), where P_I is the orthogonal projection onto span$\{e_i\}_{i\in I}$.

Proposition 1.48. *The finitely full spark families of vectors in ℓ_2 are dense in the infinite families of vectors in ℓ_2. In particular, there are Riesz bases for ℓ_2 which are finitely full spark, and these families cannot do phase retrieval.*

Proof. Let $\{y_i\}_{i=1}^{\infty}$ be a family of vectors in ℓ_2 and fix $\epsilon > 0$. We will construct the vectors by induction. To get started, choose a vector x_1 with all non-zero coordinates so that $\|x_1 - y_1\|^2 < \frac{\epsilon}{2}$. Now assume we have constructed vectors $\{x_i\}_{i=1}^{m}$ so that for every $I \subset \mathbb{N}$ with $|I| < \infty$, $\{P_I x_i\}_{i=1}^{m}$ is full spark and $\|x_i - y_i\|^2 < \frac{\epsilon}{2^{i+1}}$. For each finite subset $I \subset \mathbb{N}$, let

$$\mathcal{G}_I = \bigcup \left\{ \overline{\text{span}} \left[\{P_I x_i\}_{i\in I'} \cup \{e_i\}_{i\in I^c} \right] : I' \subset [m], \begin{cases} |I'| = m \text{ if } m+1 \leq |I| \\ |I'| = |I| - 1 \text{ if } |I| \leq m \end{cases} \right\}.$$

Let $\mathcal{F} = \bigcup_{n=1}^{\infty} \bigcup_{|I|=n} \mathcal{G}_I$, then $\mathcal{F}$ is a countable union of proper subspaces of ℓ_2 and hence there exists a vector y_{m+1} not in $\mathcal{F}$ and $\|x_{m+1} - y_{m+1}\|^2 < \frac{\epsilon}{2^{m+1}}$. This provides the required family of finitely full spark vectors. ∎

One interpretation of the definition of full spark is that any minimal number of vectors in the set which could possibly span, must span, i.e. any subset of n-vectors must span. The corresponding statement for ℓ_2 is:

Definition 1.49. A family of vectors $\{x_i\}_{i=1}^{\infty}$ is **full spark** in ℓ_2 if every infinite subset spans ℓ_2.

A full spark set clearly has complement property and hence does phase retrieval in the infinite dimensional case.

Theorem 1.50. *There exist full spark families of vectors in ℓ_2 which then do phase retrieval.*

Proof. Such an example can be found in Theorem 2 of [36]. For another example, consider $L_2[0,1]$. It is known that if a sequence $a_n \neq a$ of numbers (real or complex) tends to a when $n \to \infty$, then the sequence of functions $f_n(t) = e^{a_n t}$ spans $L_2[0,1]$ (this is a standard application of the Hahn-Banach theorem together with the uniqueness theorem for holomorphic functions, see more in Appendix III of [31].) Since every subsequence of a_n also has the same limit, every subsequence of f_n also spans $L_2[0,1]$. ∎

In the following, we will show how to create a new phase retrieval set by translating the vectors of the original one in the same direction. First, we will need a lemma.

Lemma 1.51. *If $\{x_i\}_{i=1}^{\infty}$ is Bessel in ℓ_2, then for every $v \in \ell_2$,*

$$\lim_{i\to\infty} \langle v, x_i \rangle = 0. \tag{1.19}$$

Proof. Given a vector v, we have $\sum_{i=1}^{\infty} |\langle v, x_i \rangle|^2 < \infty$, hence $\lim_{i\to\infty} |\langle v, x_i \rangle| = 0$. ∎

Note that if any $\{x_i\}_{i=1}^{\infty}$ does phase retrieval, then $\left\{ \frac{1}{\|x_i\| 2^i} x_i \right\}_{i=1}^{\infty}$ is Bessel and also does phase retrieval.

Theorem 1.52. *Assume $\{x_i\}_{i=1}^{\infty}$ is a Bessel sequence in ℓ_2 and does phase retrieval. Then for every $v \in \ell_2$, $\{x_i + v\}_{i=1}^{\infty}$ does phase retrieval.*

Proof. Assume

$$|\langle x, x_i + v\rangle| = |\langle y, x_i + v\rangle|, \text{ for all } i = 1, 2, \ldots.$$

Let $I = \{i : \langle x, x_i + v\rangle = \langle y, x_i + v\rangle\}$, then either $|I|$ or $|I^c|$ is infinite. By the complement property, either $\{x_i\}_{i\in I}$ or $\{x_i\}_{i\in I^c}$ spans the space. Without loss of generality, assume $\{x_i\}_{i\in I}$ spans ℓ_2. Now,

$$\langle x, x_i + v\rangle = \langle y, x_i + v\rangle, \text{ for all } i \in I,$$

and so $\langle x - y, x_i\rangle = \langle y - x, v\rangle$, for all $i \in I$. By Lemma 1.51,

$$\langle y - x, v\rangle = 0 = \langle x - y, x_i\rangle, \text{ for all } i \in I. \tag{1.20}$$

It follows that $x - y = 0$. ∎

Note that $\{x_i+v\}_{i=1}^{\infty}$ is not Bessel. But we can scale it to be Bessel and it still does phase retrieval. By repeating the argument in the previous corollary, it is possible to effectively "delete" a finite number of vectors by translating the system and scaling the set so they are Bessel.

Proposition 1.53. *There is a family of vectors in ℓ_2 doing phase retrieval where each of the vectors has all non-zero coordinates with respect to the unit vectors.*

Proof. Let $\{x_i\}_{i=1}^{\infty}$ do phase retrieval. Let $\{e_i\}_{i=1}^{\infty}$ be the unit vectors. Recall $x_{ij} = x_i(j)$ denotes the j^{th} coordinate of x_i. For any $j = 1, 2, \ldots$ the family $\{x_{ij}\}_{i=1}^{\infty}$ is a countable set of real numbers, so we may choose a real number $a_j \neq -x_{ij}$ and $0 < a_j < \frac{1}{2^j}$ for all $i = 1, 2, \ldots$. Let $v = (a_1, a_2, \ldots)$. Then $\{x_i + v\}_{i=1}^{\infty}$ does phase retrieval and each vector has all non-zero coordinates. ∎

1.3.1 Lifting

The following theorem is an infinite dimensional version of Theorem 1.9. Note that this is not a classification of the liftable phase retrieving frames, but is a sufficient condition.

Theorem 1.54. *Let $\mathcal{X} = \{x_i\}_{i=1}^{\infty}$ be a frame for ℓ_2 doing phase retrieval and let $\mathcal{Y} = \{y_i\}_{i=1}^{\infty}$ be a linearly dependent spanning set in ℓ_2. Then $\mathcal{X} \cup \mathcal{Y}$ can be lifted to a phase retrieving frame for ℓ_2.*

Proof. Let $\mathcal{X} = \{x_i\}_{i=1}^{\infty}$ and $\mathcal{Y} = \{y_i\}_{i=1}^{\infty}$ be as in the theorem. We show that we can lift this union to one higher dimension and maintain phase retrieval. Let L be the right shift operator on ℓ_2, i.e. if $x = (a_1, a_2, \ldots)$, then $Lx = (0, a_1, a_2, \ldots)$. Replace vectors in $\mathcal{X}$ by $\hat{\mathcal{X}} = \{\hat{x}_i\}_{i=1}^{\infty}$ where $\hat{x}_i = Lx_i$. The idea for $\hat{\mathcal{Y}} = \{\hat{y}_i\}_{i=1}^{\infty}$ is very similar to the proof in Theorem 1.9. We show existence of a vector $v = (b_1, b_2, \ldots) \in \ell_2$ such that $\hat{y}_i = b_i e_1 + Ly_i$ will have the desired property to assure $\hat{\mathcal{X}} \cup \hat{Y}$ does phase retrieval.

Since $\{y_i\}_{i=1}^{\infty}$ is linearly dependent, there exists a sequence of scalars $\alpha = \{\alpha_i\}_{i=1}^{\infty}$ with all but a finite number equal to zero, such that $\sum_{i=1}^{\infty} \alpha_i y_i = 0$. Denote $H_i = e_i^{\perp} \subset \ell_2$ and note that by the Baire Category Theorem

$$\left[(\cup_{i=1}^{\infty} H_i) \cup \alpha^{\perp}\right]^c \neq \emptyset. \tag{1.21}$$

Let $v = (b_1, b_2, \ldots) \in \left[(\cup_{i=1}^{\infty} H_i) \cup \alpha^{\perp}\right]^c$ and define $\hat{y}_i$ as stated above. Note that v has all non-zero coordinates and $\langle v, \alpha \rangle \neq 0$. Moreover,

$$\sum_{i=1}^{\infty} \alpha_i \hat{y}_i = \sum_{i=1}^{\infty} (\alpha_i (b_i e_1 + Ly_i) = \langle \alpha, v \rangle e_1. \tag{1.22}$$

Let any $j \geq 1$ and $\epsilon > 0$. Since $\{y_i\}_{i=1}^{\infty}$ spans ℓ_2, there is a finite subset I_j and scalars $\{\beta_k\}_{k \in I_j}$ such that

$$\|e_j - \sum_{k \in I_j} \beta_k y_k\| < \epsilon. \tag{1.23}$$

This implies

$$\|e_{j+1} - \sum_{k \in I_j} \beta_k (\hat{y}_k - b_k e_1)\| < \epsilon, \text{ for all } j \geq 1. \tag{1.24}$$

Since $e_1 \in \text{span}\{\hat{y}_i\}_{i=1}^{\infty}$, $e_j \in \overline{\text{span}}\{\hat{y}_i\}_{i=1}^{\infty}$ for all j, $\hat{\mathcal{Y}} = \{\hat{y}_i\}_{i=1}^{\infty}$ spans ℓ_2.

Now we will show that $\hat{\mathcal{X}} \cup \hat{\mathcal{Y}}$ satisfies Edidin's theorem. Since $\langle e_1, \hat{y}_i \rangle = b_i \neq 0$, the projection of e_1 on the vectors of $\hat{\mathcal{Y}}$ span ℓ_2. Let any non-zero vector $x \neq e_1$, the projection of x onto the vectors $\hat{x}_i$ will spans $e_1^{\perp} \subset \ell_2$. Note that x cannot be orthogonal to all $\hat{y}_i$ since these vectors span ℓ_2. Let $\hat{y}_j$ be one such vector. Since $\hat{y}_j$ is outside of $e_1^{\perp} \subset \ell_2$, the projection of x onto the vectors of $\hat{\mathcal{X}} \cup \hat{\mathcal{Y}}$ span ℓ_2 as well. Hence $\hat{\mathcal{X}} \cup \hat{\mathcal{Y}}$ does phase retrieval. ∎

1.3.2 Sets Which do Phase Retrieval in ℓ_2

Theorem 1.55. *Assume we have subspaces $W_1 \subset W_2 \subset \cdots \subset \ell_2$ and vectors $\{x_{(i,j)}\}_{j \in I_i}$ doing phase retrieval in W_i for every i. Finally, assume $\cup_{i=1}^{\infty} W_i$ is dense in ℓ_2. Then $\{x_{(i,j)}\}_{i=1, j \in I_i}^{\infty}$ does phase retrieval in ℓ_2.*

Proof. We will check the complement property. Observe that a partition of vectors $\{x_{(i,j)}\}_{i=1, j \in I_i}^{\infty}$ induces a partition for vectors $\{x_{(i,j)}\}_{j \in I_i} \subset W_i$. By assumption $\{x_{(i,j)}\}_{j \in I_i}$ does phase retrieval on W_i, therefore for each $i = 1, 2, \ldots$

$$\text{either } W_i \subset \overline{\text{span}}\{x_{(i,j)}\}_{(i,j) \in I} \text{ or } W_i \subset \overline{\text{span}}\{x_{(i,j)}\}_{(i,j) \in I^c}. \tag{1.25}$$

Then either I or I^c contains infinitely many W_i, without loss of generality we assume it is I. This means that for infinitely many i,

$$W_i \subset \overline{\text{span}}\{x_{(i,j)}\}_{(i,j)\in I}. \tag{1.26}$$

Since $W_i \subset W_{i+1}$ for all i,

$$\cup_{i=1}^{\infty} W_i \subset \overline{\text{span}}\{x_{(i,j)}\}_{(i,j)\in I}, \tag{1.27}$$

and so the closure of the right hand set is ℓ_2. This shows our family of vectors have complement property and hence do phase retrieval on ℓ_2. ∎

In finite dimensions it is known [3] that any family of vectors doing phase retrieval must contain at least $(2n-1)$-vectors. It follows that a full spark family of vectors $\{x_i\}_{i=1}^{2n-1}$ does phase retrieval (since it has complement property) but if we delete any vector it fails phase retrieval. The following example shows that there is a family of vectors in ℓ_2 which does phase retrieval but we cannot drop any vector and maintain phase retrieval. First, we need the following lemma:

Lemma 1.56. *Let $\{e_i\}_{i=1}^{\infty}$ be the canonical orthonormal basis for ℓ_2. For any fixed i, if x is orthogonal to $e_i + e_j$ for infinitely many $j > i$, then $\langle x, e_i\rangle = \langle x, e_j\rangle = 0$ for all such j.*

Proof. Let $K = \{j : j > i, \langle x, e_i + e_j\rangle = 0\}$, then by assumption, the cardinality of K is infinite.

It is clear that $|\langle x, e_i\rangle| = |\langle x, e_j\rangle|$ for all $j \in K$. Suppose by a contradition that $|\langle x, e_j\rangle| > 0$ for all $j \in K$. Then we have $\|x\|^2 \geq \sum_{j\in K} |\langle x, e_j\rangle|^2 = \infty$, a contradiction. ∎

Example 1.57. Let the family of vectors $\mathcal{X} = \{e_i + e_j\}_{i<j}$. Then $\mathcal{X}$ does phase retrieval in ℓ_2 but we cannot drop any vector of $\mathcal{X}$ and maintain phase retrieval.

Proof. Let I be any subset of the set $\{(i,j) : i < j\}$, and we can assume that $(1,j) \in I$ for infinitely many j. We will show that either $\{e_i + e_j\}_{(i,j)\in I}$ or $\{e_i + e_j\}_{(i,j)\in I^c}$ spans ℓ_2. Suppose $\{e_i + e_j\}_{(i,j)\in I}$ does not span ℓ_2. We will show that $\{e_i + e_j\}_{(i,j)\in I^c}$ spans ℓ_2.

Let any $x = (x(1), x(2), \ldots)$ be such that $\langle x, e_i + e_j\rangle = 0$ for all $(i,j) \in I^c$.

By assumption, there is $y = (y(1), y(2), \ldots), y \neq 0$ and $\langle y, e_i + e_j\rangle = 0$ for all $(i,j) \in I$. Let s be the smallest number such that $y(s) \neq 0$. By Lemma 1.56, $(s,j) \notin I$ for infinitely many $j > s$. Hence there is $t > s$ such that $(s,j) \in I^c$ for all $j \geq t$. Again, by Lemma 1.56, we get

$$x(s) = x(j) = 0 \text{ for all } j \geq t.$$

We will now show that $x(j) = 0$ for all $j = 1, 2 \ldots t-1$. Suppose there is $1 \leq j < s$ such that $x(j) \neq 0$. This implies $(j,s) \notin I^c$. Thus $(j,s) \in I$ and hence $y(j) \neq 0$. But this contradicts the way we chose s. So $x(j) = 0$ for all $1 \leq j < s$.

Now let any $s < j < t$. If $(s,j) \in I^c$, then $x(j) = 0$. If $(s,j) \in I$, then $y(j) \neq 0$. Note that by assumption, $(1,j) \in I$ for infinitely many j, and hence by Lemma 1.56, we get that $y(1) = 0$. Thus, $(1,j) \notin I$. Therefore $(1,j) \in I^c$ and so $x(j) = x(1) = 0$. This completes the proof that $\{e_i + e_j\}_{(i,j)\in I^c}$ span ℓ_2.

Now we will show that we cannot drop any vector of $\mathcal{X}$ and maintain phase retrieval.

Fix any $(k,\ell), k < \ell$. Consider $\mathcal{Y} = \{e_i + e_j : i < j, (i,j) \neq (k,\ell)\}$. Let

$x = e_k + e_\ell,\ y = e_k - e_\ell$. Clearly, $x \neq \pm y$. For any vector $e_i + e_j \in \mathcal{Y}$, we compute:

$$\langle x, e_i + e_j \rangle = \langle e_k, e_i \rangle + \langle e_k, e_j \rangle + \langle e_\ell, e_i \rangle + \langle e_\ell, e_j \rangle,$$

$$\langle y, e_i + e_j \rangle = \langle e_k, e_i \rangle + \langle e_k, e_j \rangle - \langle e_\ell, e_i \rangle - \langle e_\ell, e_j \rangle.$$

If $i = k$ then $j \neq \ell$, $i < \ell$ and $k < j$. Thus $\langle x, e_i + e_j \rangle = \langle y, e_i + e_j \rangle = 1$. If $j = k$, then $i < j = k < \ell$. So $\langle x, e_i + e_j \rangle = \langle y, e_i + e_j \rangle = 1$.

Consider the case $i, j \neq k$. If $i = \ell$ then $j \neq \ell$. Hence

$$\langle x, e_i + e_j \rangle = 1, \text{ and } \langle y, e_i + e_j \rangle = -1. \tag{1.28}$$

If $i \neq \ell$ and $j = \ell$ then $\langle x, e_i + e_j \rangle = 1$, and $\langle y, e_i + e_j \rangle = -1$. Finally, if $i \neq \ell$ and $j \neq \ell$ then $\langle x, e_i + e_j \rangle = \langle y, e_i + e_j \rangle = 0$. Thus, in all cases, we always have that

$$|\langle x, e_i + e_j \rangle| = |\langle y, e_i + e_j \rangle|, \text{ for all } e_i + e_j \in \mathcal{Y}. \tag{1.29}$$

Since $x \neq \pm y$, $\mathcal{Y}$ cannot do phase retrieval.

∎

Theorem 1.58. *Let P_n be the orthogonal projection of ℓ_2 onto $E_n = \text{span}\{e_i\}_{i=1}^n$. There is a set of vectors $\mathcal{Y} = \{y_{(n,i)}\}_{n=1,i=1}^{\infty,\ \infty}$ that does not do phase retrieval on ℓ_2, but $\mathcal{X} = \{x_{(n,i)}\}_{n=1,i=1}^{\infty,\ \infty} = \{P_n y_{(n,i)}\}_{n=1,i=1}^{\infty,\ \infty}$ does phase retrieval in ℓ_2. Moreover, finite subsets of $\mathcal{X}$ do phase retrieval on E_n for every n.*

Proof. For each $n \in \mathbb{N}$, let $\mathcal{X}_n$ be a finite set of vectors $\{x_{(n,i)}\}_{i \in I_n}$ contained in E_n that does phase retrieval in E_n. For example consider a full spark set in E_n embedded in ℓ_2 by adding zero to all other entries. We know that $\mathcal{X} = \{x_{(n,i)}\}_{n=1, i \in I_n}^{\infty}$ does phase retrieval in ℓ_2. It is sufficient to show that for each n and i, there exists $y_{(n,i)}$, with $P_n y_{(n,i)} = x_{(n,i)}$, such that the $y_{(n,i)}$ is contained in a fixed hyperplane for all (n, i).

Let w be the vector with infinitely many non-zero coordinates. For each n, $x_{(n,i)}$ has finite support contained in the first n coordinates, for all $i \in I_n$. Then there is $j > n$ such that the j^{th} coordinate of w is non-zero. Define $y_{(n,i)} = x_{(n,i)} - \dfrac{\langle x_{(n,i)}, w\rangle}{w(j)} e_j$, for $i \in I_n$. It follows that $\langle y_{(n,i)}, w\rangle = 0$, and hence $y_{(n,i)} \subset w^{\perp}$ for all (n, i). This completes the proof. ∎

Chapter 2

Quantum Detection Problem

In this chapter we will give a complete answer to the frame quantum detection problem including the injectivity problem and state estimation problem. We will answer the problem in both the real and complex cases and in both the finite dimensional and infinite dimensional cases. To explain exactly what we will solve, we need to introduce the basics of quantum detection. Let $L^\infty(\mathbb{H})$ be the space of bounded linear operators on a finite or infinite dimenional (real or complex) Hilbert space $\mathbb{H}$. For an operator $T \in L_0(\mathbb{H})$, the finite rank operators on $\mathbb{H}$, the *trace* of T is given by: $\text{tr}(T) = \sum_{i\in I}\langle Te_i, e_i\rangle$, which is finite and independent of the orthonormal basis. The trace induces a scalar product by $\langle T, S\rangle_{HS} = \text{tr}(TS^*)$. The closure of $L_0(\mathbb{H})$ with respect to this scalar product, denoted $L^2(\mathbb{H})$ is the space of the Hilbert-Schmidt operators on $\mathbb{H}$. For any $T \in L^\infty(\mathbb{H})$ let $|T| = \sqrt{TT^*}$, the positive square root of TT^*. We say that T is a *trace class operator* if $\text{tr}(|T|) < \infty$. The set of all trace class operators is denoted by $L^1(\mathbb{H})$ and forms a Banach space under the *trace norm* $\|T\|_1 = \text{tr}(|T|)$.

Let $Sym(\mathbb{H})$ denote the real Banach space of self-adjoint operators on $\mathbb{H}$, or

$$Sym(\mathbb{H}) = \{T : T \in L^\infty(\mathbb{H}),\ T = T^*\},$$

and let $Sym^+(\mathbb{H})$ denote the real cone of positive self-adjoint operators on $\mathbb{H}$

$$Sym^+(\mathbb{H}) = \{T = T^* \geq 0\}$$

The main objects to analyze these operators are the *positive operator-valued measures*. In quantum mechanics, the definition of a von Neumann measurement can be generalized using positive operator-valued measures (POVMs) [23, 22, 28]. This generalization allows one to distinguish more accurately among elements of a set of non-orthogonal quantum states.

