Annals of Mathematical Sciences and Applications
Volume 3, Number 1, 337–365, 2018

Data recovery on a manifold from linear samples: theory and computation

Jian-Feng Cai*, Yi Rong, Yang Wang†, and Zhiqiang Xu‡

Dedicated to David Mumford's 80th Birthday

Data recovery on a manifold is an important problem in many applications. Many such problems, e.g. compressive sensing, involve solving a system of linear equations knowing that the unknowns lie on a known manifold. The aim of this paper is to survey theoretical results and numerical algorithms about the recovery of signals lying on a manifold from linear measurements. Particularly, we focus on the case where signals lying on an algebraic variety. We first introduce the tools from algebraic geometry which plays an important role in studying the minimal measurement number and also show its applications. We finally introduce the numerical algorithms for solving it.

AMS 2000 subject classifications: Primary 42C15.
Keywords and phrases: Frames, phase retrieval.

1. Introduction

Solving systems of linear equations $A\mathbf{x} = \mathbf{b}$ is ubiquitous in all areas of science and engineering. This problem has been well studied even before Gauss introduced the Gaussian elimination method. Thus one may even wonder whether there is anything we don't already know about solving a system of linear equations.

Traditional systems of linear equations typically assume that the number of equations is no less than the number of unknowns. Failing it we have an under-determined system of linear equations where the solution will not

*Jian-Feng Cai was supported by Hong Kong Research Grant Council grant 16300616 and 16306317.

†Yang Wang was supported in part by the Hong Kong Research Grant Council grant 16306415 and 16317416.

‡Zhiqiang Xu was supported by NSFC grant (11422113, 91630203, 11331012) and by National Basic Research Program of China (973 Program 2015CB856000).

be unique. Solving an under-determined system would require additional regularization. However, in recent years there has been an explosion in the study of *compressive sensing* where under the condition of sparsity one may solve a significantly under-determined system of linear equations, see e.g. [1, 2, 3] among the vast literature.

It turns out that there is a more general framework in which an under-determined system of linear equations can be solved, namely we know a priori that the solution to the system lies on certain subset of the Euclidean space. For example, often the solution may lie on a lower dimensional manifold. There are many applications under this general framework, and here we list some of the best known ones.

Example 1: phase retrieval

The classical phase retrieval problem concerns the reconstruction of a function (typically the density or structure function of certain material) from the magnitude of its Fourier transform (X-ray diffraction). In recent years this problem has been broadened to encompass all problems involving the recovery of a function or signal from the magnitude of its samples (usually linear samples). Such problems arise in many important applications in imaging, optics, communication, audio signal processing and more [4, 5, 6, 7, 8, 9, 10].

The precise statement of the phase retrieval problem in this setting is:

The Phase Retrieval Problem. *Let $\{\mathbf{f}_j\}_{j=1}^N$ be a set of vectors in $\mathbb{F}^d$, where $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$. Can we reconstruct any $\mathbf{x} \in \mathbb{F}^d$ up to a unimodular scalar from $\{|\langle \mathbf{x}, \mathbf{f}_j \rangle|^2\}$, and if so, how?*

We say that $\{\mathbf{f}_j\}_{j=1}^N$ in $\mathbb{F}^d$ have the *phase retrieval property*, or are phase retrievable, if any $\mathbf{x} \in \mathbb{F}^d$ can be recovered up to a unimodular scalar from $\{|\langle \mathbf{x}, \mathbf{f}_j \rangle|^2\}$.

Note that in phase retrieval one cannot distinguish $\mathbf{x}$ from $c\mathbf{x}$, where c is a unimodular constant in $\mathbb{F}$. This ambiguity can be removed by reformulating the phase retrieval problem as recovering the rank one Hermitian matrix $X = \mathbf{x}\mathbf{x}^* \in \mathbb{F}^{d\times d}$. Given the magnitude measurements $|\langle \mathbf{x}, \mathbf{f}_j \rangle|^2 = b_j$, $j = 1, \ldots, N$, set $F_j = \mathbf{f}_j \mathbf{f}_j^*$. Then we have

$$b_j = |\langle \mathbf{x}, \mathbf{f}_j \rangle|^2 = \mathbf{x}^* \mathbf{f}_j \mathbf{f}_j^* \mathbf{x} = \operatorname{tr}(\mathbf{x}^* F_j \mathbf{x}) = \operatorname{tr}(F_j X), \quad j = 1, \ldots, N. \tag{1.1}$$

Thus the phase retrieval problem is an example of a system of linear equations $\mathbf{L}(X) = \mathbf{b}$ where $X \in \mathbb{F}^{d\times d}$ is a rank one Hermitian matrix.

A more general version of the phase retrieval problem is to recover a vector $\mathbf{x} \in \mathbb{F}^d$ from a finite number of quadratic measurements $\{\mathbf{x}^* A_j \mathbf{x}\}_{j=1}^N$ where each A_j is a Hermitian matrix in $\mathbb{F}^{d\times d}$. Again, we say a set of Hermitian matrices $\{A_j\}_{j=1}^N$ in $\mathbb{F}^{d\times d}$ have the *phase retrieval property* if any $\mathbf{x} \in \mathbb{F}^d$ can be recovered up to a unimodular scalar from the quadratic measurements $\{\mathbf{x}^* A_j \mathbf{x}\}_{j=1}^N$. This generalized version is studied in [11], and in special cases such as for orthogonal projection matrices $\{A_j\}_{j=1}^N$ in other papers [8, 10, 12]. Let $b_j = \mathbf{x}^* A_j \mathbf{x}$ and $X = \mathbf{x}\mathbf{x}^*$. Then we have similarly

$$b_j = \mathbf{x}^* A_j \mathbf{x} = \mathrm{tr}(\mathbf{x}^* A_j \mathbf{x}) = \mathrm{tr}(A_j X), \qquad j = 1, \ldots, N. \tag{1.2}$$

Like the original phase retrieval problem, the generalized phase retrieval problem also solves a system of linear equations where the unknown X is a rank one Hermitian matrix.

Example 2: low rank matrix recovery

The matrix recovery problem is an active topic recently. The general formulation of the problem is that there is a $X \in \mathbb{F}^{p\times q}$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and we are given some measurements (also called *samples*) of X. We would like to recover the matrix X from those measurements or samples. Matrix recovery is widely used in image processing, system identification and control, Euclidean embedding, and recommender systems (see [15, 16, 17]). We state the problem as follows: For $1 \le j \le N$ let $L_j : \mathbb{F}^{p\times q} \longrightarrow \mathbb{F}$ be linear functions, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Suppose that for an $X \in \mathbb{F}^{p\times q}$ with $\mathrm{rank}(X) \le r$ we are given the values $L_j(X)$ for $1 \le j \le N$. The question is: can we recover X? This problem is another example of solving a system of linear equations $\mathbf{L}(X) = \mathbf{b}$ where $X \in \mathbb{F}^{p\times q}$, but with the a priori knowledge that X is in the manifold consisting of rank r or less matrices. Note that we can always represent the linear function L_j by $L_j(X) = \mathrm{tr}(A_j^T X)$ for some $A_j \in \mathbb{F}^{d\times d}$.

Example 3: compressive sensing

In compressive sensing, we aim to solve a system of linear equations $A\mathbf{x} = \mathbf{b}$, where $A \in \mathbb{F}^{N\times d}$ and $\mathbf{x} \in \mathbb{F}^N, \mathbf{b} \in \mathbb{F}^d$ with $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$, with the knowledge that $\mathbf{x}$ being sparse with sparsity $\|\mathbf{x}\|_0 \le k \ll d$. Here $\|\mathbf{x}\|_0$ denotes the number of nonzero entries of $\mathbf{x}$ (the *sparsity*). Let

$$\mathbb{F}_k^d = \Big\{\mathbf{x} \in \mathbb{F}^d : \ \|\mathbf{x}\|_0 \le k\Big\}.$$

Clearly $\mathbb{F}_k^d$ is a finite union of k-dimensional subspaces in $\mathbb{F}^d$. Thus compressive sensing is equivalent to solving a system of linear equations where the solution $\mathbf{x}$ is known to lie on $\mathbb{F}_k^d$.

Example 4: the projection retrieval problem

Assume that we have a real or complex orthogonal projection matrix $P \in \mathbb{F}^{d\times d}$ with rank r which means P satisfies $P^* = P$ and $P^2 = P$. The Projection Retrieval Problem considers the following question: Let $\mathbf{v}_1, \ldots, \mathbf{v}_N \in \mathbb{F}^d$ be sample points where we measure $\|P\mathbf{v}_j\| = a_j$ for $1 \le j \le N$. Can we determine the projection matrix P from these measurements $\{a_j\}_{j=1}^N$?