Let X denote a set of outcomes (e.g. this could be a finite or infinite subset of $\mathbb{Z}^d$ or $\mathbb{R}^d$) and β denote a sigma algebra of subsets of X.

Definition 2.1. A **positive operator-valued measure** (**POVM**) is a function $\Pi : \beta \to Sym^+(\mathbb{H})$ satisfying:

1. $\Pi(\emptyset) = 0$ (the zero operator).

2. For every at most countable disjoint family $\{U_i\}_{i\in I} \subset \beta$, $x, y \in \mathbb{H}$ we have

$$\langle \Pi\left(\cup_{i\in I} U_i\right) x, y\rangle = \sum_{i\in I} \langle \Pi(U_i)x, y\rangle.$$

3. $\Pi(X) = I$ (the identity operator).

A *quantum system* is defined as a von Neumann algebra $\mathcal{A}$ of operators acting on $\mathbb{H}$. The set of *states* on $\mathbb{H}$ is

$$\mathcal{S}(\mathbb{H}) = \{T \in L^1(\mathbb{H}),\ T = T^* \geq 0,\ \mathrm{tr}(T) = 1\},$$

and represents the reservoir of *quantum states* for any quantum system.

The set of *quantum states* $\mathcal{S}(\mathcal{A})$ associated to a quantum system $\mathcal{A}$ is obtained by identifying states that differ by a null state with respect to $\mathcal{A}$. Thus, the set of quantum states are in one-to-one correspondance with the linear functionals on $\mathcal{A}$ of the form:

$$\rho : \mathcal{A} \to \mathbb{C}, \text{ for some } S \in \mathcal{S}(\mathbb{H}),\ \rho(T) = \mathrm{tr}(TS), \quad \text{for every } T \in \mathcal{A}.$$

A quantum state $\rho \in \mathcal{S}(\mathcal{A})$ is called a *pure state* if it is an extreme point in the convex *weak** compact set of quantum states $\mathcal{S}(\mathcal{A})$. We say a POVM Π is *associated to a von Neumann algebra* $\mathcal{A}$ if $\Pi : \beta \to \mathcal{A} \cap Sym^+(\mathbb{H})$.

Given a quantum state ρ, the *quantum measurement* performed by the POVM Π is the map $p : \beta \to \mathbb{R}$ defined by $p(U) = \rho(\Pi(U)) = \mathrm{tr}(\Pi(U)T)$, where $T \in \mathcal{S}(\mathbb{H})$ is in the equivalence class associated to ρ.

2.0.1 The Quantum Detection Problem

Let $L(\beta, \mathbb{R})$ denote the set of real-valued bounded functions defined on β. Given a POVM Π associated to a von Neumann algebra $\mathcal{A}$, the *quantum detection problem* asks if there is a unique quantum state $\rho \in \mathcal{S}(\mathcal{A})$ compatible with the set of quantum measurements performed by the POVM Π?

More specifically, the quantum detection problem asks two questions:

1. *Injectivity, or state separability*: Is the following map injective

$$\mathbb{M} : \mathcal{S}(\mathcal{A}) \to L(\beta, \mathbb{R}), \quad \mathbb{M}(\rho)(U) = \rho(\Pi(U))?$$

2. *Range analysis, or state estimation*: Assume $\mathbb{M}$ is injective. Then, given a map $p \in L(\beta, \mathbb{R})$, determine if p is in the range of $\mathbb{M}$, hence is of the form

$p = \mathbb{M}(\rho)$ for some unique $\rho \in \mathcal{S}(\mathcal{A})$. If not, find a quantum state ρ that best approximates p in some sense (e.g. robustness to noise).

It should be mentioned that, in this context, a significant amount of work has been put into computing the probability of detection error [9, 28, 27, 32, 30].

Shortly we will introduce the Hilbert space frame version of the quantum detection problem, but first some prelimarinaries about frame POVMs. For some background on frame POVMs see [7, 23, 26, 19].

If $\{x_k\}_{k\in I}$ is a Parseval frame for a Hilbert space $\mathbb{H}$, it naturally induces a POVM Π on $X = I$ with $\beta = 2^I$ (the power set of I):

$$\Pi(U) = \sum_{k\in U} x_k x_k^*, \text{ where } x_k^* : \mathbb{H} \to \mathbb{C}, \ x_k^*(x) = \langle x, x_k\rangle,$$

with strong convergence for any $U \subset I$.

Given a state $T \in \mathcal{S}(\mathbb{H})$ (i.e. a unit-trace, trace class, positive, self-adjoint operator on $\mathbb{H}$), the frame induced quantum measurement is given by the function

$$p : \beta \to \mathbb{R}, \quad p(U) = \sum_{k\in U} \mathrm{tr}(T x_k x_k^*) = \sum_{k\in U} \langle T x_k, x_k\rangle.$$

For the Von Neumann algebra $\mathcal{A} = L^\infty(\mathbb{H})$, the quantum states coincide with the convex set of states $\mathcal{S}(\mathbb{H})$. In this case, we may pose the injectivity and state estimation problems as follows:

1. *Injectivity, or state separability*: Is there a Parseval frame $\mathcal{X} = \{x_k\}_{k\in I}$ so that the map $\mathbb{M} : \mathcal{S}(\mathbb{H}) \to L(\beta, \mathbb{R})$ defined by $\mathbb{M}(T)(U) = \sum_{k\in U}\langle T x_k, x_k\rangle$ for $U \subset I$ is injective?

2. *State Estimation Problem*: Given an injective Parseval frame $\{x_k\}_{k\in I}$ and a

function $p : \beta \to \mathbb{R}$, is there any $T \in \mathcal{S}(\mathbb{H})$ so that $\mathbb{M}(T) = p$? If not, find a quantum state T that best approximates p.

We will first work on a more general problem and then investigate the added conditions. In particular, we first consider

1. Self-adjoint operators which may not be positive.

2. Operators which are not trace one, but are Hilbert Schmidt.

3. Frames which are not Parseval.

It will be shown that the approach used here to solve this more general problem will also solve the original problem. First, we need a definition.

Definition 2.2. A family of vectors $\mathcal{X} = \{x_k\}_{k\in I}$ in a Hilbert space $\mathbb{H}$ is said to be **injective** if given a Hilbert Schmidt self-adjoint operator T satisfying $\langle Tx_k, x_k\rangle = 0$ for all $k \in I$, then $T = 0$.

First we notice that if a family of vectors gives injectivity in a Hilbert space $\mathbb{H}^n$, then it is a frame for $\mathbb{H}^n$.

Proposition 2.3. *Let $\{x_k\}_{k=1}^m$ be a family of vectors in $\mathbb{H}^n$ which is injective. Then* $\text{span}\{x_k\}_{k=1}^m = \mathbb{H}^n$.

Proof. Suppose by contradiction that $W := \text{span}\{x_k\}_{k=1}^m \neq \mathbb{H}^n$. Let P be the orthogonal projection onto $W^\perp$. Then $\langle Px_k, x_k\rangle = 0$ for all k, but $P \neq 0$, a contradiction. ∎

Proposition 2.4. *Let $\{x_k\}_{k\in I}$ be a frame for $\mathbb{H}$ which gives injectivity. If F is a bounded invertible operator on $\mathbb{H}$, then $\{Fx_k\}_{k\in I}$ also gives injectivity.*

Proof. Let T be a Hilbert Schmidt self-adjoint operator such that

$$\langle TFx_k, Fx_k\rangle = 0, \text{ for all } k.$$

Then $\langle F^*TFx_k, x_k\rangle = 0$, for all k. Note that F^*TF is also a Hilbert Schmidt self-adjoint operator. Therefore, $F^*TF = 0$ and hence $T = 0$. ∎

The previous result shows that injectivity is preserved by bounded invertible operators. That is, we do not need to find Parseval frames for the quantum detection problem. If we have an injective frame, then its canonical Parseval frame is injective and we have theh following corollary.

Corollary 2.5. *Let $\{x_k\}_{k\in I}$ be a frame with frame operator S. If $\{x_k\}_{k\in I}$ gives injectivity, then the canonical Parseval frame $\{S^{-1/2}x_k\}_{k\in I}$ also gives injectivity.*

2.1 The Finite Dimensional Case

In this section we will solve the finite dimensional injectivity problem and the state estimation problem for both the real and complex cases. These problems were originally solved by Scott [35] (See also [8]) where the solutions are called *informationally complete quantum measurements*. The approach presented here provides more information about the solutions and easily generalizes to infinite dimensions. The following theorem shows that we do not need to work with positive operators.

Theorem 2.6. *Given a family of vectors $\mathcal{X} = \{x_k\}_{k=1}^m$ in $\mathbb{H}^n$, the following are equivalent:*

1. *Whenever T, S are positive and self-adjoint, and*

$$\langle Tx_k, x_k\rangle = \langle Sx_k, x_k\rangle, \text{ for all } k,$$

then $T = S$.

2. *Whenever* T, S *are self-adjoint, and*

$$\langle Tx_k, x_k\rangle = \langle Sx_k, x_k\rangle, \text{ for all } k,$$

then $T = S$.

3. $\mathcal{X}$ *is injective.*

Proof. (1) $\Rightarrow$ (2): Let T, S be self-adjoint operators such that

$$\langle Tx_k, x_k\rangle = \langle Sx_k, x_k\rangle, \text{ for all } k.$$

Set

$$m_1 := \inf_{\|x\|=1} \langle Tx, x\rangle, \quad m_2 := \inf_{\|x\|=1} \langle Sx, x\rangle$$

then $m_1, m_2 \in \mathbb{R}$. Set $m = \min\{m_1, m_2\}$.

Now let $P = T - mI$, $Q = S - mI$. Then for any $x \in \mathbb{H}, \|x\| = 1$, we have

$$\langle Px, x\rangle = \langle (T - mI)x, x\rangle = \langle Tx, x\rangle - m \geq 0.$$

Hence, P is positive. Similarly, Q is positive.

We have

$$\begin{aligned}\langle Px_k, x_k\rangle &= \langle (T - mI)x_k, x_k\rangle \\ &= \langle Tx_k, x_k\rangle - m\|x_k\|^2 \\ &= \langle Sx_k, x_k\rangle - m\|x_k\|^2 \\ &= \langle Qx_k, x_k\rangle.\end{aligned}$$

By (1) we get $P = Q$ and therefore $T = S$.

(2) $\Rightarrow$ (3) and (3) $\Rightarrow$ (1) are clear since $T - S$ is also a self-adjoint operator. ∎

If we further require that the operators are trace one, then to prove injectivity we only need to show that if T is trace zero and $\langle Tx_k, x_k\rangle = 0$ for all $k = 1, 2, \ldots$, then $T = 0$. This is because the trace is linear. That is, if T, S are trace one and $\langle Tx_k, x_k\rangle = \langle Sx_k, x_k\rangle$ for all k

$$\langle (T-S)x_k, x_k\rangle = 0, \text{ for all } k \text{ and } \operatorname{tr}(T-S) = 0.$$

The real case

We start with a propositon which motivates the classification for injectivity presented here.

Proposition 2.7. *Given a self-adjoint operator $T = (a_{ij})_{i,j=1}^n$ on $\mathbb{R}^n$ and a vector $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, we have*

$$\langle Tx, x\rangle = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_i x_j = \sum_{i=1}^n a_{ii}x_i^2 + 2\sum_{i=1}^n \sum_{j=i+1}^n a_{ij}x_i x_j.$$

Proof. First we compute:

$$Tx = \left(\sum_{j=1}^n a_{1j}x_j, \sum_{j=1}^n a_{2j}x_j, \ldots, \sum_{j=1}^n a_{nj}x_j\right) \tag{2.1}$$

$$\langle Tx, x\rangle = \sum_{j=1}^n a_{1j}x_j x_1 + \sum_{j=1}^n a_{2j}x_j x_2 + \cdots + \sum_{j=1}^n a_{nj}x_j x_n. \tag{2.2}$$

Using the fact that T is self-adjoint:

$$\begin{aligned}\langle Tx, x\rangle &= \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_i x_j \\ &= \sum_{i=1}^n a_{ii}x_i^2 + 2\sum_{1\le i<j\le n} a_{ij}x_i x_j.\end{aligned}$$

∎

This proposition inspired the following definition:

Definition 2.8. Given a vector $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, we define the following associated vector $\tilde{x}$ in $\mathbb{R}^{\frac{n(n+1)}{2}}$ by:

$$\tilde{x} = (x_1x_1, x_1x_2, \ldots, x_1x_n; x_2x_2, x_2x_3, \ldots, x_2x_n; \ldots; x_{n-1}x_{n-1}, x_{n-1}x_n; x_nx_n).$$

To a self-adjoint operator $T = (a_{ij})_{i,j=1}^n$ on $\mathbb{R}^n$, we associate a vector $\tilde{T}$ in $\mathbb{R}^{\frac{n(n+1)}{2}}$ by:

$$\tilde{T} = (a_{11}, 2a_{12}, \ldots, 2a_{1n}; a_{22}, 2a_{23}, \ldots, 2a_{2n}; \ldots; a_{(n-1)(n-1)}, 2a_{(n-1)n}; a_{nn}).$$

Proposition 2.7 now becomes:

Corollary 2.9. *Given a self-adjoint operator $T = (a_{ij})_{i,j=1}^n$ on $\mathbb{R}^n$ and a vector $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, we have $\langle Tx, x\rangle = \langle \tilde{T}, \tilde{x}\rangle$.*

We are now able to give a classification of the frames χ which give injectivity for the quantum detection problem.

Theorem 2.10. *Let $\mathcal{X} = \{x_k\}_{k=1}^m$ be a frame for $\mathbb{R}^n$. The following are equivalent:*

1. *$\mathcal{X}$ gives injectivity.*

2. *We have that $\{\tilde{x}_k\}_{k=1}^m$ spans $\mathcal{K} := \mathbb{R}^{\frac{n(n+1)}{2}}$.*

Proof. (1) $\Rightarrow$ (2): Let a vector

$$a = (a_{11}, a_{12}, \ldots, a_{1n}; a_{22}, a_{23}, \ldots, a_{2n}; \ldots; a_{(n-1)(n-1)}, a_{(n-1)n}; a_{nn}) \in \mathcal{K}$$

be such that $\langle a, \tilde{x}_k\rangle = 0$ for all k.

Define an operator $T = (b_{ij})_{i,j=1}^n$ on $\mathbb{R}^n$, where $b_{ii} = a_{ii}$ for $i = 1, 2, \ldots, n$ and $b_{ij} = b_{ji} = \dfrac{1}{2}a_{ij}$ for $i < j$. Then T is a self-adjoint operator with $\tilde{T} = a$. Hence by the previous Corollary 2.9 we have that $\langle Tx, x\rangle = \langle a, \tilde{x}\rangle$ for any $x \in \mathbb{R}^n$. Therefore, $\langle Tx_k, x_k\rangle = \langle a, \tilde{x}_k\rangle = 0$ for all k. This implies $T = 0$ and hence $a = 0$.

(2) $\Rightarrow$ (1): This direction is immediate by Corollary 2.9. If $\langle \tilde{T}, \tilde{x}_k \rangle = \langle Tx_k, x_k \rangle = 0$ for all k, then because $\{\tilde{x}_k\}_{k=1}^m$ spans $\mathcal{K}$, we have that $\tilde{T} = 0$. Hence, $T = 0$. ∎

The spanning condition in the previous theorem gives a lower limit on the number of vectors needed to achieve injectivity.

Corollary 2.11. *If a frame $\mathcal{X} = \{x_k\}_{k=1}^m$ gives injectivity in $\mathbb{R}^n$, then $m \geq \dfrac{n(n+1)}{2}$.*

As a consequence (see [17]):

Corollary 2.12. *Given a frame $\{x_k\}_{k=1}^m$ for $\mathbb{R}^n$, the following are equivalent:*

1. *The family $\{x_k x_k^*\}_{k=1}^m$ spans the class of self-adjoint operators.*
2. *The family of vectors $\{\tilde{x}_k\}_{k=1}^m$ spans $\mathbb{R}^{\frac{n(n+1)}{2}}$.*

Proof. This is immediate since for every $x \in \mathbb{R}^n$ and self-adjoint operator T, we have

$$\langle T, xx^* \rangle = \operatorname{tr}(Txx^*) = \langle Tx, x \rangle.$$

∎

In the standard frame quantum detection problem, there is the added assumption that the trace of the operators is one. As mentioned before, by linearity of the trace we only need to consider trace zero operators. The following simple example shows that with the added assumption $\operatorname{tr}(T) = 0$, we reduce the number of measurements.

Example 2.13. Let $\mathcal{X} = \{(1,0), (1,1)\}$ in $\mathbb{R}^2$. Then $\mathcal{X}$ gives injectivity in $\mathbb{R}^2$ for all self-adjoint operators of trace one.

Let $T = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ be a self-adjoint matrix of trace zero such that

$$\langle T(1,0), (1,0) \rangle = \langle T(1,1), (1,1) \rangle = 0.$$

Then $a = \langle T(1,0),(1,0)\rangle = 0$. Since $a + c = 0$ and

$$\langle T(1,1),(1,1)\rangle = \langle (a+b, b+c),(1,1)\rangle = a + 2b + c,$$

we find $b = c = 0$. Therefore, $T = 0$.

From this example, one may expect that for trace 0 operators we only need $\mathbb{R}^{\frac{n(n+1)}{2}-1}$ measurements. The following classification confirms this fact.

Theorem 2.14. *Let $\mathcal{X} = \{x_k\}_{k=1}^m$ be a frame for $\mathbb{R}^n$. The following are equivalent:*

1. *$\mathcal{X}$ gives injectivity for all self-adjoint operators of trace one.*

2. *We have that $\{\tilde{x}_k\}_{k=1}^m$ spans $\mathcal{K} := \mathbb{R}^{\frac{n(n+1)}{2}-1}$.*

Proof. For this version of the theorem we must consider a new associated vector $\tilde{x}$. Let $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Define

$$\tilde{x} = (x_1x_2, \ldots, x_1x_n; x_2^2 - x_1^2, x_2x_3, \ldots, x_2x_n; \ldots; x_{n-1}^2 - x_1^2, x_{n-1}x_n; x_n^2 - x_1^2).$$

As previously observed, it is sufficient to show that if T is a self-adjoint operator of trace zero and $\langle Tx_k, x_k\rangle = 0$ for all $k = 1, 2, \ldots$, then $T = 0$.

(1) $\Rightarrow$ (2): Let a vector

$$a = (a_{12}, \ldots, a_{1n}; a_{22}, \ldots, a_{2n}; \ldots; a_{(n-1)(n-1)}, a_{(n-1)n}; a_{nn}) \in \mathcal{K}$$

be such that $\langle a, \tilde{x}_k\rangle = 0$ for all k.

Define an operator $T = (b_{ij})_{i,j=1}^n$, where $b_{11} = -\sum_{i=2}^n a_{ii}$, $b_{ii} = a_{ii}$ for $i = 2, 3, \ldots, n$ and $b_{ij} = b_{ji} = \dfrac{1}{2}a_{ij}$ for $i < j$. Then T is self-adjoint and $\mathrm{tr}(T) = 0$.

For any $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ we have

$$\langle Tx, x\rangle = \sum_{i=1}^n \sum_{j=1}^n b_{ij}x_ix_j = \sum_{i=1}^n b_{ii}x_i^2 + 2\sum_{1\le i<j\le n} b_{ij}x_ix_j$$

$$= \left(-\sum_{i=2}^{n} a_{ii} \right) x_1^2 + \sum_{i=2}^{n} a_{ii} x_i^2 + \sum_{1 \le i < j \le n} a_{ij} x_i x_j$$
$$= \langle a, \tilde{x} \rangle.$$

Therefore, $\langle Tx_k, x_k \rangle = \langle a, \tilde{x}_k \rangle = 0$ for all k. This implies $T = 0$ and hence $a = 0$.

(2) $\Rightarrow$ (1): Let $T = (a_{ij})_{i,j=1}^n$ be a self-adjoint operator with $\mathrm{tr}(T) = 0$ and such that $\langle Tx_k, x_k \rangle = 0$ for all k. Then $a_{11} = -\sum_{i=2}^n a_{ii}$.

Define

$$\tilde{T} = (2a_{12}, 2a_{13}, \ldots, 2a_{1n}; a_{22}, 2a_{23}, \ldots, 2a_{2n}; \ldots; a_{(n-1)(n-1)}, 2a_{(n-1)n}; a_{nn}).$$

Then $\tilde{T} \in \mathcal{K}$ and

$$\langle \tilde{T}, \tilde{x}_k \rangle = \langle Tx_k, x_k \rangle = 0, \text{ for all } k.$$

Since $\{\tilde{x}_k\}_{k=1}^m$ spans $\mathcal{K}$, then $\tilde{T} = 0$. Hence $T = 0$. ∎

The complex case

To show the analogous result in the complex case we need to adjust the $\tilde{x}$ vector again so that the conclusion of Corollary 2.9 holds.

Definition 2.15. Given $x = (x_1, x_2, \ldots, x_n) \in \mathbb{C}^n$, define

$$\tilde{x} = (|x_1|^2, \mathrm{Re}(\bar{x}_1 x_2), \mathrm{Im}(\bar{x}_1 x_2), \ldots, \mathrm{Re}(\bar{x}_1 x_n), \mathrm{Im}(\bar{x}_1 x_n);$$
$$|x_2|^2, \mathrm{Re}(\bar{x}_2 x_3), \mathrm{Im}(\bar{x}_2 x_3), \ldots, \mathrm{Re}(\bar{x}_2 x_n), \mathrm{Im}(\bar{x}_2 x_n); \ldots;$$
$$|x|_{n-1}^2, \mathrm{Re}(\bar{x}_{n-1} x_n), \mathrm{Im}(\bar{x}_{n-1} x_n); |x_n|^2) \in \mathbb{R}^{n^2}.$$

Now we can give our classification theorem for injectivity in the quantum detection problem for the complex case.

Theorem 2.16. *Let $\mathcal{X} = \{x_k\}_{k=1}^m$ be a frame for $\mathbb{C}^n$. The following are equivalent:*

1. *$\mathcal{X}$ gives injectivity.*

2. *We have that $\{\tilde{x}_k\}_{k=1}^m$ spans $\mathbb{R}^{n^2}$.*

Proof. Given $x = (x_1, x_2, \ldots, x_n) \in \mathbb{C}^n$, define

$$\begin{aligned}\tilde{x} = (&|x_1|^2, \mathrm{Re}(\bar{x}_1 x_2), \mathrm{Im}(\bar{x}_1 x_2), \ldots, \mathrm{Re}(\bar{x}_1 x_n), \mathrm{Im}(\bar{x}_1 x_n);\\ &|x_2|^2, \mathrm{Re}(\bar{x}_2 x_3), \mathrm{Im}(\bar{x}_2 x_3), \ldots, \mathrm{Re}(\bar{x}_2 x_n), \mathrm{Im}(\bar{x}_2 x_n); \ldots;\\ &|x|_{n-1}^2, \mathrm{Re}(\bar{x}_{n-1} x_n), \mathrm{Im}(\bar{x}_{n-1} x_n); |x_n|^2) \in \mathbb{R}^{n^2}.\end{aligned}$$

(1) $\Rightarrow$ (2): Let a be any vector

$$\begin{aligned}a = (&a_{11}, u_{12}, v_{12}, \ldots, u_{1n}, v_{1n}; a_{22}, u_{23}, v_{23}, \ldots, u_{2n}, v_{2n}; \ldots;\\ &a_{(n-1)(n-1)}, u_{(n-1)n}, v_{(n-1)n}; a_{nn}) \in \mathbb{R}^{n^2}\end{aligned}$$

such that $\langle a, \tilde{x}_k \rangle = 0$ for all k.

Define an operator $T = (b_{ij})_{i,j=1}^n$ with $b_{ii} = a_{ii}$ for $i = 1, 2, \ldots, n$ and $b_{ij} = \bar{b}_{ji} = \frac{1}{2}(u_{ij} - \iota v_{ij})$ for $i < j$. Then T is a self-adjoint operator.

For any $x = (x_1, x_2, \ldots, x_n) \in \mathbb{C}^n$ we have

$$\begin{aligned}\langle Tx, x \rangle &= \sum_{i=1}^n \sum_{j=1}^n b_{ij} \bar{x}_i x_j\\ &= \sum_{i=1}^n b_{ii} |x_i|^2 + \sum_{1 \le i < j \le n} b_{ij} \bar{x}_i x_j + \sum_{1 \le j < i \le n} b_{ij} \bar{x}_i x_j\\ &= \sum_{i=1}^n b_{ii} |x_i|^2 + \sum_{1 \le i < j \le n} b_{ij} \bar{x}_i x_j + \sum_{1 \le j < i \le n} \bar{b}_{ji} \bar{x}_i x_j\\ &= \sum_{i=1}^n b_{ii} |x_i|^2 + \sum_{1 \le i < j \le n} b_{ij} \bar{x}_i x_j + \sum_{1 \le i < j \le n} \bar{b}_{ij} \bar{x}_j x_i\\ &= \sum_{i=1}^n b_{ii} |x_i|^2 + 2 \sum_{1 \le i < j \le n} \mathrm{Re}(b_{ij} \bar{x}_i x_j)\\ &= \sum_{i=1}^n b_{ii} |x_i|^2 + 2 \sum_{1 \le i < j \le n} (\mathrm{Re}(b_{ij}) \mathrm{Re}(\bar{x}_i x_j) - \mathrm{Im}(b_{ij}) \mathrm{Im}(\bar{x}_i x_j))\end{aligned}$$

$$= \sum_{i=1}^{n} a_{ii}|x_i|^2 + \sum_{1\le i<j\le n} (u_{ij}\operatorname{Re}(\bar{x}_i x_j) + v_{ij}\operatorname{Im}(\bar{x}_i x_j))$$
$$= \langle a, \tilde{x}\rangle.$$

Therefore, $\langle Tx_k, x_k\rangle = \langle a, \tilde{x}_k\rangle = 0$ for all k. This implies $T = 0$ and hence $a = 0$.

(2) $\Rightarrow$ (1): Let $T = (a_{ij})_{i,j=1}^n$ be a self-adjoint operator such that

$$\langle Tx_k, x_k\rangle = 0, \text{ for all } k.$$

Define

$$\tilde{T} = (a_{11}, 2\operatorname{Re}(a_{12}), -2\operatorname{Im}(a_{12}), \ldots, 2\operatorname{Re}(a_{1n}), -2\operatorname{Im}(a_{1n});$$
$$a_{22}, 2\operatorname{Re}(a_{23}), -2\operatorname{Im}(a_{23}), \ldots, 2\operatorname{Re}(a_{2n}), -2\operatorname{Im}(a_{2n}); \ldots;$$
$$a_{(n-1)(n-1)}, 2\operatorname{Re}(a_{(n-1)n}), -2\operatorname{Im}(a_{(n-1)n}); a_{nn}) \in \mathbb{R}^{n^2}.$$

Then we have

$$\langle \tilde{T}, \tilde{x}_k\rangle = \langle Tx_k, x_k\rangle = 0, \text{ for all } k.$$

Since $\{\tilde{x}_k\}_{k=1}^m$ spans $\mathcal{K}$ we have that $\tilde{T} = 0$ and so $T = 0$; i.e. $\{x_k\}_{k=1}^m$ gives injectivity.

■

Corollary 2.17. *If a frame $\mathcal{X} = \{x_k\}_{k=1}^m$ gives injectivity in $\mathbb{C}^n$, then $m \ge n^2$.*

Similar to the real case, we have a classification for injectivity for positive self-adjoint operators of trace one in a complex Hilbert space. This requires another definition of $\tilde{x}$ to fit this case and we will see that this requires one less measurement, as in the real case.

Theorem 2.18. *Let $\mathcal{X} = \{x_k\}_{k=1}^m$ be a frame for $\mathbb{C}^n$. The following are equivalent:*

1. *$\mathcal{X}$ gives injectivity for all self-adjoint operators of trace one.*

2. We have that $\{\tilde{x}_k\}_{k=1}^m$ spans $\mathbb{R}^{n^2-1}$.

Proof. Given $x = (x_1, x_2, \ldots, x_n) \in \mathbb{C}^n$, define

$$\begin{aligned}\tilde{x} = (&\operatorname{Re}(\bar{x}_1 x_2), \operatorname{Im}(\bar{x}_1 x_2), \ldots, \operatorname{Re}(\bar{x}_1 x_n), \operatorname{Im}(\bar{x}_1 x_n);\\ &|x_2|^2 - |x_1|^2, \operatorname{Re}(\bar{x}_2 x_3), \operatorname{Im}(\bar{x}_2 x_3), \ldots, \operatorname{Re}(\bar{x}_2 x_n), \operatorname{Im}(\bar{x}_2 x_n); \ldots;\\ &|x_{n-1}|^2 - |x_1|^2, \operatorname{Re}(\bar{x}_{n-1} x_n), \operatorname{Im}(\bar{x}_{n-1} x_n); |x_n|^2 - |x_1|^2) \in \mathbb{R}^{n^2-1}.\end{aligned}$$

(1) $\Rightarrow$ (2): Let a vector

$$\begin{aligned}a = (&u_{12}, v_{12}, \ldots, u_{1n}, v_{1n}; a_{22}, u_{23}, v_{23}, \ldots, u_{2n}, v_{2n}; \ldots;\\ &a_{(n-1)(n-1)}, u_{(n-1)n}, v_{(n-1)n}; a_{nn}) \in \mathbb{R}^{n^2-1}\end{aligned}$$

be such that $\langle a, \tilde{x}_k \rangle = 0$ for all k.

Define an operator $T = (b_{ij})_{i,j=1}^n$ with $b_{11} = -\sum_{i=2}^n a_{ii}$, $b_{ii} = a_{ii}$ for $i = 2, \ldots, n$ and $b_{ij} = \bar{b}_{ji} = \dfrac{1}{2}(u_{ij} - \iota v_{ij})$ for $i < j$. Then T is a self-adjoint operator and $\operatorname{tr}(T) = 0$.

For any $x = (x_1, x_2, \ldots, x_n) \in \mathbb{C}^n$ we have

$$\begin{aligned}\langle Tx, x \rangle &= \sum_{i=1}^n b_{ii}|x_i|^2 + 2 \sum_{1 \le i < j \le n} (\operatorname{Re}(b_{ij}) \operatorname{Re}(\bar{x}_i x_j) - \operatorname{Im}(b_{ij}) \operatorname{Im}(\bar{x}_i x_j))\\ &= \left(-\sum_{i=2}^n a_{ii} \right) |x_1|^2 + \sum_{i=2}^n a_{ii}|x_i|^2 + \sum_{1 \le i < j \le n} (u_{ij} \operatorname{Re}(\bar{x}_i x_j) + v_{ij} \operatorname{Im}(\bar{x}_i x_j))\\ &= \langle a, \tilde{x} \rangle.\end{aligned}$$

Therefore, $\langle Tx_k, x_k \rangle = \langle a, \tilde{x}_k \rangle = 0$ for all k. This implies $T = 0$ and hence $a = 0$.

(2) $\Rightarrow$ (1): Let $T = (a_{ij})_{i,j=1}^n$ be a self-adjoint operator such that $\operatorname{tr}(T) = 0$ and $\langle Tx_k, x_k \rangle = 0$ for all k. Then $a_{11} = -\sum_{i=2}^n a_{ii}$.

Define

$$\begin{aligned}\tilde{T} = (&2 \operatorname{Re}(a_{12}), -2 \operatorname{Im}(a_{12}), \ldots, 2 \operatorname{Re}(a_{1n}), -2 \operatorname{Im}(a_{1n});\end{aligned}$$

$$a_{22}, 2\,\mathrm{Re}(a_{23}), -2\,\mathrm{Im}(a_{23}), \ldots, 2\,\mathrm{Re}(a_{2n}), -2\,\mathrm{Im}(a_{2n}); \ldots;$$

$$a_{(n-1)(n-1)}, 2\,\mathrm{Re}(a_{(n-1)n}), -2\,\mathrm{Im}(a_{(n-1)n}); a_{nn}) \in \mathbb{R}^{n^2-1}.$$

Then we have that

$$\langle \tilde{T}, \tilde{x}_k \rangle = \langle Tx_k, x_k \rangle = 0, \text{ for all } k.$$

Since $\{\tilde{x}_k\}_{k=1}^m$ spans $\mathbb{R}^{n^2-1}$ then $\tilde{T} = 0$. Hence $T = 0$. ∎

Next we give second classification of injectivity for the quantum detection problem. This classification has the disadvantage that the requirements are not easily verified in practice. However, the advantage is that injective frames must satisfy the following:

Theorem 2.19. *Let $\mathcal{X} = \{x_k\}_{k=1}^m$ be a frame for a real or complex Hilbert space $\mathbb{H}^n$. The following are equivalent:*

1. *$\mathcal{X}$ gives injectivity.*

2. *For every orthonormal basis $\mathcal{E} = \{e_j\}_{j=1}^n$ for $\mathbb{H}^n$ we have:*

$$H(\mathcal{E}) =: \mathrm{span}\{(|\langle x_k, e_1\rangle|^2, |\langle x_k, e_2\rangle|^2, \ldots, |\langle x_k, e_n\rangle|^2) : k = 1, 2, \ldots, m\} = \mathbb{R}^n.$$

Proof. (1) ⇒ (2): We prove the contrapositive. Suppose that (2) fails. Then there is an orthonormal basis $\mathcal{E} = \{e_j\}_{j=1}^n$ so that $H(\mathcal{E}) \neq \mathbb{R}^n$. Hence there is a non-zero vector $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{R}^n$ such that $\lambda \perp H(\mathcal{E})$.

Define an operator on $\mathbb{H}^n$ by

$$Te_j = \lambda_j e_j, j = 1, 2, \ldots, n.$$

Then T is a non-zero self-adjoint operator and satisfies $\langle Tx_k, x_k \rangle = 0$, for all $k = 1, 2, \ldots, m$, which is a contradiction.

(2) $\Rightarrow$ (1): Let T be a self-adjoint operator such that $\langle Tx_k, x_k\rangle = 0$, for all k. Let $\mathcal{E} = \{e_j\}_{j=1}^n$ be an eigenbasis for T with respective eigenvalues $\{\lambda_j\}_{j=1}^n$. Then for every $k = 1, 2, \ldots, m$ we have

$$\langle Tx_k, x_k\rangle = \sum_{j=1}^{n} \lambda_j |\langle x_k, e_j\rangle|^2 = 0.$$

That is,

$$(\lambda_1, \lambda_2, \ldots, \lambda_n) \perp H(\mathcal{E}) = \mathbb{R}^n \text{ by assumption (2).}$$

Therefore, $\lambda_j = 0$ for all $j = 1, 2, \ldots, n$ and so $T = 0$. ∎

2.1.1 Constructing the Solutions to the Injectivity Problem

Using the classifications given we will construct large classes of frames which give injectivity for the quantum detection problem (in both $\mathbb{R}$ and $\mathbb{C}$).

Theorem 2.20. *Let $\{x_k\}_{k=1}^n$ be a linearly independent set in $\mathbb{R}^n$ such that the first coordinates of these vectors are non-zero. Now choose $(n-1)$ linearly independent vectors $\{x_k\}_{k=n+1}^{2n-1}$ in $\mathbb{R}^n$ such that each vector is zero in the first coordinate and is non-zero in the second coordinate. Continuing this procedure we get a frame $\{x_k\}_{k=1}^{\frac{n(n+1)}{2}}$ which gives injectivity.*

Proof. We will show that $\{\tilde{x}_k\}_{k=1}^{\frac{n(n+1)}{2}}$ is a basis for $\mathbb{R}^{\frac{n(n+1)}{2}}$. Suppose that $\sum_{k=1}^{\frac{n(n+1)}{2}} \alpha_k \tilde{x}_k = 0$ for some scalars $\{\alpha_k\}$. Since for $n+1 \leq k \leq \frac{n(n+1)}{2}$, $\tilde{x}_k$ are zero in the first coordinate, we get

$$\sum_{k=1}^{n} \alpha_k x_{k1} x_k = 0.$$

Since $\{x_k\}_{k=1}^n$ are linearly independent, $\alpha_k x_{k1} = 0$ for all k and since $x_{k1} \neq 0$, $\alpha_k = 0$ for $k = 1, 2, \ldots, n$.

Now do this argument for the next $(n-1)$ vectors and continue we get $\alpha_k = 0$ for all $k = 1, 2, \ldots, \frac{n(n+1)}{2}$. ∎

The following example satisfies the previous construction.

Example 2.21. The frame

$$\{e_i\}_{i=1}^n \cup \{e_i + e_j : i < j\}_{i,j=1}^n$$

gives injectivity.

For the complex case, we have the following construction. The proof is as in the real case.

Theorem 2.22. *Let $\{x_k\}_{k=1}^{2n-1}$ be a basis for $\mathbb{R}^{2n-1}$, where*

$$x_k = (u_{k1}, u_{k2}, v_{k2}, \ldots, u_{kn}, v_{kn})$$

and $u_{k1} \neq 0$, $k = 1, \ldots, 2n-1$.

Define $(2n-1)$ vectors $\{z_k\}_{k=1}^{2n-1}$ in $\mathbb{C}^n$ by

$$z_k = (u_{k1}, u_{k2} + \iota v_{k2}, \ldots, u_{kn} + \iota v_{kn}).$$

Now let $\{x_k\}_{k=2n}^{4n-4}$ be a basis for $\mathbb{R}^{2n-3}$, where

$$x_k = (u_{k2}, u_{k3}, v_{k3}, \ldots, u_{kn}, v_{kn})$$

and $u_{k2} \neq 0$, $k = 2n, \ldots, 4n-4$.

Define $(2n-3)$ vectors $\{z_k\}_{k=2n}^{4n-4}$ in $\mathbb{C}^n$ by

$$z_k = (0, u_{k2}, u_{k3} + \iota v_{k3}, \ldots, u_{kn} + \iota v_{kn}).$$

Continuing this procedure we get n^2 vectors $\{z_k\}_{k=1}^{n^2}$ in $\mathbb{C}^n$ and they give injectivity.

As we have seen, we can get Parseval frames giving injectivity by taking $\{S^{-1/2}x_k\}_{k=1}^m$, where $\{x_k\}_{k=1}^m$ gives injectivity and has frame operator S. However, the construction above can be adjusted to directly construct Parseval frames giving injectivity.

Theorem 2.23. *Let $\{\lambda_{ij}\}_{i=1,j=i}^{n\ \ n}$ be non-negative numbers satisfying:*

1. *$\lambda_{ij} = 0$ if and only if $j < i$.*

2. *For each $j = 1, 2, \ldots, n$ we have $\sum_{i=1}^n \lambda_{ij} = 1$.*

Let $\mathcal{E} = \{e_j\}_{j=1}^n$ be the canonical basis of $\mathbb{R}^n$. Let $\{x_k\}_{k=1}^{\frac{n(n+1)}{2}}$ be vectors in $\mathbb{R}^n$ which satisfy:

1. *$\{x_k\}_{k=1}^n$ is a linearly independent set with $x_{k1} \neq 0$ for all $k = 1, \ldots, n$ and it has frame operator S_1 with eigenvectors $\mathcal{E}$ and respective eigenvalues $\{\lambda_{1j}\}_{j=1}^n$ (See [15].)*

2. *$\{x_k\}_{k=n+1}^{2n-1}$ is a linearly independent set with $x_{k1} = 0$, for all k, $x_{k2} \neq 0$ for all k, and it has frame operator S_2 with eigenvectors $\mathcal{E}$ and respective eigenvalues $\{\lambda_{2j}\}_{j=1}^n$.*

3. *continue.*

Then the vectors $\{x_k\}_{k=1}^{\frac{n(n+1)}{2}}$ form a Parseval frame for $\mathbb{R}^n$ which is injective.

Proof. This is injective by Theorem 2.20. To see that it is Parseval, observe that the frame operator of this frame is $\sum_{i=1}^n S_i$. Now, let $y \in \mathbb{R}^n$ and compute:

$$\sum_{i=1}^n S_i y = \sum_{i=1}^n S_i \left(\sum_{j=1}^n \langle y, e_j \rangle e_j \right) = \sum_{i=1}^n \sum_{j=1}^n \langle y, e_j \rangle S_i e_j$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \langle y, e_j \rangle \lambda_{ij} e_j = \sum_{j=1}^{n} \langle y, e_j \rangle e_j \sum_{i=1}^{n} \lambda_{ij}$$
$$= \sum_{j=1}^{n} \langle y, e_j \rangle e_j = y.$$

■

Similarly, we have the following theorem for the complex case.

Theorem 2.24. *Fix $\{\lambda_{ij}\}_{i=1,j=i}^{n\ n}$ be non-negative numbers satisfying:*

1. *$\lambda_{ij} = 0$ if and only if $j < i$.*

2. *For each $j = 1, 2, \ldots, n$ we have $\sum_{i=1}^{n} \lambda_{ij} = 1$.*

Let $\mathcal{E} = \{e_j\}_{j=1}^{n}$ be the canonical basis of $\mathbb{C}^n$. Let $\{z_k\}_{k=1}^{n^2}$ be vectors in $\mathbb{C}^n$ which satisfy:

1. *For each $k = 1, \ldots, 2n-1$, z_k has the form*

$$z_k = (u_{k1}, u_{k2} + \iota v_{k2}, \ldots, u_{kn} + \iota v_{kn}),$$

where $u_{k1} \neq 0$ and the set $\{(u_{k1}, u_{k2}, v_{k2}, \ldots, u_{kn}, v_{kn})\}_{k=1}^{2n-1}$ is linearly independent in $\mathbb{R}^{2n-1}$. Moreover $\{z_k\}_{k=1}^{2n-1}$ has frame operator S_1 with eigenvectors $\mathcal{E}$ and respective eigenvalues $\{\lambda_{1j}\}_{j=1}^{n}$.

2. *For each $k = 2n, \ldots, 4n-4$, z_k has the form*

$$z_k = (0, u_{k2}, u_{k3} + \iota v_{k3}, \ldots, u_{kn} + \iota v_{kn}),$$

where $u_{k2} \neq 0$ and the set $\{(u_{k2}, u_{k3}, v_{k3}, \ldots, u_{kn}, v_{kn})\}_{k=2n}^{4n-4}$ is linearly independent in $\mathbb{R}^{2n-3}$. Moreover $\{z_k\}_{k=2n}^{4n-4}$ has frame operator S_2 with eigenvectors $\mathcal{E}$ and respective eigenvalues $\{\lambda_{2j}\}_{j=1}^{n}$.

3. continue.

Then the vectors $\{z_k\}_{k=1}^{n^2}$ *form a Parseval frame for* $\mathbb{C}^n$ *which is injective.*

By varying the above construction we can find frames which give injectivity and have prescribed eigenvalues for their frame operators. Another class of examples arise in the form of mutually unbiased bases. Recall the definition of mutually unbiased bases:

Definition 2.25. Two orthonormal bases $\{x_k\}_{k=1}^n$ and $\{y_k\}_{k=1}^n$ are **mutually unbiased** if

$$|\langle x_k, y_j\rangle| = \frac{1}{\sqrt{n}}, \text{ for all } i, j = 1, 2, \ldots, n.$$

A family of orthonormal bases is **mutually unbiased** if each pair is mutually unbiased.