This problem is related to both phase retrieval and low rank matrix recovery (see [13, 14]). It is also an example of solving a system of linear equations on a manifold. Let $A_j = \mathbf{v}_j\mathbf{v}_j^*$. Note that $P^2 = P$ and $P^* = P$ from the orthogonal projection property. We have

$$a_j^2 = \|P\mathbf{v}_j\|^2 = \mathbf{v}_j^* P^* P\mathbf{v}_j = \mathrm{tr}(\mathbf{v}_j^* P\mathbf{v}_j) = \mathrm{tr}(F_j P) \qquad \text{for all } j.$$

The Projection Retrieval Problem is thus an example of solving a system of linear equations where the unknown P lies on the set of all rank r orthogonal projections.

Example 5: the missing distance problem

Consider a set of points $S = \{\mathbf{x}_j\}_{j=0}^N$ in $\mathbb{F}^d$ and let $a_{ij} := \|\mathbf{x}_i - \mathbf{x}_j\|^2$ for all $0 \le i, j \le N$. It is well known and easy to show that the values $\{a_{ij}\}$ uniquely determine the point set S up to an Euclidean isometry. The *Missing Distance Problem* asks whether S can be uniquely determined up to an Euclidean isometry from only a subset of $\{a_{ij}\}$.

The Missing Distance Problem can also be formulated as solving a system of linear equations on a manifold. First through a translation we can normalize the set S by having $\mathbf{x}_0 = 0$. Under this normalization let

$$F_S := [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N], \qquad \text{and } X = F_S^* F_S,$$

i.e. $F_S \in \mathbb{F}^{d\times N}$ has $\mathbf{x}_j$ as its j-th column. Then the missing distance problem is equivalent to recovering the N by N matrix X from a subset of $\{a_{ij}\}$.

Now observe that $a_{ij} = a_{ji}$ and we have

$$a_{0j} = \|\mathbf{x}_j - \mathbf{x}_0\|^2 = \|\mathbf{x}_j\|^2 = x_{jj} \quad \text{if } j \ge 1,$$

$$\begin{aligned} a_{ij} &= \|\mathbf{x}_i - \mathbf{x}_j\|^2 = \mathbf{x}_i^*\mathbf{x}_i + \mathbf{x}_j^*\mathbf{x}_j - \mathbf{x}_i^*\mathbf{x}_j - \mathbf{x}_j^*\mathbf{x}_i \\ &= x_{ii} + x_{jj} - x_{ij} - x_{ji} \quad \text{if } i, j \geq 1. \end{aligned}$$

Thus each a_{ij} is a linear measurement of X, and the Missing Distance Problem is also a problem of solving a system of linear equations where the unknown $X \in \mathbb{F}^{N\times N}$ is on the manifold of positive semi-definite Hermitian matrices of rank at most d.

In this paper we examine the general problem of recovering an unknown X on a manifold from a system of its linear measurements, from both the theoretical and computational angle. We assume that X is in the Euclidean space $\mathbb{F}^d$ where $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$, and it lies on a lower dimensional manifold $\mathcal{M}$ in $\mathbb{F}^d$. A main question we ask is whether we can recover $X \in \mathcal{M}$ from a significantly under-determined system of linear measurements. This question, putting in the context of phase retrieval and low rank matrix recovery, is one of the fundamental questions still being actively studied today. Of course with enough measurements, e.g. when the system is not under-determined in $\mathbb{F}^d$, we can always recover X. So the real question is: Can we still fully recover $X \in \mathcal{M}$ with a significantly reduced number of linear measurements? The answer is yes in many cases. We may also ask a weaker question: Can we recover X for *almost all* $X \in \mathcal{M}$ (but not all) with even fewer linear measurements? Both questions have been studied for phase retrieval and matrix recovery.

Definition 1.1. Let $\mathcal{M} \subset \mathbb{F}^d$ where $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$. Let $\mathbf{L} : \mathbb{F}^d \longrightarrow \mathbb{F}^N$ be a linear map. We say $\mathbf{L}$ has the *$\mathcal{M}$-recovery property* if $\mathbf{L}$ is injective on $\mathcal{M}$. It has the *almost everywhere $\mathcal{M}$-recovery property* if for almost every $X \in \mathcal{M}$, we have $\mathbf{L}^{-1}\big(\mathbf{L}(X)\big) \cap \mathcal{M} = \{X\}$.

In other words $\mathbf{L}$ has the $\mathcal{M}$-recovery property if any $X \in \mathcal{M}$ is uniquely determined by $\mathbf{L}(X)$, and $\mathbf{L}$ has the almost everywhere $\mathcal{M}$-recovery property if almost all $X \in \mathcal{M}$ is uniquely determined by $\mathbf{L}(X)$. Here the easiest way to define "almost everywhere" and "almost all" is through the Hausdorff measure on $\mathcal{M}$. But since our study only focuses on $\mathcal{M}$ that are "nice" such as manifolds or algebraic varieties there should be no ambiguity.