It is known that the maximal number of mutually unbiased bases in $\mathbb{H}^n$ is n+1 and this is rarely achieved. It holds if $n = p^m$ for a prime p. It is observed in [35] and [7] that a maximal family of mutually unbiased bases will give injectivity in the quantum detection problem.

2.1.2 The Solutions are Open and Dense

In this section we will show that the family of m-element frames which solve the quantum detection injectivity problem is open and dense in the family of all m-element frames. For this, we need to measure the distance between m-element frames using the standard metric.

Definition 2.26. Given frames $\mathcal{X} = \{x_k\}_{k=1}^m$ and $\mathcal{Y} = \{y_k\}_{k=1}^m$ for a real or complex Hilbert space $\mathbb{H}^n$, the **distance** between them is

$$d(\mathcal{X}, \mathcal{Y})^2 = \sum_{k=1}^{m} \|x_k - y_k\|^2.$$

Note this metric is the finite version of Definition 1.42.

Theorem 2.27. *The set of all m-element frames on $\mathbb{H}^n$ that give injectivity in the frame quantum detection problem is open and dense in the space of all m-element frames on $\mathbb{H}^n$.*

Proof. Let a frame $\{x_k\}_{k=1}^{\frac{n(n+1)}{2}} \subset \mathbb{R}^n$ give injectivity. By Theorem 2.10, this is equivalent to the determinant of the matrix whose rows are $\tilde{x}_k$, for $k = 1, 2, \ldots, \frac{n(n+1)}{2}$, being non-zero.

The determinant of this matrix is a polynomial of $\frac{n^2(n+1)}{2}$ variables x_{ki} for $1 \leq k \leq \frac{n(n+1)}{2}$ and $1 \leq i \leq n$. Since the complement of the zero set of this polynomial is dense in $\mathbb{R}^{\frac{n^2(n+1)}{2}}$, the set of all $\frac{n(n+1)}{2}$-element frames which give injectivity is dense in the space of all $\frac{n(n+1)}{2}$-element frames on $\mathbb{R}^n$.

Now let any m-element frame $\{x_k\}_{k=1}^m$ in $\mathbb{R}^n$ with $m \geq \frac{n(n+1)}{2}$ and $\delta > 0$. Then there exists a subframe containing $\frac{n(n+1)}{2}$ vectors. We can assume that this subframe is $\{x_k\}_{k=1}^{\frac{n(n+1)}{2}}$. By denseness above, there is an injective frame $\{y_k\}_{k=1}^{\frac{n(n+1)}{2}}$ such that

$$\sum_{k=1}^{n(n+1)/2} \|x_k - y_k\|^2 < \delta.$$

Now define a new frame $\{\phi_k\}_{k=1}^m$, where $\phi_k = y_k$ for $k = 1, \ldots, \frac{n(n+1)}{2}$ and $\phi_k = x_k$ for $k > \frac{n(n+1)}{2}$. Then the frame $\{\phi_k\}_{k=1}^m$ is injective and

$$\sum_{k=1}^{m} \|x_k - \phi_k\|^2 < \delta.$$

This completes the proof for the real case. In the real case it is known that the complement of the zero set of a nontrivial polynomial of n variables is dense in $\mathbb{R}^n$. In the complex case, we see that given a polynomial $P(z_1, ..., z_n)$ on $\mathbb{C}^n$, we may write P as

$$P'(x_1, y_1, ..., x_n, y_n) + \iota P''(x_1, y_1, ..., x_n, y_n)$$

where $z_j = x_j + \iota y_j$. Hence P' and P'' are polynomials on $\mathbb{R}^{2n}$. P has a zero if and only if P' and P'' have a common zero. We see that the complement of the intersection of the zero sets of P' and P'' is dense in $\mathbb{R}^{2n}$ and hence is dense in $\mathbb{C}^n$ after natural identification of $\mathbb{R}^{2n}$ with $\mathbb{C}^n$. The complex case follows similarly. ∎

Theorem 2.28. *The family of all m-element frames on $\mathbb{H}^n$ that give injectivity in the frame quantum detection problem is open in the space of all m-element frames on $\mathbb{H}^n$.*

Proof. As above we will prove the real case and the complex case follows similarly.

Denote by $\mathcal{F}$ the space of all $\frac{n(n+1)}{2}$-element frames for $\mathbb{R}^n$. Consider the map:

$$f : \mathcal{F} \longrightarrow \mathbb{R}$$

$$\mathcal{X} = \{x_k\}_{k=1}^{\frac{n(n+1)}{2}} \longmapsto f(\mathcal{X}) = \det\{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_{\frac{n(n+2)}{2}}\}.$$

Then f is a continuous function. Since $f^{-1}(0)$ is a closed set, by Theorem 2.10, the set of all $\frac{n(n+1)}{2}$-element frames is open in $\mathcal{F}$.

Now let $\mathcal{X} = \{x_k\}_{k=1}^{m}$ in $\mathbb{R}^n, (m \geq \frac{n(n+1)}{2})$ be an m-element frame which gives injectivity. Then there is a subframe $\mathcal{Y}$ containing $\frac{n(n+1)}{2}$ vectors, which is also injec-

tive. Therefore, there exists $\epsilon > 0$ such that every $\frac{n(n+1)}{2}$-element frame in the ball $B(\mathcal{Y}, \epsilon)$ is injective. This implies that every m-element frame in the ball $B(\mathcal{X}, \epsilon)$ is also injective. ∎

To show that the Parseval frames giving injectivity in the quantum detection problem are dense in the Parseval frames, we will first prove a very general problem about frames.

Theorem 2.29. *Let $\mathcal{P}$ be a property of Hilbert space frames and assume:*

1. *The set of all m-element frames in $\mathbb{H}^n$ having property $\mathcal{P}$ is dense in the set of all m-element frames.*

2. *If a frame $\{x_k\}_{k=1}^m$ with frame operator S has property $\mathcal{P}$, then $\{S^{-1/2}x_k\}_{k=1}^m$ has property $\mathcal{P}$.*

Then the set of all m-element Parseval frames with property $\mathcal{P}$ is dense in the set of all m-element Parseval frames.

Proof. Fix $\epsilon > 0$ and let $\delta > 0$ so that

$$2m\delta^2 + 8(m\delta)^2 m(1+\delta)^2 < \epsilon,\ 2m\delta < 1.$$

Let $\{x_k\}_{k=1}^m$ be any Parseval frame for $\mathbb{H}^n$. By denseness, we can choose a frame $\{y_k\}_{k=1}^m$ having property $\mathcal{P}$ and satisfying $\|x_k - y_k\| \leq \delta$, for all $k = 1, 2, \ldots, m$. Since $\|x_k\| \leq 1$, we have that $\|y_k\| \leq 1 + \delta$. Let S_1 be the frame operator of $\{y_k\}_{k=1}^m$. Then,

$$\begin{aligned}\langle S_1 x, x\rangle^{1/2} &= \left(\sum_{k=1}^m |\langle x, y_k\rangle|^2\right)^{1/2} \\ &\leq \left(\sum_{k=1}^m |\langle x, x_k\rangle|^2\right)^{1/2} + \left(\sum_{k=1}^m |\langle x, x_k - y_k\rangle|^2\right)^{1/2}\end{aligned}$$

$$\leq \|x\| + \|x\| \left(\sum_{k=1}^{m} \|x_k - y_k\|^2 \right)^{1/2}$$
$$\leq \|x\|(1 + m\delta).$$

Therefore

$$\sum_{k=1}^{m} |\langle x, y_k \rangle|^2 \leq \|x\|^2 (1 + m\delta)^2.$$

Similarly,

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

I.e. $(1 - m\delta)^2 I \leq S_1 \leq (1 + m\delta)^2 I$. Hence, $(1 - m\delta)I \leq S_1^{1/2} \leq (1 + m\delta)I$ and so $(1 + m\delta)^{-1} I \leq S_1^{-1/2} \leq (1 - m\delta)^{-1} I$. Finally,

$$I - (1 - m\delta)^{-1} I \leq I - S_1^{-1/2} \leq I - (1 + m\delta)^{-1} I,$$

and so

$$-2m\delta I \leq \frac{-m\delta}{1 - m\delta} I \leq I - S_1^{-1/2} \leq \frac{m\delta}{1 + m\delta} I \leq 2m\delta I.$$

Now, $\{S_1^{-1/2} y_k\}_{k=1}^m$ is a Parseval frame with property $\mathcal{P}$ and

$$\sum_{k=1}^{m} \|x_k - S_1^{-1/2} y_k\|^2 \leq 2 \sum_{k=1}^{m} \|x_k - y_k\|^2 + 2 \sum_{k=1}^{m} \|(I - S_1^{-1/2}) y_k\|^2$$
$$\leq 2m\delta^2 + 2 \sum_{k=1}^{m} (2m\delta)^2 \|y_k\|^2$$

$$\leq 2m\delta^2 + 8(m\delta)^2 m(1 + \delta)^2 < \epsilon.$$

■

Corollary 2.30. *The set of all m-element Parseval frames which give injectivity is dense in the set of all m-element Parseval frames.*

2.1.3 Solution to the State Estimation Problem

In this section we will give a classification of injective Parseval frames for which the state estimation problem is solvable. Recall that for an injective Parseval frame $\{x_k\}_{k\in I}$ and $\beta = 2^I$, the map $\mathbb{M}$ (which maps a quantum state $T \in \mathcal{S}(\mathbb{H})$ to a function $p \in L(\beta, \mathbb{R})$) is injective. Given a function $p \in L(\beta, \mathbb{R})$, if $p = \mathbb{M}(T)$ for some $T \in \mathcal{S}(\mathbb{H})$, then for any $U \in \beta$, we must have

$$p(U) = \mathbb{M}(T)(U) = \sum_{k\in U} \langle Tx_k, x_k\rangle = \sum_{k\in U} p(\{k\}).$$

Thus, p must be additive and is determined by its value at the singleton sets $\{k\}$ for all $k \in I$. Therefore, for the state estimation problem in the finite case, we will ask:

The State Estimation Problem: Given an injective Parseval frames $\{x_k\}_{k=1}^m$ on $\mathbb{H}^n$ and a measurement vector $a = (a_1, a_2, \ldots, a_m) \in \mathbb{R}^m$, can we find a positive self-adjoint trace one operator T so that

$$\langle Tx_k, x_k\rangle = a_k, \text{ for all } k?$$

As before, we will not require the operator T of the problem to be positive and trace one. This will be considered as a special case of the problem. Hence, we will say that the state estimation problem is solvable if there exists a self-adjoint operator T so that

$$\langle Tx_k, x_k\rangle = a_k, \text{ for all } k.$$

We will give a complete classification of injective Parseval frames for which the state estimation problem is solvable.

Theorem 2.31. *Let $\mathcal{X} = \{x_k\}_{k=1}^m$ be an injective Parseval frame for $\mathbb{R}^n$, and $a = (a_1, a_2, \ldots, a_m) \in \mathbb{R}^m$, the following are equivalent:*

1. *The state estimation problem is solvable.*
2. $\operatorname{rank}(A) = \operatorname{rank}(B)$, *where A is a matrix whose the k^{th}-row is $\tilde{x}_k$, and $B = [A, a]$.*

Proof. For a vector $x \in \mathbb{R}^n$, the vector $\tilde{x}$ is defined as in the Definition 2.8. Note that a self-adjoint operator T is determined by the values $\langle Te_i, e_j\rangle$ for all $i \leq j$. Then the state estimation problem is solvable if and only if there exists a self-adjoint operator T so that for every k

$$a_k = \langle Tx_k, x_k\rangle = \left\langle T(\sum_{i=1}^n \langle x_k, e_i\rangle e_i), \sum_{j=1}^n \langle x_k, e_j\rangle e_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \langle x_k, e_i\rangle \langle x_k, e_j\rangle \langle Te_i, e_j\rangle.$$

This is equivalent to the linear system with unknowns $\langle Te_i, e_j\rangle$:

$$\sum_{i=1}^n x_{ki}^2 \langle Te_i, e_i\rangle + 2\sum_{i<j} x_{ki}x_{kj}\langle Te_i, e_j\rangle = a_k,\ k = 1, 2, \ldots, m$$

having a solution, and hence is equivalent to $\operatorname{rank}(A) = \operatorname{rank}(B)$. ∎

In the case where the number of frame vectors equals $\frac{n(n+1)}{2}$, we have a unique solution to the state estimation problem.

Corollary 2.32. *Let $\mathcal{X} = \{x_k\}_{k=1}^{\frac{n(n+1)}{2}} \subset \mathbb{R}^n$ be an injective Parseval frame. Then the state estimation problem has a unique solution for all choices of vectors $a = (a_1, a_2, \ldots, a_{\frac{n(n+1)}{2}})$.*

Proof. By Theorem 2.10, $\mathcal{X}$ is injective is equivalent to $\{\tilde{x}_k\}_{k=1}^{\frac{n(n+1)}{2}}$ is linearly independent. Hence

$$\operatorname{rank} A = \operatorname{rank} B = \frac{n(n+1)}{2}.$$

The conclusion then follows by Theorem 2.31. ∎

For the completeness of the state estimation problem, we will state the classification in the case that the operator T is required to be positive, self-adjoint operator of trace one. First, we need to recall *Sylvester's Criterion* [29].

Theorem 2.33. *A self-adjoint matrix T is positive if and only if all of its principal minors are nonnegative.*

Now we have the following classification:

Theorem 2.34. *Let $\mathcal{X} = \{x_k\}_{k=1}^m$ be an injective Parseval frame for $\mathbb{R}^n$, and $a = (a_1, a_2, \ldots, a_m) \in \mathbb{R}^m$, the following are equivalent:*

1. *The state estimation problem is solvable for a positive, self-adjoint operator of trace one.*

2. *The linear system*

$$\sum_{i=1}^{n} x_{ki}^2 \langle Te_i, e_i\rangle + 2\sum_{i<j} x_{ki}x_{kj}\langle Te_i, e_j\rangle = a_k,\ k = 1, 2, \ldots, m$$

has a solution $\{\langle Te_i, e_j\rangle : i \leq j\}$, which determines a self-adjoint matrix T such that all of its principal minors are nonnegative, and $\sum_{i=1}^n \langle Te_i, e_i\rangle = 1$.

It should be noted that all of the theorems above still hold for the complex case with the corresponding $\tilde{x}_k$, defined as in Definition 2.15. We state one of them here, the other are similar to the real case.

Theorem 2.35. *Let $\mathcal{X} = \{x_k\}_{k=1}^m$ be an injective Parseval frame for $\mathbb{C}^n$, and $a = (a_1, a_2, \ldots, a_m) \in \mathbb{R}^m$, the following are equivalent:*

1. *The state estimation problem is solvable.*

2. $\operatorname{rank}(A) = \operatorname{rank}(B)$, *where* A *is a matrix whose the* k*-row is* $\tilde{x}_k$, *and* $B = [A, a]$.

Proof. In the complex case, a self-adjoint operator T is determined by the values of the real part and imaginary part of $\langle Te_j, e_i\rangle$ for all $i \leq j$. Then the state estimation problem is solvable if and only if there exists a self-adjoint operator T so that

$$\begin{aligned} a_k &= \langle Tx_k, x_k\rangle \\ &= \langle T(\sum_{i=1}^{n}\langle x_k, e_i\rangle e_i), \sum_{j=1}^{n}\langle x_k, e_j\rangle e_j\rangle \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\langle x_k, e_i\rangle\overline{\langle x_k, e_j\rangle}\langle Te_i, e_j\rangle \\ &= \sum_{i=1}^{n}|x_{ki}|^2\langle Te_i, e_i\rangle + 2\sum_{i<j}[\operatorname{Re}(\bar{x}_{ki}x_{kj})\operatorname{Re}\langle Te_j, e_i\rangle - \operatorname{Im}(\bar{x}_{ki}x_{kj})\operatorname{Im}\langle Te_j, e_i\rangle] \end{aligned}$$

for all k.

This is equivalent to the following linear system:

$$\sum_{i=1}^{n}|x_{ki}|^2\langle Te_i, e_i\rangle + 2\sum_{i<j}[\operatorname{Re}(\bar{x}_{ki}x_{kj})\operatorname{Re}\langle Te_j, e_i\rangle - \operatorname{Im}(\bar{x}_{ki}x_{kj})\operatorname{Im}\langle Te_j, e_i\rangle] = a_k,$$

$k = 1, 2, \ldots, m$ with unknowns $\operatorname{Re}\langle Te_j, e_i\rangle, \operatorname{Im}\langle Te_j, e_i\rangle, i \leq j$ having a solution, and hence is equivalent to $\operatorname{rank}(A) = \operatorname{rank}(B)$. ∎

If a frame $\{x_k\}_{k=1}^m$ has $m > \frac{n(n+1)}{2}$ in the real case (or $m > n^2$ in the complex case) it is unlikely the state estimation is solvable because of redundancy. This is because if two of the x_k are equal while the corresponding a_k are not, then the problem is not solvable. More generally, if some of the x_k are linearly dependent then at least one of the corresponding a_k is uniquely determined. However, in this case there is a natural way to find the best estimate for the problem. We consider the real case. Note

that there always exists a subset $I \subset \{1, 2, \ldots, m\}$ of size $\frac{n(n+1)}{2}$, and a self-adjoint operator T so that $\langle Tx_k, x_k \rangle = a_k$, for all $k \in I$. Therefore, if the state estimation problem is not solvable, it is natural to find a T that best approximates the solution. That is, we minimize the distance to the measurement vector a using the following:

$$\sum_{k=1}^{m} |\langle Tx_k, x_k \rangle - a_k|^2$$

To do this, let $\mathcal{S}$ be the set of all bases of $\mathbb{R}^{\frac{n(n+1)}{2}}$ that are subsets of $\{\tilde{x}_k\}_{k=1}^m$. This set is obviously finite. Since each element $\{\tilde{x}_k\}_{k \in I}$ in $\mathcal{S}$ determines a unique self-adjoint operator T satisfying $\langle Tx_k, x_k \rangle = a_k$, for all $k \in I$, we can find the quantum state T that gives the best approximation to the measurement vector a by choosing the set which minimizes the distance above.

2.2 The Infinite Dimensional Case

In infinite dimensions we will work with both the trace class operators and Hilbert-Schmidt operators (i.e. operators $T = (a_{ij})_{i,j=1}^{\infty}$ with $\sum_{i,j=1}^{\infty} |a_{ij}|^2 < \infty$). This class contains the trace class operators. As in the finite case, we will solve the following frame injectitivity problem:

Injectivity Problem: For what frames $\{x_k\}_{k=1}^{\infty}$ in real or complex infinite dimensional Hilbert space $\mathbb{H}$ do we have the property: Whenever T, S are Hilbert Schmidt positive self-adjoint operators on $\mathbb{H}$ and $\langle Tx_k, x_k \rangle = \langle Sx_k, x_k \rangle$, for all $k = 1, 2, \ldots$, then $T = S$.

We will not require our operators to be trace class and trace one. These requirements will be considered as a special case of our problem.

2.2.1 The Solution to the Injectivity Problem

In this subsection we will solve the injectivity problem for infinite dimensional Hilbert spaces. Similar to the finite case, we first show that we only need to work with self adjoint operators. In Theorem 2.6, the "(1) implies (2)" direction does not hold for the infinite case. So we provide another proof here, but the other implications are as in the finite case.

Theorem 2.36. *Given a family of vectors $\mathcal{X} = \{x_k\}_{k=1}^{\infty}$ in a real or complex Hilbert space $\mathbb{H}$, the following are equivalent:*

1. *Whenever T, S are Hilbert Schmidt, positive and self-adjoint, and*

$$\langle Tx_k, x_k\rangle = \langle Sx_k, x_k\rangle, \text{ for all } k,$$

   then $T = S$.

2. *Whenever T, S are Hilbert Schmidt self-adjoint, and*

$$\langle Tx_k, x_k\rangle = \langle Sx_k, x_k\rangle, \text{ for all } k,$$

   then $T = S$.

3. *$\mathcal{X}$ is injective.*

Proof. We will show that (1) implies (2). Let T, S be Hilbert Schmidt self-adjoint operators such that

$$\langle Tx_k, x_k\rangle = \langle Sx_k, x_k\rangle, \text{ for all } k.$$

Set $R = T - S$. Then R is also a Hilbert Schmidt self-adjoint operator. Let $\{e_j\}_{j=1}^{\infty}$ be an orthonormal basis for $\mathbb{H}$ and let $\{u_j\}_{j=1}^{\infty}$ be an eigenbasis for R with respective

eigenvalues $\{\lambda_j\}_{j=1}^{\infty}$. Define operators U and D on $\mathbb{H}$ by $Ue_j = u_j$ and $De_j = \lambda_j e_j$, for $j = 1, 2, \ldots$. Then U is a unitary operator, D is Hilbert Schmidt self-adjoint operator, and

$$R = UDU^*.$$

Now let $r_j = |\lambda_j|, s_j = |\lambda_j| - \lambda_j, j = 1, 2, \ldots$ be non-negative numbers. Then $\lambda_j = r_j - s_j$. Let D_1, D_2 be operators defined by

$$D_1 e_j = r_j e_j,\ D_2 e_j = s_j e_j \text{ for } j = 1, 2, \ldots.$$

Note that since R is Hilbert Schmidt, $\sum_{j=1}^{\infty} \lambda_j^2$ converges. Hence D_1, D_2 are Hilbert-Schmidt positive self-adjoint and we have

$$R = UDU^* = U(D_1 - D_2)U^* = UD_1U^* - UD_2U^*.$$

Moreover, $UD_1U^*,\ UD_2U^*$ are Hilbert Schmidt positive self-adjoint operators. Since

$$0 = \langle Rx_k, x_k\rangle = \langle UD_1U^*x_k, x_k\rangle - \langle UD_2U^*x_k, x_k\rangle,$$

we have that $UD_1U^* = UD_2U^*$. Thus, $R = 0$ and hence $T = S$. ∎

If our operators are trace class, then we will have the following theorem. The proof of Theorem 2.36 is still valid here by noticing that $\sum_{j=1}^{\infty} |\lambda_j| < \infty$ for the trace class operator R.

Theorem 2.37. *Given a family of vectors $\mathcal{X} = \{x_k\}_{k=1}^{\infty}$ in a infinite dimensional Hilbert space $\mathbb{H}$, the following are equivalent:*

1. *Whenever T, S are trace class positive and self-adjoint, and*

   $$\langle Tx_k, x_k\rangle = \langle Sx_k, x_k\rangle,\ \textit{for all } k,$$

   then $T = S$.

2. *Whenever T, S are trace class self-adjoint, and*

$$\langle Tx_k, x_k\rangle = \langle Sx_k, x_k\rangle, \text{ for all } k,$$

then $T = S$.

3. *Whenever T is trace class self-adjoint, and*

$$\langle Tx_k, x_k\rangle = 0, \text{ for all } k,$$

then $T = 0$.

Similar to the finite case, we will first give the following classification of injectivity for Hilbert Schmidt operators.

Theorem 2.38. *Let $\mathcal{X} = \{x_k\}_{k=1}^{\infty}$ be a frame for an infinite dimensional real or complex Hilbert space $\mathbb{H}$. The following are equivalent:*

1. *$\mathcal{X}$ is injective.*

2. *For every orthonormal basis $\mathcal{E} = \{e_j\}_{j=1}^{\infty}$ for $\mathbb{H}$ we have:*

$$H(\mathcal{E}) =: \overline{\text{span}}\{(|\langle x_k, e_1\rangle|^2, |\langle x_k, e_2\rangle|^2, \ldots) : k = 1, 2, \ldots\} = \ell_2.$$

Proof. (1) $\Rightarrow$ (2): We prove the result by way of contradiction. Suppose that (2) is false. Then there is an orthonormal basis $\mathcal{E} = \{e_j\}_{j=1}^{\infty}$ so that $H(\mathcal{E}) \neq \ell_2$. Hence there is a non-zero vector $\lambda = (\lambda_1, \lambda_2, \ldots) \in \ell_2$ such that $\lambda \perp H(\mathcal{E})$.

Define an operator on $\mathbb{H}$ by

$$Te_j = \lambda_j e_j, \text{ for all } j = 1, 2, \ldots.$$

Then T is a non-zero Hilbert Schmidt operator. We also have: $\langle Tx_k, x_k\rangle = 0$, for all $k = 1, 2, \ldots$. This is a contradiction.

(2) $\Rightarrow$ (1): Let T be a Hilbert Schmidt self-adjoint operator such that

$$\langle Tx_k, x_k\rangle = 0, \text{ for all } k.$$

Since T is Hilbert Schmidt and hence compact, there is an eigenbasis $\mathcal{E} = \{e_j\}_{j=1}^{\infty}$ for T with respective eigenvalues $\{\lambda_j\}_{j=1}^{\infty}$. Then for every $k = 1, 2, \ldots$, we have

$$\langle Tx_k, x_k\rangle = \sum_{j=1}^{\infty} \lambda_j |\langle x_k, e_j\rangle|^2 = 0.$$

Since T is Hilbert Schmidt then

$$\sum_{j=1}^{\infty} |\lambda_j|^2 = \sum_{j=1}^{\infty} \|Te_j\|^2 < \infty.$$

That is, $(\lambda_1, \lambda_2, \ldots) \in \ell_2$. Since

$$(\lambda_1, \lambda_2, \ldots) \perp H(\mathcal{E}) = \ell_2 \text{ by assumption (2).}$$

Therefore, $\lambda_j = 0$ for all $j = 1, 2, \ldots$, and so $T = 0$. ∎

If we consider operators which are trace class, then we have the following classification for the infinite dimensions.

Theorem 2.39. *Let $\mathcal{X} = \{x_k\}_{k=1}^{\infty}$ be a frame for an infinite dimensional real or complex Hilbert space $\mathbb{H}$. The following are equivalent:*

1. *$\mathcal{X}$ satisfies the following property: The only trace class self-adjoint operator T such that*

   $$\langle Tx_k, x_k\rangle = 0, \text{ for all } k,$$

   is $T = 0$.

2. *For every $\lambda = (\lambda_1, \lambda_2, \ldots) \in \ell_1$ and for every orthonormal basis $\{e_j\}_{j=1}^{\infty}$ for $\mathbb{H}$, if $\sum_{j=1}^{\infty} \lambda_j |\langle x_k, e_j\rangle|^2 = 0$ for all k then $\lambda = 0$.*

Proof. (1) $\Rightarrow$ (2): We prove the result by way of contradiction. Suppose that (2) is false. Then there is an $\lambda = (\lambda_1, \lambda_2, \ldots) \in \ell_1$ and an orthonormal basis $\{e_j\}_{j=1}^{\infty}$ so that $\sum_{j=1}^{\infty} \lambda_j |\langle x_k, e_j \rangle|^2 = 0$ for all k but $\lambda \neq 0$.

Define an operator on $\mathbb{H}$ by

$$Te_j = \lambda_j e_j, \text{ for all } j = 1, 2, \ldots.$$

Then T is a non-zero self-adjoint operator. Moreover,

$$|T|e_j = \sqrt{TT^*}e_j = |\lambda_j|e_j, \text{ for all } j.$$

Therefore,

$$\sum_{j=1}^{\infty} \langle |T|e_j, e_j \rangle = \sum_{j=1}^{\infty} |\lambda_j| < \infty.$$

Thus, T is a non-zero trace class self-adjoint operator. Moreover, we have that $\langle Tx_k, x_k \rangle = \sum_{j=1}^{\infty} \lambda_j |\langle x_k, e_j \rangle|^2 = 0$, for all $k = 1, 2, \ldots$. This is a contradiction.

(2) $\Rightarrow$ (1): Let T be a trace class self-adjoint operator such that

$$\langle Tx_k, x_k \rangle = 0, \text{ for all } k.$$

Since T is trace class and hence compact, there is an eigenbasis $\{e_j\}_{j=1}^{\infty}$ for T with respective eigenvalues $\{\lambda_j\}_{j=1}^{\infty}$. Then for every $k = 1, 2, \ldots$, we have

$$\sum_{j=1}^{\infty} \lambda_j |\langle x_k, e_j \rangle|^2 = \langle Tx_k, x_k \rangle = 0, \text{ for all } k.$$

Since T is trace class then

$$\sum_{j=1}^{\infty} |\lambda_j| = \sum_{j=1}^{\infty} |\langle Te_j, e_j \rangle| < \infty.$$

That is, $\lambda = (\lambda_1, \lambda_2, \ldots) \in \ell_1$. By assumption (2) we get $\lambda = 0$ and hence $T = 0$. ∎

Finally, by normalizing the trace, we can give a classification for the injectivity problem if we require further that our operators are trace one. First, we need to justisfy Theorem 2.36 so that we can use it for this case.

Theorem 2.40. *Given a family of vectors $\mathcal{X} = \{x_k\}_{k=1}^{\infty}$ in the real or complex Hilbert space $\mathbb{H}$, the following are equivalent:*

1. *Whenever T, S are trace class positive and self-adjoint of trace one, and*

   $$\langle Tx_k, x_k\rangle = \langle Sx_k, x_k\rangle, \text{ for all } k,$$

   then $T = S$.

2. *Whenever T, S are trace class self-adjoint of trace one, and*

   $$\langle Tx_k, x_k\rangle = \langle Sx_k, x_k\rangle, \text{ for all } k,$$

   then $T = S$.

3. *Whenever T is trace class self-adjoint of trace zero, and*

   $$\langle Tx_k, x_k\rangle = 0, \text{ for all } k,$$

   then $T = 0$.

Proof. (1) $\Rightarrow$ (2): Let T, S be trace class self-adjoint operators of trace one such that

$$\langle Tx_k, x_k\rangle = \langle Sx_k, x_k\rangle, \text{ for all } k.$$

Set $R = T - S$ then R is a trace class self-adjoint operator of trace zero. Let $\{e_j\}_{j=1}^{\infty}$ be an orthonormal basis for $\mathbb{H}$ and let $\{u_j\}_{j=1}^{\infty}$ be an eigenbasis for R with respective eigenvalues $\{\lambda_j\}_{j=1}^{\infty}$. Then $\sum_{j=1}^{\infty} \lambda_j = 0$. Define operators U and D on $\mathbb{H}$ by $Ue_j = u_j$

and $De_j = \lambda_j e_j$, for $j = 1, 2, \ldots$. Then U is an unitary operator and D is a trace class self-adjoint operator of trace zero, and

$$R = UDU^*.$$

Now define non-negative numbers

$$r_1 = \frac{1 + |\lambda_1|}{A}, s_1 = \frac{1 + |\lambda_1| - \lambda_1}{A}, r_j = \frac{|\lambda_j|}{A}, s_j = \frac{|\lambda_j| - \lambda_j}{A}, j = 2, 3, \ldots,$$

where

$$A = 1 + \sum_{j=1}^{\infty} |\lambda_j| = 1 + \sum_{j=1}^{\infty} |\lambda_j| - \sum_{j=1}^{\infty} \lambda_j > 0$$

then $\lambda_j = r_j - s_j$ for all j. Let D_1, D_2 be operators defined by

$$D_1 e_j = r_j e_j, \ D_2 e_j = s_j e_j \text{ for } j = 1, 2, \ldots.$$

Then D_1, D_2 are trace class positive self-adjoint of trace one and we have

$$R = UDU^* = U(D_1 - D_2)U^* = UD_1U^* - UD_2U^*.$$

Moreover, UD_1U^*, UD_2U^* are trace class positive self-adjoint operators of trace one. Since

$$0 = \langle Rx_k, x_k \rangle = \langle UD_1U^*x_k, x_k \rangle - \langle UD_2U^*x_k, x_k \rangle,$$

then $UD_1U^* = UD_2U^*$. Thus, $R = 0$ and hence $T = S$.

(2) $\Rightarrow$ (3): Let T be any trace class operator of trace zero such that

$$\langle Tx_k, k_k \rangle = 0 \text{ for all } k.$$

Define an operator S on $\mathbb{H}$ by

$$Se_1 = e_1, Se_j = 0, \text{ for } j = 2, 3, \ldots.$$

Then S and $T+S$ are trace class self-adjoint operators of trace one.

Since $\langle (T+S)x_k, x_k\rangle = \langle Sx_k, x_k\rangle$ for all k, $T+S=S$ and hence $T=0$.

(3) $\Rightarrow$ (1): Let T, S are trace class positive self-adjoint operators of trace one such that

$$\langle Tx_k, x_k\rangle = \langle Sx_k, x_k\rangle \text{ for all } k.$$

Then $\langle (T-S)x_k, x_k\rangle = 0$ for all k. Since $T-S$ is a trace class seft-adjoint operator of trace zero, $T=S$ by (3). ∎

Now we are ready to give a classification for the Injectivity problem for operators of trace one. First, we need a definition.

Definition 2.41. We define a subspace of the real space ℓ_1 as follows:

$$W := \{(\lambda_1, \lambda_2, \ldots) \in \ell_1 : \sum_{j=1}^{\infty} \lambda_j = 0\}.$$

Theorem 2.42. *Let $\mathcal{X} = \{x_k\}_{k=1}^{\infty}$ be a frame for an infinite dimensional real or complex Hilbert space $\mathbb{H}$. The following are equivalent:*

1. *If T is a trace class self-adjoint operator of trace zero such that*

$$\langle Tx_k, x_k\rangle = 0, \text{ for all } k,$$

then $T=0$.

2. *For every $\lambda = (\lambda_1, \lambda_2, \ldots) \in W$ and for every orthonormal basis $\{e_j\}_{j=1}^{\infty}$ for $\mathbb{H}$, if $\sum_{j=1}^{\infty} \lambda_j |\langle x_k, e_j\rangle|^2 = 0$ for all k then $\lambda = 0$.*

Proof. (1) $\Rightarrow$ (2): We prove the contrapositive. Suppose that (2) is false. Then there is an $\lambda = (\lambda_1, \lambda_2, \ldots) \in W$ and an orthonormal basis $\{e_j\}_{j=1}^{\infty}$ so that $\sum_{j=1}^{\infty} \lambda_j |\langle x_k, e_j\rangle|^2 = 0$ for all k but $\lambda \neq 0$.

Define an operator on $\mathbb{H}$ by

$$Te_j = \lambda_j e_j, \text{ for all } j = 1, 2, \ldots.$$

Then T is a non-zero trace class self-adjoint operator of trace zero. Moreover, we have that $\langle Tx_k, x_k\rangle = \sum_{j=1}^{\infty} \lambda_j |\langle x_k, e_j\rangle|^2 = 0$, for all $k = 1, 2, \ldots$. This is a contradiction.

(2) $\Rightarrow$ (1): Let T be a trace class self-adjoint operator of trace zero such that

$$\langle Tx_k, x_k\rangle = 0, \text{ for all } k.$$

Let $\{e_j\}_{j=1}^{\infty}$ be an eigenbasis for T with respective eigenvalues $\{\lambda_j\}_{j=1}^{\infty}$. Then for every $k = 1, 2, \ldots$, we have

$$\sum_{j=1}^{\infty} \lambda_j |\langle x_k, e_j\rangle|^2 = \langle Tx_k, x_k\rangle = 0, \text{ for all } k.$$

Since T is trace class,

$$\sum_{j=1}^{\infty} |\lambda_j| = \sum_{j=1}^{\infty} |\langle Te_j, e_j\rangle| < \infty.$$

Moreover, $\sum_{j=1}^{\infty} \lambda_j = 0$. Thus, $\lambda = (\lambda_1, \lambda_2, \ldots) \in W$. By assumption (2) we get $\lambda = 0$ and hence $T = 0$. ∎

From now on, we will also work in the direct sum of infinitely many copies of ℓ_2.

Definition 2.43. Denote by $\tilde{\mathbb{H}}$ the direct sum of the real Hilbert spaces ℓ_2:

$$\tilde{\mathbb{H}} = \left(\sum_{i=1}^{\infty} \oplus \ell_2\right)_{\ell_2} = \left\{\{z_i\}_{i=1}^{\infty} \ : \ z_i \in \ell_2, \ \sum_i \|z_i\|^2 < \infty\right\}.$$

To avoid confusion with earlier notation, a vector in this sum will be written in the form:

$$\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n, \ldots),$$

and we have

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{\infty} \langle \mathbf{x}_i, \mathbf{y}_i \rangle.$$

We also need the following lemma for both the real and complex cases.

Lemma 2.44. *Let $A = (a_{ij})_{i,j=1}^{\infty}$ be a real or complex infinite matrix such that $\sum_{i,j=1}^{\infty} |a_{ij}|^2 < \infty$. Then the operator T_A defined in ℓ_2 by*

$$T_A(x_1, x_2, \ldots) = (y_1, y_2, \ldots),$$

where

$$y_i = \sum_{j=1}^{\infty} a_{ij} x_j, i = 1, 2, \ldots,$$

is a bounded operator. Moreover, T_A is self-adjoint if and only if $a_{ji} = \bar{a}_{ij}$ for all i, j.

Proof. Let $x = \{x_i\}_{i=1}^{\infty} \in \ell_2$. For each $i = 1, 2 \ldots$, we have

$$|y_i|^2 \leq \left(\sum_{j=1}^{\infty} |a_{ij} x_j| \right)^2 \leq \left(\sum_{j=1}^{\infty} |a_{ij}|^2 \right) \left(\sum_{j=1}^{\infty} |x_j|^2 \right) = \left(\sum_{j=1}^{\infty} |a_{ij}|^2 \right) \|x\|^2.$$

Hence,

$$\|T_A x\|^2 = \sum_{i=1}^{\infty} |y_i|^2 \leq \left(\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |a_{ij}|^2 \right) \|x\|^2.$$

This shows that T_A is a bounded operator on ℓ_2.

Suppose that T is self-adjoint. Then

$$a_{ji} = \langle T_A e_i, e_j \rangle = \langle e_i, T_A e_j \rangle = \overline{\langle T_A e_j, e_i \rangle} = \bar{a}_{ij},$$

for all i, j.

Conversely, if $a_{ji} = \bar{a}_{ij}$ for all i, j, then

$$\langle T_A^* e_i, e_j \rangle = \langle e_i, T_A e_j \rangle = \overline{\langle T_A e_j, e_i \rangle} = \bar{a}_{ij} = a_{ji} = \langle T_A e_i, e_j \rangle,$$

for all i, j. Hence $T_A^* = T_A$. ∎

The real case

Now we will solve the infinite dimensional injectivity problem in the real case. To avoid confusion between coordinates of a vector in ℓ_2 and vectors in $\tilde{\mathbb{H}}$ we define:

Definition 2.45. For $x = \{x_i\}_{i=1}^{\infty} \in \ell_2$, we define

$$\tilde{x} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n, \ldots) \in \tilde{\mathbb{H}},$$

where

$$\mathbf{x}_1 = (x_1x_1, x_1x_2, \ldots);\ \mathbf{x}_2 = (x_2x_2, x_2x_3, \ldots); \ldots; \mathbf{x}_n = (x_nx_n, x_nx_{n+1}, \ldots); \ldots$$

We first observe that these vectors are actually in $\tilde{\mathbb{H}}$.

Lemma 2.46. *If $x = \{x_i\}_{i=1}^{\infty} \in \ell_2$, then $\tilde{x} \in \tilde{\mathbb{H}}$.*

Proof. We have that

$$\sum_{j=i}^{\infty} |x_ix_j|^2 = |x_i|^2 \sum_{j=i}^{\infty} |x_j|^2 \leq |x_i|^2\|x\|^2,$$

for $i = 1, 2, \ldots$. Hence $\mathbf{x}_i \in \ell_2$ for all i.

Moreover, since

$$\sum_{i=1}^{\infty} \|\mathbf{x}_i\|^2 \leq \sum_{i=1}^{\infty} |x_i|^2\|x\|^2 = \|x\|^4,$$

then $\tilde{x} \in \tilde{\mathbb{H}}$. ∎

Now we are ready for the classification of the solutions to the injectivity problem in the infinite dimensional case.

Theorem 2.47. *Let $\mathcal{X} = \{x_k\}_{k=1}^{\infty}$ be a frame in the real Hilbert space ℓ_2. The following are equivalent:*

1. $\mathcal{X}$ is injective.

2. $\overline{\text{span}}\{\tilde{x}_k\}_{k=1}^{\infty} = \tilde{\mathbb{H}}$.

Proof. (1) $\Rightarrow$ (2): Let any $a = (\mathbf{a}_1, \mathbf{a}_2, \ldots) \in \tilde{\mathbb{H}}$ be such that $a \perp \overline{\text{span}}\{\tilde{x}_k\}_{k=1}^{\infty}$. Then $\langle a, \tilde{x}_k \rangle = 0$ for all k.

We denote

$$\mathbf{a}_1 = (a_{11}, a_{12}, \ldots);\ \mathbf{a}_2 = (a_{22}, a_{23}, \ldots); \ldots, \mathbf{a}_n = (a_{nn}, a_{n(n+1)}, \ldots); \ldots.$$

Define an infinite matrix $B = (b_{ij})_{i,j=1}^{\infty}$, where $b_{ii} = a_{ii}$ for all i and $b_{ij} = b_{ji} = \frac{1}{2}a_{ij}$ for all $i < j$.

Then by Lemma 2.44, the operator T_B defined by B is a Hilbert Schmidt self-adjoint operator.

For any $x = \{x_i\}_{i=1}^{\infty} \in \ell_2$, we have

$$\begin{aligned}
\langle T_B x, x \rangle &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} b_{ij} x_i x_j \\
&= \sum_{i=1}^{\infty} b_{ii} x_i^2 + 2 \sum_{i<j} b_{ij} x_i x_j \\
&= \sum_{i=1}^{\infty} a_{ii} x_i^2 + \sum_{i<j} a_{ij} x_i x_j \\
&= \sum_{i=1}^{\infty} \langle \mathbf{a}_i, \mathbf{x}_i \rangle \\
&= \langle a, \tilde{x} \rangle.
\end{aligned}$$

Hence, $\langle T_B x_k, x_k \rangle = \langle a, \tilde{x}_k \rangle = 0$ for all k. This implies $T_B = 0$ by (1) and therefore $a = 0$.

(2) $\Rightarrow$ (1): Let T be a Hilbert Schmidt self-adjoint operator on ℓ_2 such that $\langle T x_k, x_k \rangle = 0$ for all k, and recall that $\{e_i\}_{i=1}^{\infty}$ is the canonical orthonormal basis for ℓ_2.

Denote

$$a_{ij} = \langle Te_j, e_i \rangle, i, j = 1, 2 \ldots,$$

and

$$\tilde{T} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n, \ldots),$$

where

$$\mathbf{a}_1 = (a_{11}, 2a_{12}, 2a_{13}, \ldots); \quad \mathbf{a}_2 = (a_{22}, 2a_{23}, 2a_{24}, \ldots); \ldots;$$

$$\mathbf{a}_n = (a_{nn}, 2a_{n(n+1)}, 2a_{n(n+2)}, \ldots); \ldots.$$

Since T is a Hilbert Schmidt operator, $\tilde{T} \in \tilde{\mathbb{H}}$. Moreover, we have

$$\langle \tilde{T}, \tilde{x}_k \rangle = \langle Tx_k, x_k \rangle = 0, \text{ for all } k.$$

Since $\overline{\text{span}}\{\tilde{x}_k\}_{k=1}^{\infty} = \tilde{\mathbb{H}}$, we get $\tilde{T} = 0$. So $T = 0$.