The theoretical part of this paper studies the following questions: Let $\mathcal{M} \subset \mathbb{F}^d$ where $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$. Let $\mathbf{L} : \mathbb{F}^d \longrightarrow \mathbb{F}^N$ be a linear map. How large should N be so that $\mathbf{L}$ has the $\mathcal{M}$-recovery property or the almost everywhere $\mathcal{M}$-recovery property? For example, there is an extensive literature in phase retrieval on choosing the measurement vectors $\{\mathbf{f}_j\}$ to be i.i.d. Gaussian. Here we will provide a framework for answering these questions. For phase retrieval and matrix completion we have developed techniques to make substantial progresses recently [21, 11, 18]. Our goal for this paper is more appropriately described as the combination of a survey and putting past work into a more unified framework. We also include some new results, such as for projection retrieval. On the computational part of this paper, we present some new techniques developed in [29, 30]. We hope that this paper will provide useful ideas and techniques to those researchers working in these aforementioned areas.

2. The algebraic geometry connection

The manifold recovery problem essentially comes down to examining the intersection of a set of hyperplanes defined by the system of linear equations with the manifold on which the unknown data lie. This is one of the classical areas in algebraic geometry, provided that the manifold in question is an algebraic variety. Fortunately this is precisely the case in most of the applications we are interested in. For example, for low rank matrix recovery we are studying recovery on the set $\mathcal{M}$ of all rank r or less matrices

$$\mathcal{M}_{p\times q,r}(\mathbb{F}) := \Big\{Q \in \mathbb{F}^{p\times q} : \mathrm{rank}(Q) \le r\Big\}, \qquad \mathbb{F} = \mathbb{R} \text{ or } \mathbb{C}, \tag{2.1}$$

which is known as a *determinantal variety* for $\mathbb{F} = \mathbb{C}$. In our study instead of considering general manifolds, our manifolds will actually be projective varieties. Before proceeding to the main results, we first introduce some basic notations related to projective spaces and varieties.

An algebraic variety (affine variety) $V \subseteq \mathbb{C}^d$ is the locus of a finite collection of polynomials in $\mathbb{C}[\mathbf{x}]$. In this paper we shall primarily consider *projective varieties*. They lie in the projective space $\mathbb{P}(\mathbb{C}^d)$, which is the space of all one dimensional subspaces of $\mathbb{C}^d$. Let $\sigma : \mathbb{C}^d\setminus\{0\} \longrightarrow \mathbb{P}(\mathbb{C}^d)$ be the canonical map $\sigma(\mathbf{x}) = [\mathbf{x}]$, where $[\mathbf{x}] \in \mathbb{P}(\mathbb{C}^d)$ denotes the line through $\mathbf{x}$. We shall also often consider the *projectivization* of a set $S \subset \mathbb{C}^d\setminus\{0\}$, to be $[S] = \sigma(S)$. A projective variety is the projectivization of an affine variety defined by homogeneous polynomials. But for simplicity, in this paper we adopt a looser terminology. Whenever there is no confusion, the phrase *projective variety in* $\mathbb{C}^d$ means an affine variety in $\mathbb{C}^d$ that is the locus of a finite collection of homogeneous polynomials. We shall use a *projective variety in* $\mathbb{P}(\mathbb{C}^d)$ to describe a true projective variety. A variety V is *irreducible* if it cannot be decomposed into $V = \bigcup_{j=1}^k V_j$ where $k > 1$ and V_j are distinct proper subvarieties. A reducible variety can be written as finite union of distinct irreducible subvarieties (irreducible components). Throughout the paper, by a *generic* point $\mathbf{x}$ in an algebraic variety V we mean $\mathbf{x} \in V \setminus Z$ where $Z \subset V$ is a subvariety with $\dim(Z) < \dim(V)$.

A set $U \subset \mathbb{C}^d$ is a *quasi-projective variety* if there exist two projective varieties V and Y with $Y \subset V$ such that $U = V \setminus Y$. The concept of dimension for a quasi-projective variety in $\mathbb{C}^d$ is very well defined (see [19]).

Note that a complex algebraic variety V may contain real points. We use $V_{\mathbb{R}}$ to denote the real points of V. The set $V_{\mathbb{R}}$ is itself a real algebraic variety, and its (real) dimension is well defined (see e.g. [20]). We use $\dim_{\mathbb{R}}(V_{\mathbb{R}})$ to denote the real dimension of $V_{\mathbb{R}}$. The following lemma is a key result in this area of study.

Lemma 2.1. *[8, 11] Let V be an algebraic variety in $\mathbb{C}^d$. Then $\dim_{\mathbb{R}}(V_{\mathbb{R}}) \leq \dim(V)$.*

The following theorem, which concerns the intersection of a hyperplane and a projective variety in $\mathbb{C}^d$, is well known and plays an important role in our study. Its proof can be found in any standard textbook in algebraic geometry, see [19].

Theorem 2.2. *Let V be a projective variety and P be a subspace in $\mathbb{C}^d$ with $\dim(P) = d-1$. Then $\dim(V \cap P) \geq \dim(V) - 1$. Furthermore, if P does not contain an irreducible component of V then $\dim(V \cap P) = \dim(V) - 1$.*

Note that the above theorem fails for real projective varieties. As a result, it is often easier to prove results for data recovery on a complex projective variety. We illustrate how the above results from algebraic geometry can be applied to give a very simple proof to the following result for matrix recovery, which was first proved in [21, 18]. The corresponding result does not hold for real matrix recovery.

Theorem 2.3. *Assume that $1 \leq r \leq \frac{1}{2}\min(p,q)$ and let $A_1, \dots, A_N \in \mathbb{C}^{p\times q}$. Define $\mathbf{L} : \mathbb{C}^{p\times q} \longrightarrow \mathbb{C}^N$ by $\mathbf{L}(X) = (\mathrm{tr}(A_1^T X), \dots, \mathrm{tr}(A_N^T X))$.*

(1) *If $N < 2r(p+q) - 4r^2$ then $\mathbf{L}$ does not have the $\mathcal{M}_{p\times q,r}(\mathbb{C})$-recovery property.*
(2) *Let $N \geq 2r(p+q) - 4r^2$ and $\{A_j\}_{j=1}^N$ be independently randomly chosen under an absolutely continuous probability distribution in $\mathbb{C}^{q\times p}$. Then with probability one $\mathbf{L}$ does have the $\mathcal{M}_{p\times q,r}(\mathbb{C})$-recovery property.*

Proof. First it is well known that $\dim \mathcal{M}_{p\times q,r}(\mathbb{C}) = r(p+q) - r^2$ ([19, Prop. 12.2]). Note that $\mathbf{L}$ is injective on $\mathcal{M}_{p\times q,r}(\mathbb{C})$ if and only if $\mathbf{L}(X - Y) \neq 0$ for any $X \neq Y$ in $\mathcal{M}_{p\times q,r}(\mathbb{C})$, which is equivalent to $\mathbf{L}(Z) = 0$ for $Z \in \mathcal{M}_{p\times q,2r}(\mathbb{C})$ if and only if $Z = 0$. Now the set

$$W = \Big\{ Z \in \mathcal{M}_{p\times q,2r}(\mathbb{C}) : \ \mathbf{L}(Z) = 0 \Big\}$$

is the intersection of $\mathcal{M}_{p\times q,2r}(\mathbb{C})$ with N subspaces of dimension $pq-1$. By Theorem 2.2 it has dimension at least

$$\dim W \geq \dim \mathcal{M}_{p\times q,r}(\mathbb{C}) - N = 2r(p+q) - 4r^2 - N.$$

For (1) we have $\dim W > 0$, and hence it contains a nonzero element. So $\mathbf{L}$ cannot be injective on $\mathcal{M}_{p\times q,r}(\mathbb{C})$.

For (2) by choosing $\{A_j\}$ independently, with probability one for each k the subspace defined by $\operatorname{tr}(A_k^T Z) = 0$ does not contain an irreducible component of the projective variety

$$W_{k-1} = \Big\{ Z \in \mathcal{M}_{p\times q,2r}(\mathbb{C}) : \ \operatorname{tr}(A_j^T Z) = 0 \text{ for } j = 1, \ldots, k-1 \Big\}.$$

In fact this holds for any given projective variety, not just for W_{k-1}. By Theorem 2.2, with probability one we have $\dim W = 0$, which implies that $W = \{0\}$ (see [19]). Hence $\mathbf{L}$ is injective on $\mathcal{M}_{p\times q,r}(\mathbb{C})$ with probability one. $\square$