∎

Remark 2.48. We have that (2) ⇒ (1) in the theorem holds for trace class operators. But in general (1) ⇒ (2) since the operators we construct may not be trace class.

The complex case

For the complex case of the injectivity problem, we need a new variation of the tilde vectors.

Definition 2.49. For $x = \{x_i\}_{i=1}^{\infty} \in \ell_2$, we define

$$\tilde{x} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n, \ldots),$$

where

$$\mathbf{x}_1 = (|x_1|^2, \text{Re}(\bar{x}_1 x_2), \text{Im}(\bar{x}_1 x_2), \text{Re}(\bar{x}_1 x_3), \text{Im}(\bar{x}_1 x_3), \ldots);$$

$$\mathbf{x}_2 = (|x_2|^2, \mathrm{Re}(\bar{x}_2 x_3), \mathrm{Im}(\bar{x}_2 x_3), \mathrm{Re}(\bar{x}_2 x_4), \mathrm{Im}(\bar{x}_2 x_4), \ldots); \ldots;$$

$$\mathbf{x}_n = (|x_n|^2, \mathrm{Re}(\bar{x}_n x_{n+1}), \mathrm{Im}(\bar{x}_n x_{n+1}), \mathrm{Re}(\bar{x}_n x_{n+2}), \mathrm{Im}(\bar{x}_n x_{n+2}), \ldots); \ldots.$$

We first need to verify that our vectors are in $\tilde{\mathbb{H}}$.

Lemma 2.50. *If $x = \{x_i\}_{i=1}^{\infty} \in \ell_2$, then $\tilde{x} \in \tilde{\mathbb{H}}$.*

Proof. For each $i = 1, 2, \ldots$, we have

$$\begin{aligned}
\|\mathbf{x}_i\|^2 &= |x_i|^4 + \sum_{j=i+1}^{\infty} |\,\mathrm{Re}(\bar{x}_i x_j)|^2 + \sum_{j=i+1}^{\infty} |\,\mathrm{Im}(\bar{x}_i x_j)|^2 \\
&= |x_i|^4 + \sum_{j=i+1}^{\infty} |\bar{x}_i x_j|^2 \\
&= |x_i|^2 \left(|x_i|^2 + \sum_{j=i+1}^{\infty} |x_j|^2 \right) \\
&\leq |x_i|^2 \|x\|^2.
\end{aligned}$$

It follows that:

$$\sum_{i=1}^{\infty} \|\mathbf{x}_i\|^2 \leq \sum_{i=1}^{\infty} |x_i|^2 \|x\|^2 = \|x\|^4.$$

This implies $\tilde{x} \in \tilde{\mathbb{H}}$. ∎

Now we give the classification theorem for injectivity in the infinite dimensional case.

Theorem 2.51. *Let $\mathcal{X} = \{x_k\}_{k=1}^{\infty}$ be a frame in the complex Hilbert space ℓ_2. The following are equivalent:*

1. *$\mathcal{X}$ gives injectivity.*

2. *$\overline{\mathrm{span}}\{\tilde{x}_k\}_{k=1}^{\infty} = \tilde{\mathbb{H}}$.*

Proof. (1) $\Rightarrow$ (2): Let any $a = (\mathbf{a}_1, \mathbf{a}_2, \ldots) \in \tilde{\mathbb{H}}$ be such that $a \perp \overline{\operatorname{span}}\{\tilde{x}_k\}_{k=1}^{\infty}$. Then $\langle a, \tilde{x}_k \rangle = 0$ for all k.

Denote

$$\mathbf{a}_1 = (a_{11}, u_{12}, v_{12}, u_{13}, v_{13}, \ldots);\ \mathbf{a}_2 = (a_{22}, u_{23}, v_{23}, u_{24}, v_{24}, \ldots); \ldots;$$

$$\mathbf{a}_n = (a_{nn}, u_{n(n+1)}, v_{n(n+1)}, u_{n(n+2)}, v_{n(n+2)}, \ldots); \ldots.$$

Define an infinite matrix $B = (b_{ij})_{i,j=1}^{\infty}$, where $b_{ii} = a_{ii}$ for all i and $b_{ij} = \bar{b}_{ji} = \frac{1}{2}(u_{ij} - \iota v_{ij})$ for all $i < j$.

We have

$$\sum_{i,j=1}^{\infty} |b_{ij}|^2 = \sum_{i=1}^{\infty} |a_{ii}|^2 + 2\sum_{i<j} |b_{ij}|^2 = \sum_{i=1}^{\infty} |a_{ii}|^2 + \frac{1}{2}\sum_{i<j} \left(|u_{ij}|^2 + |v_{ij}|^2\right) < \infty.$$

Then by Lemma 2.44, the operator T_B defined by B is Hilbert Schmidt and self-adjoint.

For any $x = \{x_i\}_{i=1}^{\infty} \in \ell_2$, we have

$$\begin{aligned}
\langle T_B x, x \rangle &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} b_{ij}\bar{x}_i x_j \\
&= \sum_{i=1}^{\infty} b_{ii}|x_i|^2 + 2\sum_{i<j} \operatorname{Re}(b_{ij}\bar{x}_i x_j) \\
&= \sum_{i=1}^{\infty} b_{ii}|x_i|^2 + 2\sum_{i<j} (\operatorname{Re}(b_{ij})\operatorname{Re}(\bar{x}_i x_j) - \operatorname{Im}(b_{ij})\operatorname{Im}(\bar{x}_i x_j)) \\
&= \sum_{i=1}^{\infty} a_{ii}|x_i|^2 + \sum_{i<j} (u_{ij}\operatorname{Re}(\bar{x}_i x_j) + v_{ij}\operatorname{Im}(\bar{x}_i x_j)) \\
&= \sum_{i=1}^{\infty} \langle \mathbf{a}_i, \mathbf{x}_i \rangle \\
&= \langle a, \tilde{x} \rangle.
\end{aligned}$$

Hence, $\langle T_B x_k, x_k \rangle = \langle a, \tilde{x}_k \rangle = 0$ for all k. This implies $T_B = 0$ by (1) and therefore $a = 0$.

(2) $\Rightarrow$ (1): Let T be a Hilbert Schmidt self-adjoint operator such that $\langle Tx_k, x_k\rangle = 0$ for all k.

Denote $a_{ij} = \langle Te_j, e_i\rangle$ for $i, j = 1, 2, \ldots$ and $\tilde{T} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n, \ldots)$, where

$$\mathbf{a}_1 = (a_{11}, 2\operatorname{Re}(a_{12}), -2\operatorname{Im}(a_{12}), 2\operatorname{Re}(a_{13}), -2\operatorname{Im}(a_{13}), \ldots);$$

$$\mathbf{a}_2 = (a_{22}, 2\operatorname{Re}(a_{23}), -2\operatorname{Im}(a_{23}), 2\operatorname{Re}(a_{24}), -2\operatorname{Im}(a_{24}), \ldots); \ldots;$$

$$\mathbf{a}_n = (a_{nn}, 2\operatorname{Re}(a_{n(n+1)}), -2\operatorname{Im}(a_{n(n+1)}), 2\operatorname{Re}(a_{n(n+2)}), -2\operatorname{Im}(a_{n(n+2)}), \ldots); \ldots$$

Since T is Hilbert Schmidt, $\tilde{T} \in \tilde{\mathbb{H}}$.

For any $x = \sum_{j=1}^{\infty} x_j e_j$ we have

$$\begin{aligned}
\langle Tx, x\rangle &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \bar{x}_i x_j \langle Te_j, e_i\rangle \\
&= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \bar{x}_i x_j a_{ij} \\
&= \sum_{i=1}^{\infty} a_{ii}|x_i|^2 + 2\sum_{i<j} \operatorname{Re}(a_{ij}\bar{x}_i x_j) \\
&= \sum_{i=1}^{\infty} a_{ii}|x_i|^2 + 2\sum_{i<j} (\operatorname{Re}(a_{ij})\operatorname{Re}(\bar{x}_i x_j) - \operatorname{Im}(a_{ij})\operatorname{Im}(\bar{x}_i x_j)) \\
&= \langle \tilde{T}, \tilde{x}\rangle.
\end{aligned}$$

Hence

$$\langle \tilde{T}, \tilde{x}_k\rangle = \langle Tx_k, x_k\rangle = 0, \text{ for all } k.$$

Since $\overline{\operatorname{span}}\{\tilde{x}_k\}_{k=1}^{\infty} = \tilde{\mathbb{H}}$, $\tilde{T} = 0$. So $T = 0$. This completes the proof. ∎

As a consequence we have:

Corollary 2.52. *For a frame $\{x_k\}_{k=1}^{\infty}$ in ℓ_2 the following are equivalent:*

1. *The family $\{x_k x_k^*\}_{k=1}^{\infty}$ spans the family of real self-adjoint Hilbert Schmidt operators on ℓ_2.*

2. The family $\{\tilde{x}_k\}_{k=1}^{\infty}$ spans $\tilde{\mathbb{H}}$.

As we have seen in the proof of Theorem 2.47 for the real case and Theorem 2.51 for the complex case, for a vector $a \in \tilde{\mathbb{H}}$, there is a Hilbert Schmidt self-adjoint operator T so that

$$\langle Tx, x\rangle = \langle a, \tilde{x}\rangle, \text{ for all } x \in \ell_2.$$

Conversely, for a Hilbert Schmidt self-adjoint operator T, there is a vector $\tilde{T} \in \tilde{\mathbb{H}}$ satisfying

$$\langle \tilde{T}, \tilde{x}\rangle = \langle Tx, x\rangle, \text{ for all } x \in \ell_2.$$

Is is easy to see that the canonical orthonormal basis is not injective. Actually, the family $\{\tilde{e}_i\}_{i=1}^{\infty}$ forms an orthonormal set in $\tilde{\mathbb{H}}$. In finite dimensions, the definition of a frame is synonymous with a spanning set. However, in infinite dimensions this does not hold. The following theorem shows that for a given Bessel sequence, the associated tilde vectors will not produce a frame in $\tilde{\mathbb{H}}$. In particular, injective frames do not produce frames in $\tilde{\mathbb{H}}$.

Theorem 2.53. *For any Bessel sequence $\{x_k\}_{k=1}^{\infty}$ for the real or complex space ℓ_2, the family $\{\tilde{x}_k\}_{k=1}^{\infty}$ is a Bessel sequence in $\tilde{\mathbb{H}}$. However, $\{\tilde{x}_k\}_{k=1}^{\infty}$ is not a frame for $\tilde{\mathbb{H}}$.*

Proof. We may assume that $\|x_k\| \leq 1$ for all k. Let B be the Bessel bound of $\{x_k\}_{k=1}^{\infty}$.

Given any finite real scalar sequence $\{a_k\}_{k=1}^{\infty}$ we will compute the real case and the complex case separately.

The real case: Using Definition 2.45 for the tilde vector $\tilde{x}_k$, we have

$$\|\sum_{k=1}^{\infty} a_k\tilde{x}_k\|^2 = \sum_{i=1}^{\infty}\sum_{j=i}^{\infty}\left(\sum_{k=1}^{\infty} a_k x_{ki}x_{kj}\right)^2$$

$$\leq \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\left(\sum_{k=1}^{\infty} a_k x_{ki}x_{kj}\right)^2$$
$$= \sum_{i=1}^{\infty} \| \sum_{k=1}^{\infty} a_k x_{ki} x_k\|^2.$$

Using the fact that $\{x_k\}_{k=1}^{\infty}$ is Bessel with bound B, we get

$$\begin{aligned}
\| \sum_{k=1}^{\infty} a_k \tilde{x}_k\|^2 &\leq B \sum_{i=1}^{\infty}\sum_{k=1}^{\infty}(a_k x_{ki})^2 \\
&= B \sum_{k=1}^{\infty} a_k^2 \sum_{i=1}^{\infty} x_{ki}^2 \\
&= B \sum_{k=1}^{\infty} a_k^2 \|x_k\|^2 \\
&\leq B \sum_{k=1}^{\infty} a_k^2.
\end{aligned}$$

The complex case: Now we need to use Definition 2.49 for the tilde vectors $\tilde{x}_k$.

We have that

$$\begin{aligned}
\| \sum_{k=1}^{\infty} a_k \tilde{x}_k\|^2 &= \sum_{i=1}^{\infty}\left(\sum_{k=1}^{\infty} a_k |x_{ki}|^2\right)^2 + \sum_{i=1}^{\infty}\sum_{j=i+1}^{\infty}\left(\sum_{k=1}^{\infty} a_k \operatorname{Re}(\bar{x}_{ki}x_{kj})\right)^2 \\
&\quad + \sum_{i=1}^{\infty}\sum_{j=i+1}^{\infty}\left(\sum_{k=1}^{\infty} a_k \operatorname{Im}(\bar{x}_{ki}x_{kj})\right)^2 \\
&= \sum_{i=1}^{\infty}\left(\sum_{k=1}^{\infty} a_k |x_{ki}|^2\right)^2 + \sum_{i=1}^{\infty}\sum_{j=i+1}^{\infty}\left(\operatorname{Re}\left(\sum_{k=1}^{\infty} a_k \bar{x}_{ki}x_{kj}\right)\right)^2 \\
&\quad + \sum_{i=1}^{\infty}\sum_{j=i+1}^{\infty}\left(\operatorname{Im}\left(\sum_{k=1}^{\infty} a_k \bar{x}_{ki}x_{kj}\right)\right)^2 \\
&\leq 2\sum_{i=1}^{\infty}\left(\sum_{k=1}^{\infty} a_k |x_{ki}|^2\right)^2 + 2\sum_{i=1}^{\infty}\sum_{j=i+1}^{\infty}\left|\sum_{k=1}^{\infty} a_k \bar{x}_{ki}x_{kj}\right|^2 \\
&\leq 2\sum_{i=1}^{\infty}\left(\sum_{k=1}^{\infty} a_k |x_{ki}|^2\right)^2 + 2\sum_{i=1}^{\infty}\sum_{j=1,j\neq i}^{\infty}\left|\sum_{k=1}^{\infty} a_k \bar{x}_{ki}x_{kj}\right|^2 \\
&= 2\sum_{i=1}^{\infty} \| \sum_{k=1}^{\infty} a_k \bar{x}_{ki} x_k\|^2 \leq 2B \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} a_k^2 |x_{ki}|^2 \\
&= 2B \sum_{k=1}^{\infty} a_k^2 \sum_{i=1}^{\infty} |x_{ki}|^2 = 2B \sum_{k=1}^{\infty} a_k^2 \|x_k\|^2 \leq 2B \sum_{k=1}^{\infty} a_k^2.
\end{aligned}$$

Hence, $\{\tilde{x}_k\}_{k=1}^{\infty}$ is a Bessel sequence for the both cases.

Now we will show that $\{\tilde{x}_k\}_{k=1}^{\infty}$ fails to have a lower frame bound. We will prove the real case, the complex case is similar. By our assumption, we have that

$$\sum_{i=1}^{\infty} |x_{ki}|^2 < \infty, \text{ for all } k.$$

Also,

$$\sum_{k=1}^{\infty} |x_{ki}|^2 = \sum_{k=1}^{\infty} |\langle e_i, x_k \rangle|^2 < \infty, \text{ for all } i.$$

Fix $\epsilon > 0$ and choose n so that $\sum_{k=n}^{\infty} |x_{k1}|^2 < \epsilon$. Now choose m so that $\sum_{k=1}^{n-1} |x_{km}|^2 < \epsilon$. Let

$$\tilde{e}_{1m} := (e_m; \mathbf{0}; \mathbf{0}, \ldots) \in \tilde{\mathbb{H}}.$$

Then we have

$$\begin{aligned}
\sum_{k=1}^{\infty} |\langle \tilde{e}_{1m}, \tilde{x}_k \rangle|^2 &= \sum_{k=1}^{\infty} |x_{k1}|^2 |x_{km}|^2 \\
&= \sum_{k=1}^{n-1} |x_{k1}|^2 |x_{km}|^2 + \sum_{k=n}^{\infty} |x_{k1}|^2 |x_{km}|^2 \\
&\leq \sum_{k=1}^{n-1} |x_{km}|^2 + \sum_{k=n}^{\infty} |x_{k1}|^2 \\
&< 2\epsilon.
\end{aligned}$$

It follows that $\{\tilde{x}_k\}_{k=1}^{\infty}$ does not have a lower frame bound. ∎

2.2.2 Constructing the Solutions to the Injectivity Problem

For the construction of solutions to the injectivity problem, we will follow the outline for the finite dimensional case. But this construction is much more complicated because of problems with convergence, problems with keeping the upper frame bound

finite, and the fact that we cannot show spanning in ℓ_2 by just checking linear independence. Also, we proved in the finite dimensional case that the $\tilde{x}_i$ span by showing they are independent and have enough vectors to span $\tilde{\mathbb{H}}$. This does not work in the infinite dimensional case. Note that the following construction works for trace class operators and for Hilbert Schmidt operators.

Theorem 2.54. *Let $\{e_i\}_{i=1}^{\infty}$ be the canonial basis for the real Hilbert space ℓ_2 and let $a_i \neq 0$ for $i = 1, 2, \ldots$ be such that $\sum_{i=1}^{\infty} a_i^2 < \infty$. Define*

$$x_k = a_k(e_1 + e_{k+1}), \text{ for } k = 1, 2, \ldots.$$

Let L be the right shift operator on ℓ_2. Then the family

$$\{e_i\}_{i=1}^{\infty} \cup \{\frac{1}{2^i} L^i x_k\}_{i=0,k=1}^{\infty,\infty}$$

is a frame for ℓ_2 which gives injectivity.

Proof. First we need to see that our family of vectors forms a frame for ℓ_2. Since our family contains an orthonormal basis for ℓ_2, we automatically have a lower frame bound. So we need to check that our family is Bessel, and since $\{e_i\}_{i=1}^{\infty}$ is already Bessel, we only need to check that $\{\frac{1}{2^i} L^i x_k\}_{i=0,k=1}^{\infty,\infty}$ is Bessel.

For any $x \in \ell_2$, we have

$$\begin{aligned}
\sum_{i=0}^{\infty}\sum_{k=1}^{\infty} |\langle x, \frac{1}{2^i} L^i x_k\rangle|^2 &\leq \sum_{i=0}^{\infty}\sum_{k=1}^{\infty} \frac{1}{4^i}\|x\|^2\|L^i x_k\|^2 \\
&\leq \sum_{i=0}^{\infty}\sum_{k=1}^{\infty} \frac{1}{4^i}\|x\|^2 4a_k^2 \\
&= \left(\sum_{i=0}^{\infty} \frac{1}{4^{i-1}} \sum_{k=1}^{\infty} a_k^2\right)\|x\|^2.
\end{aligned}$$

So our family is a Bessel sequence.

To see our frame is injective, let T be a Hilbert Schmidt self-adjoint operator such that

$$\langle Te_k, e_k\rangle = 0 \text{ and } \langle T(L^i x_k), L^i x_k\rangle = 0, \text{ for } i = 0, 1\ldots;\ k = 1, 2, \ldots.$$

Note that

$$L^i x_k = a_k(e_{1+i} + e_{1+i+k}) \text{ for all } i, k.$$

Hence

$$\begin{aligned}\langle T(L^i x_k), L^i x_k\rangle &= a_k^2\langle Te_{1+i}, e_{1+i}\rangle + 2a_k^2\langle Te_{1+i}, e_{1+i+k}\rangle + a_k^2\langle Te_{1+i+k}, e_{1+i+k}\rangle \\ &= 2a_k^2\langle Te_{1+i}, e_{1+i+k}\rangle,\end{aligned}$$

for all i, k.

This implies $\langle Te_j, e_k\rangle = 0$ for all $j, k = 1, 2, \ldots$, and hence $T = 0$. ∎

The complex version of this construction looks like:

Theorem 2.55. *Let $\{e_i\}_{i=1}^\infty$ be the canonical orthonormal basis for complex ℓ_2, and let $\{a_i\}_{i=1}^\infty, \{b_i\}_{i=1}^\infty \in \ell_2$, $|a_i|, |b_i| \neq 0$ for all i. Then the following frame gives injectivity:*

$$\{e_i\}_{i=1}^\infty \cup \{\frac{1}{2^i}L^i(a_k(e_1 + e_{k+1}))\}_{i=0,k=1}^\infty \cup \{\frac{1}{2^i}L^i(b_k(e_1 + \iota e_{k+1}))\}_{i=0,k=1}^\infty.$$

The above frames are unbounded. The following theorem shows that we can easily adjust unbounded injective frames to produce bounded injective frames.