Tying the recovery property with the dimension of varieties we easily have

Theorem 2.4. *Let $\mathcal{M}$ be a projective variety in $\mathbb{C}^d$ with $\dim(\mathcal{M}) = K$. Let $\ell_1(\mathbf{x}), \ldots, \ell_N(\mathbf{x})$ be linear functions on $\mathbb{F}^d$ where $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$. Set $\mathbf{L}(\mathbf{x}) = (\ell_1(\mathbf{x}), \ldots, \ell_N(\mathbf{x}))^T$ and*

$$Y := \Big\{ (\mathbf{x}, \mathbf{y}) : \ \mathbf{x}, \mathbf{y} \in \mathcal{M}, \ \mathbf{x} \neq \mathbf{y}, \ \ell_j(\mathbf{x} - \mathbf{y}) = 0 \text{ for } 1 \leq j \leq N \Big\}.$$

(A) *For $\mathbb{F} = \mathbb{C}$, $\mathbf{L}$ has the $\mathcal{M}$-recovery property if and only if $Y = \emptyset$. If the (complex) quasi-projective variety Y has $\dim(Y) < K$ then $\mathbf{L}$ has the almost everywhere $\mathcal{M}$-recovery property.*

(B) *For $\mathbb{F} = \mathbb{R}$ let $\mathcal{M}_{\mathbb{R}}$ and $Y_{\mathbb{R}}$ be the set of real points in $\mathcal{M}$ and Y, respectively. Then $\mathbf{L}$ has the $\mathcal{M}_{\mathbb{R}}$-recovery property if and only if $Y_{\mathbb{R}} = \emptyset$. If $\dim(\mathcal{M}) = \dim_{\mathbb{R}}(\mathcal{M}_{\mathbb{R}}) = K$ and $\dim(Y) < K$ then $\mathbf{L}$ has the almost everywhere $\mathcal{M}_{\mathbb{R}}$-recovery property.*

Proof. For both (A) and (B), the conclusions on $\mathcal{M}$-recovery and $\mathcal{M}_{\mathbb{R}}$-recovery are rather clear. For the almost everywhere recovery in the complex case, let Z denote the set of $\mathbf{x} \in \mathcal{M}$ such that there exists a $\mathbf{y} \neq \mathbf{x}$ in $\mathcal{M}$ such that $\ell_j(\mathbf{y}) = \ell_j(\mathbf{x})$ for all $1 \leq j \leq N$. Observe that the set Z is the projection of Y onto the first coordinate. Since projections cannot increase dimension (see [19, Cor.11.13]), it follows that $\dim Z < K = \dim \mathcal{M}$. Hence Z is a null set in $\mathcal{M}$ (with respect to the Hausdorff measure).

For $\mathbb{F} = \mathbb{R}$, we already stated that the real dimension of $Y_{\mathbb{R}}$ is no larger than the (complex) dimension of Y. Thus $\dim_{\mathbb{R}}(Y_{\mathbb{R}}) < K = \dim_R(\mathcal{M}_{\mathbb{R}})$. The same argument now applies to show that $Z_{\mathbb{R}}$ is a null set in $\mathcal{M}_{\mathbb{R}}$. The theorem is proved. $\square$

In many of the data recovery problems the measurements are restricted to special settings. Often the measurement vectors are on a projective variety themselves. Such are the cases for phase retrieval, matrix recovery, and projection retrieval among the examples we listed. Techniques presented above cannot be straightforwardly extended in these special settings. To extend the techniques broadly we introduce the notion of an admissible algebraic variety with respect to a family of linear functions. This was first done in [11], and it proves to be very useful for the study of data recovery on projective varieties.

Definition 2.1 ([11])**.** Let V be the zero locus of a finite collection of homogeneous polynomials in $\mathbb{C}^d$ with $\dim V > 0$ and let $\{\ell_\alpha(\mathbf{x}) : \alpha \in I\}$ be a family of (homogeneous) linear functions. We say V is *admissible* with respect to $\{\ell_\alpha(\mathbf{x})\}$ if $\dim(V \cap \{\ell_\alpha(\mathbf{x}) = 0\}) < \dim V$ for all $\alpha \in I$.

It is well known in algebraic geometry that if V is irreducible in $\mathbb{C}^d$ then $\dim(V \cap Y) = \dim(V) - 1$ for any hyperplane Y that does not contain V. Thus the above admissible condition is equivalent to the property that no irreducible component of V of dimension $\dim V$ is contained in any hyperplane $\ell_\alpha(\mathbf{x}) = 0$. In general without the irreducibility condition, admissibility is equivalent to that for a generic point $\mathbf{x} \in V$, any small neighborhood U of $\mathbf{x}$ has the property that $U \cap V$ is not completely contained in any hyperplane $\ell_\alpha(\mathbf{x}) = 0$. The following theorem plays a fundamental role in our study. It was proved in [18]. For completeness we also present the original proof here.

Theorem 2.5 ([18])**.** *For $j = 1, \ldots, N$ let $L_j : \mathbb{C}^n \times \mathbb{C}^m \to \mathbb{C}$ be bilinear functions and V_j be projective varieties in $\mathbb{C}^n$. Set $V := V_1 \times \cdots \times V_N \subseteq (\mathbb{C}^n)^N$. Let $W, Y \subset \mathbb{C}^m$ be a projective varieties in $\mathbb{C}^m$ and consider the quasi-projective variety $W \setminus Y$. For each fixed j, assume that V_j is admissible with respect to the linear functions $\{f^{\mathbf{w}}(\cdot) = L_j(\cdot, \mathbf{w}) : \mathbf{w} \in W \setminus Y\}$.*

(A) Assume that $N \geq \dim W$. There exists an algebraic subvariety $Z \subseteq V$ with $\dim(Z) < \dim(V)$ such that for any $\mathbf{x} = (\mathbf{v}_j)_{j=1}^N \in V \setminus Z$, the subvariety $X_{\mathbf{x}}$ given by

$$X_{\mathbf{x}} := \Big\{\mathbf{w} \in W \setminus Y : \; L_j(\mathbf{v}_j, \mathbf{w}) = 0 \text{ for all } 1 \leq j \leq N\Big\}$$

is the empty set.

(B) Assume that $N < \dim W$. There exists an algebraic subvariety $Z \subset V$ with $\dim Z < \dim V$ such that for any $\mathbf{x} = (\mathbf{v}_j)_{j=1}^N \in V \setminus Z$, the subvariety $X_\mathbf{x}$ given by

$$X_\mathbf{x} := \left\{ \mathbf{w} \in W \setminus Y : \ L_j(\mathbf{v}_j, \mathbf{w}) = 0 \text{ for all } 1 \le j \le N \right\}$$

has $\dim X_\mathbf{x} = \dim W - N$.

Proof. We include the original proof from [18] for self-containment. First we prove (A). For $\mathbf{x} = (\mathbf{v}_j)_{j=1}^N \in V$, define $\Phi_\mathbf{x} : W \setminus Y \to \mathbb{C}^N$ by $\Phi_\mathbf{x}(\mathbf{w}) = (L_j(\mathbf{v}_j, \mathbf{w}))_{j=1}^N$. Let $\mathcal{G}$ be the subset of $[V] \times [W \setminus Y] \subset \mathbb{P}((\mathbb{C}^n)^N) \times \mathbb{P}(\mathbb{C}^m)$ such that $([X], [W]) \in \mathcal{G}$ if and only if $\Phi_\mathbf{x}(\mathbf{w}) = 0$, i.e. $L_j(\mathbf{v}_j, \mathbf{w}) = 0$ for all j. Note that $\mathcal{G}$ is a projective variety of $\mathbb{P}((\mathbb{C}^n)^N) \times \mathbb{P}(\mathbb{C}^m)$. We consider its dimension. Let π_1 and π_2 be projections from $\mathbb{P}((\mathbb{C}^n)^N) \times \mathbb{P}(\mathbb{C}^m)$ onto the first and the second coordinates, respectively, namely

$$\pi_1([\mathbf{x}], [\mathbf{w}]) = [\mathbf{v}_1, \ldots, \mathbf{v}_N], \qquad \pi_2([\mathbf{x}], [\mathbf{w}]) = [\mathbf{w}].$$

It is easy to check that $\pi_2(\mathcal{G}) = [W \setminus Y]$, the projection of $W \setminus Y$. Thus $\dim(\pi_2(\mathcal{G})) = \dim(W \setminus Y) - 1$.

We next consider the dimension of the preimage of the $\pi_2^{-1}([\mathbf{w}_0]) \subset \mathbb{P}((\mathbb{C}^n)^N)$ for a fixed $[\mathbf{w}_0] \in \mathbb{P}(\mathbb{C}^m)$. Let $V_j' := V_j \cap H_j$ where $H_j := \{\mathbf{y} \in \mathbb{C}^n : L_j(\mathbf{y}, \mathbf{w}_0) = 0\}$ is a hyperplane. The admissibility property of V_j implies that $\dim(V_j') = \dim(V_j) - 1$. Hence after projectivization the preimage $\pi_2^{-1}([\mathbf{w}_0])$ has dimension

$$\dim \pi_2^{-1}([\mathbf{w}_0]) = \sum_{j=1}^N (\dim(V_j) - 1) - 1 = \dim(V) - N - 1.$$