Theorem 2.56. *Every injective frame $\mathcal{X} = \{e_i\}_{i=1}^\infty \cup \{x_k\}_{k=1}^\infty$ of finitely supported vectors, induces a bounded injective frame.*

Proof. Recall that for each k, we denote

$$x_k = (x_{k1}, x_{k2}, \ldots, x_{ki}, \ldots).$$

Choose integers $n_1 < n_2 < \cdots$ so that $\max\{i : x_{ki} \neq 0\} < n_k$. For $k = 1, 2, \ldots$ let

$$y_{2k} = x_k + e_{n_k} \quad \text{and} \quad y_{2k+1} = x_k - e_{n_k}, \text{ for } k = 1, 2, \ldots.$$

It is clear that $\mathcal{Y} = \{e_i\}_{i=1}^{\infty} \cup \{y_k\}_{k=1}^{\infty}$ is still a frame and $\|y_k\| \geq 1$ for all $k = 1, 2, \ldots$. Since $\tilde{y}_{2k} + \tilde{y}_{2k+1} = 2\tilde{x}_k + 2\tilde{e}_{n_k}$, and $\{\tilde{e}_{n_k}\}_{k=1}^{\infty}$ are vectors in our set, it follows that $\{\tilde{x}_k\}_{k=1}^{\infty}$ is in our set of vectors and so $\mathcal{Y}$ is injective. ∎

2.2.3 The Solutions are Neither Open nor Dense

In this section we will show that the solutions to the injectivity problem in infinite dimensions are neither open nor dense in the class of frames. First we need a definition:

Definition 2.57. Given frames $\mathcal{X} = \{x_k\}_{k=1}^{\infty}$ and $\mathcal{Y} = \{y_k\}_{k=1}^{\infty}$ for ℓ_2, we define the **distance** between them by

$$d^2(\mathcal{X}, \mathcal{Y}) = \sum_{k=1}^{\infty} \|x_k - y_k\|^2.$$

Note that this distance may be infinity. The following theorem shows that the frames which give injectivity are not open in the family of frames for ℓ_2.

Theorem 2.58. *Let $\mathcal{X} = \{e_i\}_{i=1}^{\infty} \cup \{\frac{1}{2^i} L^i x_k\}_{i=0,k=1}^{\infty}$ be the injective frame for the real space ℓ_2 as in Theorem 2.54. Then for any $\epsilon > 0$, there is a frame $\mathcal{Y}$ such that $d(\mathcal{X}, \mathcal{Y}) < \epsilon$, and $\mathcal{Y}$ is not injective.*

Proof. Let any $\epsilon > 0$. Since the series $\sum_{i=0}^{\infty} \sum_{k=1}^{\infty} \|\frac{1}{2^i} L^i x_k\|^2$ converges, for any ϵ, there exists n_0 such that

$$\sum_{i=n_0+1}^{\infty} \sum_{k=1}^{\infty} \|\frac{1}{2^i} L^i x_k\|^2 < \epsilon^2.$$

Let $y_{ik} = \frac{1}{2^i} L^i x_k$ for $i = 0, 1, \ldots, n_0$ and $k = 1, 2, \ldots$, and $y_{ik} = 0$ otherwise. It is clear that

$$\mathcal{Y} = \{e_i\}_{i=1}^{\infty} \cup \{y_{ik}\}_{i=0,k=1}^{\infty}$$

cannot give injectivity by Theorem 2.47 while

$$d^2(\mathcal{X}, \mathcal{Y}) = \sum_{i=n_0+1}^{\infty} \sum_{k=1}^{\infty} \| \frac{1}{2^i} L^i x_k \|^2 < \epsilon^2.$$

This completes the proof. ∎

Remark 2.59. There is a perturbation theory for frames which looks like it should apply here. The problem is that although our vectors form a frame for ℓ_2, their tilde vectors do not form a frame to $\tilde{\mathbb{H}}$ and so the theory does not apply.

To show the solutions are not dense, we need the definition of a Riesz sequence in ℓ_2.

Definition 2.60. A family of vectors $\{x_i\}_{i=1}^{\infty}$ in the real or complex Hilbert space ℓ_2 is called a **Riesz sequence** if there are constants $0 < A \leq B < \infty$ so that for all $\{a_i\}_{i=1}^{\infty} \subset \ell_2$ we have:

$$A \sum_{i=1}^{\infty} |a_i|^2 \leq \| \sum_{i=1}^{\infty} a_i x_i \|^2 \leq B \sum_{i=1}^{\infty} |a_i|^2.$$

The constants A, B are called the **lower and upper Riesz bounds**. If the vectors span ℓ_2, this is called a **Riesz basis**.

It is known [16, 18] that a Riesz basis is a frame and the Riesz bounds are the frame bounds. Also, $\{x_i\}_{i=1}^{\infty}$ is a Riesz sequence if and only if the operator $T : \ell_2 \to \ell_2$ given by $Te_i = x_i$ is a bounded, linear, invertible operator (on its range).

To show the desired result we first need a few known results from frame theory. The first is a perturbation result.

Proposition 2.61. *Assume $\chi = \{x_i\}_{i=1}^{\infty}$ are vectors in the real or complex space ℓ_2 satisfying:*

$$\sum_{i=1}^{\infty} \|e_i - x_i\|^2 < \epsilon^2.$$

Then χ is a Riesz sequence in ℓ_2 with lower Riesz bound $(1-\epsilon)^2$.

Proof. We compute for scalars $\{a_i\}_{i=1}^{\infty}$,

$$\begin{aligned}
\|\sum_{i=1}^{\infty} a_i x_i\| &\geq \|\sum_{i=1}^{\infty} a_i e_i\| - \|\sum_{i=1}^{\infty} a_i (e_i - x_i)\| \\
&\geq \left(\sum_{i=1}^{\infty} |a_i|^2\right)^{1/2} - \sum_{i=1}^{\infty} |a_i| \|e_i - x_i\| \\
&\geq \left(\sum_{i=1}^{\infty} |a_i|^2\right)^{1/2} - \left(\sum_{i=1}^{\infty} |a_i|^2\right)^{1/2} \left(\sum_{i=1}^{\infty} \|e_i - x_i\|^2\right)^{1/2} \\
&\geq \left(\sum_{i=1}^{\infty} |a_i|^2\right)^{1/2} (1-\epsilon)
\end{aligned}$$

The upper Riesz bound is done similarly. ∎

Theorem 2.62 ([14])**.** *Let Y, Z be subspaces of a Banach space X. If $T : Y \to Z$ is a surjective linear operator with $\|I - T\| < 1$, then $codim_X Y = codim_X Z$.*

The next theorem shows that the solution set of the infinite dimensional injectivity problem is not dense in the class of all frames for ℓ_2.

Theorem 2.63. *Let $\{e_k\}_{k=1}^{\infty}$ be the canonical basis for the real space ℓ_2 and $\mathcal{X} = \{x_k\}_{k=1}^{\infty} \subset \ell_2$ be such that*

$$\sum_{k=1}^{\infty} \|e_k - x_k\|^2 \leq \frac{1}{8},$$

Then $\mathcal{X}$ is not injective.

Proof. Will will show that $\mathcal{X}$ does not satisfy Theorem 2.47. Note that $codim_{\tilde{\mathbb{H}}}\{\tilde{e}_k\}_{k=1}^{\infty}$ is infinite. Also, $\{\tilde{e}_k\}_{k=1}^{\infty}$ is an orthonormal sequence in $\tilde{\mathbb{H}}$. We have that

$$\sum_{k=1}^{\infty}\|e_k - x_k\|^2 = \sum_{k=1}^{\infty}\left((1-x_{kk})^2 + \sum_{i\neq k} x_{ki}^2\right) \leq \frac{1}{8}.$$

In particular, $\|x_k\|^2 \leq 2$. Let

$$X = \overline{\text{span}}\{\tilde{e}_k\}_{k=1}^{\infty}, \text{ and } Y = \overline{\text{span}}\{\tilde{x}_k\}_{k=1}^{\infty}.$$

For each $k = 1, 2, \ldots$ we have

$$\begin{aligned}\|\tilde{e}_k - \tilde{x}_k\|^2 &= (1-x_{kk}^2)^2 + \sum_{j\geq k+1}(x_{kk}x_{kj})^2 + \sum_{i\neq k}\sum_{j\geq i}(x_{ki}x_{kj})^2 \\ &\leq (1-x_{kk})^2(2\|x_k\|^2+2) + \|x_k\|^2\sum_{j\geq k+1}x_{kj}^2 + \|x_k\|^2\sum_{i\neq k}x_{ki}^2 \\ &\leq 6\left((1-x_{kk})^2 + \sum_{i\neq k}x_{ki}^2\right).\end{aligned}$$

Hence,

$$\sum_{k=1}^{\infty}\|\tilde{e}_k - \tilde{x}_k\|^2 \leq 6\sum_{k=1}^{\infty}\left((1-x_{kk})^2 + \sum_{i\neq k}x_{ki}^2\right) \leq \frac{3}{4}.$$

It follows that $\{\tilde{x}_k\}_{k=1}^{\infty}$ is a Riesz sequence.

Now we define $T : X \to Y$ by: for $x = \sum_{k=1}^{\infty}\langle x, \tilde{e}_k\rangle\tilde{e}_k \in X$,

$$Tx = \sum_{k=1}^{\infty}\langle x, \tilde{e}_k\rangle\tilde{x}_k.$$

Since T is mapping a Riesz sequence to a Riesz sequence, it follows that T is bounded and surjective. Now,

$$\begin{aligned}\|(I-T)x\| &= \left\|\sum_{k=1}^{\infty}\langle x, \tilde{e}_k\rangle(\tilde{e}_k - \tilde{x}_k)\right\| \\ &\leq \sum_{k=1}^{\infty}|\langle x, \tilde{e}_k\rangle|\|\tilde{e}_k - \tilde{x}_k\|\end{aligned}$$

$$\leq \left(\sum_{k=1}^{\infty} |\langle x, \tilde{e}_k\rangle|^2\right)^{1/2} \left(\sum_{k=1}^{\infty} \|\tilde{e}_k - \tilde{x}_k\|^2\right)^{1/2}$$
$$\leq \frac{\sqrt{3}}{2}\|x\|.$$

Hence, $\|I - T\| < 1$ and by Theorem 2.62, $codim_{\tilde{\mathbb{H}}} Y = codim_{\tilde{\mathbb{H}}} X = \infty$.

∎

2.2.4 The Solution to the State Estimation Problem

For the infinite dimensional case, the state estimation problem asks:

State Estimation Problem: Given an injective Parseval frame $\{x_k\}_{k=1}^{\infty}$ for ℓ_2, and a sequence of real numbers $a = \{a_k\}_{k=1}^{\infty}$, does there exist a Hilbert Schmidt self-adjoint operator T so that

$$\langle Tx_k, x_k\rangle = a_k, \text{ for all } k?$$

In fact, this problem is rarely solvable.

1. If $x_k x_k^* = x_l x_l^*$, but $a_k \neq a_l$ for some k, l, then the problem has no solution.

2. Recall a set of vectors $\{x_i\}_{i=1}^{\infty}$ is **ω-independent** if $\sum_{i=1}^{\infty} c_i x_i = 0$ implies $c_i = 0$ for all $i = 1, 2, \ldots$. If $\{x_k x_k^*\}_{k=1}^{\infty}$ is not ω-independent and $\sum_{k=1}^{\infty} c_k x_k x_k^* = 0$ but not all c_k are zero, then for $\langle Tx_k, x_k\rangle = a_k$ we need

$$\sum_{k=1}^{\infty} c_k a_k = \langle T, \sum_{k=1}^{\infty} c_k x_k x_k^*\rangle = 0.$$

For the solution of the state estimation problem we will need the notion of a separated sequence in ℓ_2.

Definition 2.64. A family of vectors $\{x_i\}_{i=1}^{\infty}$ in ℓ_2 is **separated** if for every $j \in \mathbb{N}$,

$$x_j \notin \overline{\text{span}}\{x_i\}_{i\neq j}.$$

It is δ-separated if the projection P_j onto $\overline{\text{span}}\{x_i\}_{i\neq j}$ satisfies

$$\|(I - P_j)x_j\| \geq \delta.$$

A Riesz basic sequence $\{x_k\}_{k\in I}$ is δ-separated since it is clear from the definition that

$$\text{dist}(x_k, \text{span}\{x_i\}_{i\neq k}) \geq \delta.$$

In general, a Bessel sequence which is δ-separated may not be a Riesz sequence. To see this let

$$\mathbb{H} = \left(\sum_{n=1}^{\infty} \oplus \mathbb{H}_n\right)_{\ell_2},$$

where $\mathbb{H}^n$ is an n-dimensional Hilbert space with orthonormal basis $\{e_{in}\}_{i=1}^{n}$. Let P be the orthogonal projection onto the one dimensional subspace spanned by $\sum_{i=1}^{n} e_{in}$. Then $\{(I - P)e_{in}\}_{i=1,n=1}^{n-1,\ \infty}$ as a family of vectors in $\mathbb{H}$ is δ-separated, 1-Bessel, but not a Riesz sequence. (Careful: We have thrown away the vectors $(I - P)e_{nn}$ above.)

Note also that a δ-separated sequence may not be Bessel.

Example 2.65. Let $x_i = e_1 + e_{i+1}, i = 1, 2, \ldots$. Then $\{x_i\}_{i=1}^{\infty}$ is not a Bessel sequence. However, it is δ-separated.

Indeed, let P_j be the projection onto $\overline{\text{span}}\{x_i\}_{i\neq j}$. Then

$$\|x_j - P_j x_j\|^2 = \|e_1 + e_j - P_j(e_1 + e_j)\|^2 = \|e_j + e_1 - P_j e_1\|^2 = 1 + \|e_1 - P_j e_1\|^2 \geq 1,$$

for all j. So $\{x_i\}_{i=1}^{\infty}$ is δ-separated, where $\delta = 1$.

The next proposition presents a fundamental property of separated sequences.

Proposition 2.66. *If a family of vectors $\{x_i\}_{i=1}^{\infty}$ is separated, then there are vectors $\{y_i\}_{i=1}^{\infty}$ satisfying:*

$$\langle y_i, x_j \rangle = \delta_{ij}, \text{ for all i,j.}$$

If it is δ-separated then, $\sup_i \|y_i\| < \infty$.

Proof. Fix j and let P_j be the orthogonal projection onto $\overline{\text{span}}\{x_i\}_{i \neq j}$. Note that $P_j x_j \neq x_j$ and so $(I - P_j)x_j \neq 0$.

Clearly,

$$\langle (I - P_j)x_j, x_i \rangle = 0 \text{ for } i \neq j.$$

So let $y_j = \dfrac{(I - P_j)x_j}{\|(I - P_j)x_j\|^2}$, and we get the desired sequence. For the δ-separated case, we have that $\|(I - P_j)x_j\| \geq \delta$ and the result follows. ∎

For the next result, we will need:

Proposition 2.67. *Let $\{x_i\}_{i=1}^{\infty}$ be a bounded sequence in a Hilbert space $\mathbb{H}$. The following are equivalent:*

1. *For some $\delta > 0$, $\{x_i\}_{i=1}^{\infty}$ is δ-separated.*

2. *$\{x_i\}_{i=1}^{\infty}$ is separated and $\{x_i\}_{i=n}^{\infty}$ is δ_1-separated for some $\delta_1 > 0$, for some $n \geq 1$.*

Proof. We just need to show that (2) $\Rightarrow$ (1). So assume $\{x_i\}_{i=1}^{\infty}$ is separated and $\{x_i\}_{i=n}^{\infty}$ is δ_1-separated. Let P_j be the projection onto $\overline{\text{span}}\{x_i\}_{i \neq j}$, for $j = 1, 2, \ldots,$ and let Q_j be the projection onto $\overline{\text{span}}\{x_i\}_{n \leq i \neq j}$, for $j = n, n+1, \ldots$. So

$$\|(I - Q_j)x_j\| \geq \delta_1, \text{ for all } j \geq n.$$

We need to show that there exists a $\delta > 0$ so that

$$\|(I - P_j)x_j\| \geq \delta, \text{ for all } j \geq 1.$$

We will do this in steps.

Step 1: There exists a $\delta_2 > 0$ so that

$$\|(I - P_j)x_j\| \geq \delta_2, \text{ for all } j \geq n.$$

We will do this by way of contradiction. So assume there are natural numbers $n \leq n_1 < n_2 < \cdots$ satisfying:

$$\|x_{n_j} - P_{n_j}(x_{n_j})\| < \frac{1}{j}.$$

It follows that there are vectors $y_j \in \text{span}\{x_i\}_{i=1}^{n-1}$ and $z_j \in \text{span}\{x_i\}_{n \leq i \neq n_j < \infty}$ so that $\|x_{n_j} - (y_j + z_j)\| < \frac{1}{j}$.

Claim 1: There are an $\epsilon > 0$ and $n_0 \in \mathbb{N}$ so that $\|y_j\| \geq \epsilon$, for all $j \geq n_0$.

We prove the claim by way of contradiction. If the claim fails, there are integers $j_1 < j_2 < \cdots$ so that $\|y_{j_k}\| < \frac{1}{k}$ for all $k = 1, 2, \ldots$. It follows that

$$\|x_{n_{j_k}} - z_{j_k}\| \leq \|x_{n_{j_k}} - (z_{j_k} + y_{j_k})\| + \|y_{j_k}\| < \frac{2}{k}, \text{ for all } k,$$

which contradicts the fact that $\{x_i\}_{i=n}^{\infty}$ is δ-separated.

Claim 2: There is a constant $K > 0$ so that $\|y_j\| \leq K$, for all $j \geq n_0$.

Define

$$\gamma = \inf\{\|u - v\| : u \in \text{span}\{x_i\}_{i=1}^{n-1}, v \in \overline{\text{span}}\{x_i\}_{i=n}^{\infty}, \|u\| = 1\}.$$

We will show that $\gamma > 0$. Indeed, if $\gamma = 0$ then there are sequences $\{u_j\}_{j=1}^{\infty} \subset \text{span}\{x_i\}_{i=1}^{n-1}, \|u_j\| = 1$, for all j, and $\{v_j\}_{j=1}^{\infty} \subset \overline{\text{span}}\{x_i\}_{i=n}^{\infty}$ so that

$$\|u_j - v_j\| \to 0 \text{ as } j \to \infty.$$

By switching to a subsequence if necessary, we may assume $u_j \to u \in \text{span}\{x_i\}_{i=1}^{n-1}$ and $u \neq 0$. Since

$$\|v_j - u\| \leq \|v_j - u_j\| + \|u_j - u\|,$$

we conclude that $v_j \to u \in \overline{\text{span}}\{x_i\}_{i=n}^{\infty}$. Thus,

$$u \in \text{span}\{x_i\}_{i=1}^{n-1} \cap \overline{\text{span}}\{x_i\}_{i=n}^{\infty}.$$

Since $u \in \text{span}\{x_i\}_{i=1}^{n-1}, u \neq 0$, we can write $u = \sum_{i=1}^{n-1} \alpha_i x_i$ for some scalars $\alpha_i's$ not all zero. Without loss of generality, we can assume $\alpha_1 \neq 0$. Then

$$x_1 = \frac{1}{\alpha_1}\left(u - \sum_{i=2}^{n-1} \alpha_i x_i\right) \in \overline{\text{span}}\{x_i\}_{i=2}^{\infty},$$

which contradicts the fact that $\{x_i\}_{i=1}^{\infty}$ is separated. So, $\gamma > 0$.

Now we have

$$\|\frac{y_j + z_j}{\|y_j\|}\| \geq \gamma, \text{ for all } j \geq n_0,$$

and $\sup_{j \geq 1} \|x_j\|$ is finte. Therefore, there is some $K > 0$ such that

$$\|y_j\| \leq \frac{1}{\gamma}\|y_j + z_j\| \leq \frac{1}{\gamma}(\|y_j + z_j - x_{n_j}\| + \|x_{n_j}\|) \leq K, \text{ for all } j \geq n_0.$$

The Claim 2 is proven.

Now since $\epsilon \leq \|y_j\| \leq K$ for all $j \geq n_0$, it has a convergent subsequence $y_{j_k} \to y \in \text{span}\{x_i\}_{i=1}^{n-1}$, and $y \neq 0$.

From the fact that

$$\|x_{n_{j_k}} - z_{j_k} - y\| \leq \|x_{n_{j_k}} - z_{j_k} - y_{j_k}\| + \|y_{j_k} - y\| \leq \frac{1}{j_k} + \|y_{j_k} - y\|,$$

we conclude $x_{n_{j_k}} - z_{j_k} \to y \in \overline{\text{span}}\{x_i\}_{i=n}^{\infty}$ as $k \to \infty$. Thus,

$$y \in \text{span}\{x_i\}_{i=1}^{n-1} \cap \overline{\text{span}}\{x_i\}_{i=n}^{\infty}$$

By the same argument as in the proof of Claim 2, this leads to a contradiction with the fact that $\{x_i\}_{i=1}^{\infty}$ is separated.

Step 2: There exists a $\delta > 0$ so that

$$\|(I - P_j)x_j\| \geq \delta, \text{ for all } j \geq 1.$$

Since $\{x_i\}_{i=1}^{\infty}$ is separated, for each $i = 1, 2, \ldots, n-1$, there exists $\epsilon_i > 0$ so that $\|(I - P_i)x_i\| \geq \epsilon_i$. Combined with Step 1, we have that $\{x_i\}_{i=1}^{\infty}$ is δ-separated, where $\delta = \min_{i=1,\ldots,n-1}\{\epsilon_i, \delta_2\}$. The proof of the Proposition is completed. ∎

Now we give a complete classification of when the state estimation problem is solvable for all measurement vectors in ℓ_1. Note that we have done it in complete generality and not assumed that $\{x_k\}_{k=1}^{\infty}$ is injective.

Theorem 2.68. *Let $\mathcal{X} = \{x_k\}_{k=1}^{\infty}$ be a frame for the real or complex space ℓ_2. The following are equivalent:*

1. *For every real vector $a = (a_1, a_2, \ldots) \in \ell_1$, there is a Hilbert Schmidt self-adjoint operator T so that*

   $$\langle Tx_k, x_k \rangle = a_k, \text{ for all } k = 1, 2, \ldots.$$

2. *The sequence $\{\tilde{x}_k\}_{k=1}^{\infty}$ is δ-separated.*

Proof. (1) $\Rightarrow$ (2): By (1), for each $k = 1, 2, \ldots$, there is a Hibert Schmidt self-adjoint operator R_k, and hence a vector $\tilde{R}_k \in \tilde{\mathbb{H}}$ so that

$$\langle \tilde{R}_k, \tilde{x}_l \rangle = \langle R_k x_l, x_l \rangle = \begin{cases} 1 & \text{if } k = l \\ 0 & \text{if } k \neq l. \end{cases}$$

It follows that $\tilde{x}_l \notin \overline{\text{span}}\{\tilde{x}_k\}_{k \neq l}$ and hence $\{\tilde{x}_k\}_{k=1}^{\infty}$ is separated. We now proceed by way of contradiction. Suppose that $\{\tilde{x}_k\}_{k=1}^{\infty}$ is not δ-separated. Then $\{\tilde{x}_k\}_{k=n}^{\infty}$ is not δ_n-separated for all n. Then for $n = 1$, there is $k_1 \geq 1$ such that

$$\|\tilde{x}_{k_1} - P_{k_1}(\tilde{x}_{k_1})\| < \frac{1}{2}.$$

Since $P_{k_1}(\tilde{x}_{k_1}) \in \overline{\text{span}}\{\tilde{x}_k\}_{k=1, k \neq k_1}^{\infty}$, there are some scalars $\alpha_k, k \in I$, where I is a finite subset of $\{k : k \geq 1, k \neq k_1\}$ such that

$$\|P_{k_1}(\tilde{x}_{k_1}) - \sum_{k \in I} \alpha_k \tilde{x}_k\| < \frac{1}{2}.$$

Let $y_1 = \sum_{k \in I} \alpha_k \tilde{x}_k$. Then

$$\|\tilde{x}_{k_1} - y_1\| < 1.$$

Now let $n_2 > \max\{k_1, k\}_{k \in I}$. Since $\{\tilde{x}_k\}_{k=n_2}^{\infty}$ is not δ_{n_2}-separated, similar to the above, there are numbers $n_2 \leq k_2 < n_3$ and a vector

$$y_2 \in \text{span}\{\tilde{x}_k : n_2 \leq k \neq k_2 < n_3\}$$

such that $\|\tilde{x}_{k_2} - y_2\| < \dfrac{1}{2^3}$. Continuing this procedure we can choose $1 = n_1 \leq k_1 < n_2 \leq k_2 < n_3 < \cdots$ and vectors

$$y_m \in \text{span}\{\tilde{x}_k : n_m \leq k \neq k_m < n_{m+1}\},$$

such that $\|\tilde{x}_{k_m} - y_m\| < \frac{1}{m^3}$, for all m. Now let $a = \{a_k\}_{k=1}^{\infty} \in \ell_1$, where

$$a_k = \begin{cases} \frac{1}{m^2} & \text{if } k = k_m \\ 0 & \text{otherwise .} \end{cases}$$

Then by assumption, there exists a Hilbert Schmidt self-adjoint operator T and a vector $\tilde{T} \in \tilde{\mathbb{H}}$ so that $\langle \tilde{T}, \tilde{x}_k \rangle = \langle Tx_k, x_k \rangle = a_k$ for all k. But then

$$\frac{1}{m^2} = \langle \tilde{T}, \tilde{x}_{k_m} \rangle = \langle \tilde{T}, \tilde{x}_{k_m} - y_m \rangle \leq \|\tilde{T}\| \|\tilde{x}_{k_m} - y_m\| \leq \|\tilde{T}\| \frac{1}{m^3},$$

which implies $\|\tilde{T}\| \geq m$ for all m, a contradiction.

(2) $\Rightarrow$ (1): Since $\{\tilde{x}_k\}_{k=1}^{\infty}$ is δ-separated, by Proposition 2.66, there are vectors $\{\tilde{T}_k\}_{k=1}^{\infty}$ in $\tilde{\mathbb{H}}$ satisfying

$$\langle \tilde{T}_k, \tilde{x}_l \rangle = \begin{cases} 1 & \text{if } k = l \\ 0 & \text{if } k \neq l \end{cases}$$

for all $k, l \geq 1$, and $\sup_{k \geq 1} \|\tilde{T}_k\| < \infty$. Now, fix $a = (a_1, a_2, \ldots) \in \ell_1$ and let

$$\tilde{T} = \sum_{k=1}^{\infty} a_k \tilde{T}_k.$$

This series converges since $a \in \ell_1$ and $\sup_{k \geq 1} \|\tilde{T}_k\| < \infty$. Now, let T be the Hilbert Schmidt self-adjoint operator that corresponds with $\tilde{T}$. Then we have

$$\langle Tx_k, x_k \rangle = \langle \tilde{T}, \tilde{x}_k \rangle = a_k, \text{ for all } k = 1, 2, \ldots.$$

This completes the proof. ∎

Next we show that there is no injective frame for which the state estimation problem is solvable for all measurements taken from ℓ_2. Recall that for a Hilbert Schmidt self-adjoint operator T on the Hilbert space ℓ_2, the corresponding vector $\tilde{T}$ is defined as in the proof of Theorem 2.47 for the real case and Theorem 2.51 for the complex case.

Theorem 2.69. *There is no injective frame $\mathcal{X} = \{x_k\}_{k=1}^{\infty}$ in the real or complex space ℓ_2 so that for every $a = \{a_k\}_{k=1}^{\infty} \in \ell_2$, there is a self-adjoint Hilbert Schmidt*

operator T so that

$$\langle Tx_k, x_k\rangle = a_k, \ \textit{for all } k = 1, 2, \ldots.$$

Proof. We will proceed by way of contradiction. The proof is divided into steps.

Suppose that there is an injective frame $\mathcal{X} = \{x_k\}_{k=1}^{\infty}$ for which the state estimation problem is solvable for all choices $\{a_k\}_{k=1}^{\infty} \in \ell_2$.

Step I: There are vectors $\tilde{R}_k \in \tilde{\mathbb{H}}, k = 1, 2, \ldots$ so that $\langle \tilde{R}_k, \tilde{x}_l\rangle = \delta_{kl}$.

This is immediate because by assumption, for each $k = 1, 2, \ldots$, there is a Hilbert Schmidt self-adjoint operator R_k so that

$$\langle \tilde{R}_k, \tilde{x}_l\rangle = \langle R_k x_l, x_l\rangle = \begin{cases} 1 & \text{if } k = l \\ 0 & \text{if } k \neq l. \end{cases}$$

Denote $E_n = \text{span}\{\tilde{x}_k\}_{k=1}^{n}$ and let P_n be the projection onto E_n.

Step II: If there is a real vector $\{a_k\}_{k=1}^{\infty} \in \ell_2$ satisfying $\sup_n \| \sum_{k=1}^{n} a_k \tilde{R}_k\| = \infty$, then there is a real vector $\{b_k\}_{k=1}^{\infty} \in \ell_2$ and $n_1 < n_2 < \cdots$ so that

$$\left\| P_{n_j} \left(\sum_{k=1}^{n_j} b_k \tilde{R}_k \right) \right\| \geq j.$$

Indeed, since $\sup_n \| \sum_{k=1}^{n} a_k \tilde{R}_k\| = \infty$, we can choose a sequence $m_1 < m_2 < \cdots$ so that $\left\| \sum_{k=1}^{m_j} a_k \tilde{R}_k \right\| \geq 2j$.

For any $j > 1$, we have

$$\begin{aligned} \left\| \sum_{k=1}^{m_1} a_k \tilde{R}_k - \sum_{k=m_1+1}^{m_j} a_k \tilde{R}_k \right\| &\geq \left\| \sum_{k=1}^{m_j} a_k \tilde{R}_k \right\| - 2 \left\| \sum_{k=1}^{m_1} a_k \tilde{R}_k \right\| \\ &\geq 2j - 2 \left\| \sum_{k=1}^{m_1} a_k \tilde{R}_k \right\|. \end{aligned}$$

Combining this with the fact that $E_1 \subset E_2 \subset \ldots$ and $\cup_{n=1}^{\infty} E_n$ is dense in $\tilde{\mathbb{H}}$, we can choose j large enough so that

$$\left\| P_{m_j} \left(\sum_{k=1}^{m_1} a_k \tilde{R}_k \right) \right\| \geq \frac{1}{2} \left\| \sum_{k=1}^{m_1} a_k \tilde{R}_k \right\|,$$

and

$$\left\| \sum_{k=1}^{m_1} a_k \tilde{R}_k - \sum_{k=m_1+1}^{m_j} a_k \tilde{R}_k \right\| \geq 4.$$

Since

$$\begin{aligned} &\|P_{m_j}(\sum_{k=1}^{m_1} a_k \tilde{R}_k) + P_{m_j}(\sum_{k=m_1+1}^{m_j} a_k \tilde{R}_k)\|^2 + \|P_{m_j}(\sum_{k=1}^{m_1} a_k \tilde{R}_k) - P_{m_j}(\sum_{k=m_1+1}^{m_j} a_i \tilde{R}_k)\|^2 \\ &\qquad = 2 \left(\|P_{m_j}(\sum_{k=1}^{m_1} a_k \tilde{R}_k)\|^2 + \|P_{m_j}(\sum_{k=m_1+1}^{m_j} a_k \tilde{R}_k)\|^2 \right) \\ &\qquad \geq 2\|P_{m_j}(\sum_{k=1}^{m_1} a_k \tilde{R}_k)\|^2, \end{aligned}$$

we can choose $b_i = a_i$ for $i = 1, \ldots, m_1$ and $b_i \in \{a_i, -a_i\}$ for $i = m_1 + 1, \ldots, m_j$ so that

$$\|P_{m_j}(\sum_{k=1}^{m_j} b_k \tilde{R}_k)\| \geq \|P_{m_j}(\sum_{k=1}^{m_1} b_k \tilde{R}_k)\| \geq \frac{1}{2}\|\sum_{k=1}^{m_1} b_k \tilde{R}_k\| \geq 1 \text{ and } \|\sum_{k=1}^{m_j} b_k \tilde{R}_k\| \geq 4.$$

Setting $n_1 = m_j$,

$$\|P_{n_1}(\sum_{k=1}^{n_1} b_k \tilde{R}_k)\| \geq 1 \text{ and } \|\sum_{k=1}^{n_1} b_k \tilde{R}_k\| \geq 4.$$

Now for m_j above, by the same argument, there is $m_l > m_j$ and $b_i \in \{a_i, -a_i\}$ for $i = m_j + 1, \ldots, m_l$ so that

$$\left\| P_{m_l} \left(\sum_{k=1}^{m_l} b_k \tilde{R}_k \right) \right\| \geq \left\| P_{m_l} \left(\sum_{k=1}^{m_j} b_k \tilde{R}_k \right) \right\| \geq \frac{1}{2} \left\| \sum_{k=1}^{m_j} b_k \tilde{R}_k \right\| \geq 2$$

and $\left\| \sum_{k=1}^{m_l} b_k \tilde{R}_k \right\| \geq 6$. Set $n_2 = m_l$ we get

$$\left\| P_{n_2} \left(\sum_{k=1}^{n_2} b_k \tilde{R}_k \right) \right\| \geq 2.$$

Continuing this process inductively, the result follows.

Step III: For all vectors $\{a_k\}_{k=1}^{\infty} \in \ell_2$, $\sup_n \| \sum_{k=1}^{n} a_k \tilde{R}_k \|$ is finite.

Suppose by way of contradiction that there is a vector $\{a_k\}_{k=1}^{\infty} \in \ell_2$ so that $\sup_n \| \sum_{k=1}^{n} a_k \tilde{R}_k \| = \infty$. Let $\{b_k\}_{k=1}^{\infty}$ be the vector in Step II, then there exists a vector $\tilde{T} \in \tilde{\mathbb{H}}$ so that $\langle \tilde{T}, \tilde{x}_k \rangle = b_k$, for all $k = 1, 2, \ldots$. It follows that

$$P_{n_j} \tilde{T} = P_{n_j} \left(\sum_{k=1}^{n_j} b_k \tilde{R}_k \right)$$

for all $j = 1, 2, \ldots$. Hence,

$$\infty = \sup_j \left\| P_{n_j} (\sum_{k=1}^{n_j} b_k \tilde{R}_k) \right\| = \sup_j \| P_{n_j} \tilde{T} \| \leq \| \tilde{T} \|,$$

which is a contradiction.

Step IV: $\{\tilde{R}_k\}_{k=1}^{\infty}$ is a Bessel sequence in $\tilde{\mathbb{H}}$.

For each $n \in \mathbb{N}$, define an operator

$$\begin{aligned} T_n : \ell_2 &\longrightarrow \tilde{\mathbb{H}} \\ x = (a_1, a_2, \ldots) &\longmapsto T_n(x) = \sum_{k=1}^{n} a_k \tilde{R}_k \end{aligned}$$

Then T_n is a bounded linear operator for all n.

By Step III, $\sup_n \| \sum_{k=1}^{n} a_k \tilde{R}_k \|$ is finite for all $x = \{a_k\}_{k=1}^{\infty}$. By the Uniform Boundedness Principle, $\sup_n \|T_n\| \leq B$, for some $B > 0$. For any $n, m \in \mathbb{N}, m > n$, we have

$$\left\| \sum_{k=n+1}^{m} a_k \tilde{R}_k \right\|^2 = \left\| T_m \left(\sum_{k=n+1}^{m} a_k e_k \right) \right\|^2 \leq B^2 \sum_{k=n+1}^{m} a_k^2.$$

It follows that $\sum_{k=1}^{\infty} a_k \tilde{R}_k$ converges, and hence $\{\tilde{R}_k\}_{k=1}^{\infty}$ is Bessel.

Step V: We show $\{\tilde{x}_k\}_{k=1}^{\infty}$ is a frame, a contradiction.

We have shown that under our assumption, $\{\tilde{R}_k\}_{k=1}^{\infty}$ is B^2-Bessel for some B. Now choose any $a = \{a_k\}_{k=1}^{\infty} \in \ell_2$. We have that

$$\left\| \sum_{k=1}^{\infty} a_k \tilde{R}_k \right\|^2 \leq B^2 \sum_{k=1}^{\infty} a_k^2.$$

By Theorem 2.53, $\sum_{k=1}^{\infty} a_k \tilde{x}_k$ converges. Now, we have

$$\begin{aligned} \left\| \sum_{k=1}^{\infty} a_k \tilde{x}_k \right\| &= \sup_{\|x\| \leq 1} \left| \left\langle x, \sum_{k=1}^{\infty} a_k \tilde{x}_k \right\rangle \right| \\ &\geq \frac{1}{B\|a\|} \left| \left\langle \sum_{k=1}^{\infty} a_k \tilde{R}_k, \sum_{l=1}^{\infty} a_l \tilde{x}_l \right\rangle \right| \\ &= \frac{1}{B\|a\|} \left| \sum_{k,l=1}^{\infty} a_k a_l \langle \tilde{R}_k, \tilde{x}_l \rangle \right| \\ &= \frac{1}{B} \|a\|. \end{aligned}$$

It follows that $\{\tilde{x}_k\}_{k=1}^{\infty}$ has a positive lower Riesz bound and since this family is injective, it is a Riesz basis. Hence, it is a frame sequence. This contradicts Theorem 2.53, completing the proof. ∎

As in the finite dimensional case, often times the state estimation problem is not solvable. As before there is a natural way to get a good estimation to the solution. Given a frame $\{x_k\}_{k=1}^{\infty}$ and $\{a_k\}_{k=1}^{\infty} \in \ell_2$, choose m so that $\sum_{k=m+1}^{\infty} a_k^2 \leq \epsilon$. Then apply the argument in the finite case to get the best solution for $\{a_k\}_{k=1}^{m}$.

Bibliography

[1] S. Bahmanpour, J. Cahill, P. G. Casazza, J. Jasper, and L. M. Woodland, *Phase retrieval and norm retrieval*, arXiv preprint arXiv:1409.8266, (2014).

[2] R. Balan, *Stability of frames which give phase retrieval*, Houston Journal of Mathematics, (2015).

[3] R. Balan, P. Casazza, and D. Edidin, *On signal reconstruction without phase*, Applied and Computational Harmonic Analysis, 20 (2006), pp. 345–356.

[4] A. S. Bandeira, J. Cahill, D. G. Mixon, and A. A. Nelson, *Saving phase: Injectivity and stability for phase retrieval*, Applied and Computational Harmonic Analysis, 37 (2014), pp. 106–125.

[5] R. Bates and D. Mnyama, *The status of practical fourier phase retrieval*, in Advances in Electronics and Electron physics, vol. 67, Elsevier, 1986, pp. 1–64.

[6] C. Becchetti and L. P. Ricotti, *Speech recognition theory and c++ implementation*, John WILEY&Sons, Ltd, (1999), pp. 125–137.

[7] J. J. Benedetto and A. Kebo, *The role of frame force in quantum detection*, Journal of Fourier Analysis and Applications, 14 (2008), pp. 443–474.

[8] B. Bodmann and J. Haas, *A short history of frames and quantum designs*, arXiv preprint arXiv:1709.01958, (2017).

[9] H. Bolcskei and Y. C. Eldar, *Geometrically uniform frames*, IEEE Transactions on Information Theory, 49 (2003), pp. 993–1006.

[10] S. Botelho-Andrade, P. G. Casazza, H. Van Nguyen, and J. C. Tremain, *Phase retrieval versus phaseless reconstruction*, Journal of Mathematical Analysis and Applications, 436 (2016), pp. 131–137.

[11] J. Cahill, P. Casazza, and I. Daubechies, *Phase retrieval in infinite-dimensional hilbert spaces*, Transactions of the American Mathematical Society, Series B, 3 (2016), pp. 63–76.

[12] J. Cahill, P. G. Casazza, J. Peterson, and L. Woodland, *Phase retrieval by projections*, arXiv preprint arXiv:1305.6226, (2013).

[13] P. G. Casazza and D. Cheng, *Associating vectors in $\mathbb{C}^n$ with rank 2 projections in $\mathbb{R}^{2n}$: with applications*, arXiv preprint arXiv:1703.02657, (2017).

[14] P. G. Casazza and N. J. Kalton, *Generalizing the paley-wiener perturbation theory for banach spaces*, Proceedings of the American Mathematical Society, (1999), pp. 519–527.

[15] P. G. Casazza and M. Leon, *Existence and construction of finite frames with a given frame operator*, Int. J. Pure Appl. Math, 63 (2010), pp. 149–158.

[16] P. G. Casazza and R. G. Lynch, *A brief introduction to hilbert space frame theory and its applications*, Finite Frame Theory: A Complete Introduction to Overcompleteness, 93 (2016), p. 2.

[17] P. G. Casazza, E. Pinkham, and B. Tuomanen, *Riesz outer product hilbert space frames: Quantitative bounds, topological properties, and full geometric characterization*, Journal of Mathematical Analysis and Applications, 441 (2016), pp. 475–498.

[18] O. Christensen, *An introduction to frames and Riesz bases*, Springer, 2016.

[19] D. Cochran, S. Howard, and B. Moran, *Positive-operator-valued measures: a general setting for frames*, in Excursions in Harmonic Analysis, Volume 2, Springer, 2013, pp. 49–64.

[20] J. Drenth, *Principles of protein X-ray crystallography*, Springer Science & Business Media, 2007.

[21] D. Edidin, *Projections and phase retrieval*, Applied and Computational Harmonic Analysis, 42 (2017), pp. 350–359.

[22] Y. C. Eldar, *Von Neumann measurement is optimal for detecting linearly independent mixed quantum states*, Physical Review A, 68 (2003), p. 052303.

[23] Y. C. Eldar and G. D. Forney, *Optimal tight frames and quantum measurement*, IEEE Transactions on Information Theory, 48 (2002), pp. 599–610.

[24] J. R. Fienup, *Reconstruction of an object from the modulus of its fourier transform*, Optics letters, 3 (1978), pp. 27–29.

[25] ——, *Phase retrieval algorithms: a comparison*, Applied optics, 21 (1982), pp. 2758–2769.

[26] D. Han, D. Larson, B. Liu, and R. Liu, *Operator-valued measures, dilations, and the theory of frames*, vol. 229, American Mathematical Society, 2014.

[27] P. Hausladen and W. K. Wootters, *A 'pretty good' measurement for distinguishing quantum states*, Journal of Modern Optics, 41 (1994), pp. 2385–2390.

[28] C. W. Helstrom, *Quantum detection and estimation theory*, Journal of Statistical Physics, 1 (1969), pp. 231–252.

[29] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge university press, 1985.

[30] R. Kennedy, M. Lax, and H. Yuen, *Optimum testing of multiple hypotheses in quantum detection theory*, IEEE Transactions on Information Theory, 21 (1975), pp. 125–134.

[31] B. I. Levin, *Distribution of zeros of entire functions*, vol. 5, American Mathematical Soc., 1964.

[32] A. Peres and D. R. Terno, *Optimal distinction between nonorthogonal quantum states*, J. Phys., A31 (1998), pp. 7105–7112.

[33] L. R. Rabiner, B.-H. Juang, and J. C. Rutledge, *Fundamentals of speech recognition*, vol. 14, PTR Prentice Hall Englewood Cliffs, 1993.

[34] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, *Symmetric informationally complete quantum measurements*, Journal of Mathematical Physics, 45 (2004), pp. 2171–2180.

[35] A. J. Scott, *Tight informationally complete quantum measurements*, Journal of Physics A: Mathematical and General, 39 (2006), p. 13507.

[36] C. Vinzant, *A small frame and a certificate of its injectivity*, in 2015 International Conference on Sampling Theory and Applications (SampTA), IEEE, 2015, pp. 197–200.

VITA

Sara Botelho-Andrade was born in Berkeley, California. She was raised in Memphis, TN and attended the University of Memphis for both her baccalaureate and master's degrees in mathematical sciences; with a master's thesis on the isometric equivalence problem. During her time at the University of Missouri, she has worked with Professor Peter Casazza at the Frame Research Center. After graduation, Sara will be participating in the Repperger Research Intern Program conducting research at the Air Force Research Lab in San Antonio, Texas. The following semester she will join the University of Denver as a postdoc.