According to Cor.11.13 in [19], we have

$$\begin{aligned}
\dim(\mathcal{G}) &= \dim(\pi_2(\mathcal{G})) + \dim(\pi_2^{-1}([\mathbf{w}_0])) \\
&= (\dim(W \setminus Y) - 1) + (\dim(V) - N - 1) \\
&= \dim(V) + \dim(W \setminus Y) - N - 2 \\
&\le \dim(V) + \dim(W) - N - 2.
\end{aligned}$$

If $N \ge \dim W$ then

$$\dim(\pi_1(\mathcal{G})) \le \dim(\mathcal{G}) = \dim(V) + \dim(W) - N - 2 \le \dim(V) - 2.$$

Note that $\pi_1(\mathcal{G})$ is itself a projective variety. Let Z be the lift of $\pi_1(\mathcal{G})$ into the vector space $(\mathbb{C}^n)^N$. Then

$$\dim Z \le \dim V - 1.$$

The definition of Z implies that $X_{\mathbf{x}}$ is an empty set provided $\mathbf{x} \in V \setminus Z$.

Next we prove (B). Let $K = \dim(W \setminus Y)$. Noting $K > N$, we augment $\{V_j\}_{j=1}^N$ and $\{L_j(\mathbf{v}, \mathbf{w})\}_{j=1}^N$ to $\{V_j\}_{j=1}^K$ and $\{L_j(\mathbf{v}, \mathbf{w})\}_{j=1}^K$ via $V_j = V_1$ and $L_j(\mathbf{v}, \mathbf{w}) = L_1(\mathbf{v}, \mathbf{w})$ for all $j > N$. Set $\hat{V} = V_1 \times \cdots \times V_K \subseteq (\mathbb{C}^n)^K$. By (A) there exists a subvariety $\hat{Z}$ of $\hat{V}$ with $\dim \hat{Z} < \dim \hat{V}$ such that for any $\hat{\mathbf{x}} = (\mathbf{v}_j)_{j=1}^K \in \hat{V} \setminus \hat{Z}$ and $\mathbf{w} \in W \setminus Y$, we have $L_j(\mathbf{v}, \mathbf{w}) \neq 0$ for all j. Now consider the sequence of nested varieties with $X_{\hat{\mathbf{x}},0} = W$ and

$$X_{\hat{\mathbf{x}},k} := \Big\{\mathbf{w} \in W \setminus Y : \; L_j(\mathbf{v}_j, \mathbf{w}) = 0 \text{ for all } 1 \leq j \leq k\Big\}, \quad k = 1, \ldots, K.$$

Thus the above is equivalent to $X_{\hat{\mathbf{x}},K} = \emptyset$ provided $\hat{\mathbf{x}} \in \hat{V} \setminus \hat{Z}$.

Since for each fixed $\mathbf{v}_j$ the equation $L_j(\mathbf{v}_j, \mathbf{w}) = 0$ defines a hyperplane H in $\mathbb{C}^m$, it is well known that $\dim(U \cap H) \geq \dim(U) - 1$ for any variety U in $\mathbb{C}^m$. Then we have a decreasing sequence of subvarieties of $\mathbb{C}^m$

$$W \setminus Y = X_{\hat{\mathbf{x}},0} \supseteq X_{\hat{\mathbf{x}},1} \supseteq X_{\hat{\mathbf{x}},2} \supseteq \cdots \supseteq X_{\hat{\mathbf{x}},K} = \emptyset.$$

Now $\dim(X_{\hat{\mathbf{x}},0}) = \dim W \setminus Y = K$. By Krull's Principal Ideal Theorem, at each step the dimension can only be reduced by at most 1, we must thus have $\dim(X_{\hat{\mathbf{x}},k-1}) - 1 = \dim(X_{\hat{\mathcal{A}},k})$ for $1 \leq k \leq K$. It follows that $\dim(X_{\hat{\mathbf{x}},N}) = \dim W - N = K - N$.

Thus for any $\mathbf{x} = (\mathbf{v}_j)_{j=1}^N \in V$, if there exists $\mathbf{v}_j \in V_j$ for $N < j \leq K$ such that $\hat{\mathbf{x}} = (\mathbf{v}_j)_{j=1}^K \in \hat{V} \setminus \hat{Z}$ we must have $\dim(X_{\hat{\mathbf{x}},N}) = K - N$. Since $X_{\hat{\mathbf{x}},N} = X_{\mathbf{x}}$ we then have $\dim(X_{\mathbf{x}}) = K - N$. Finally, let $Z = \{\mathbf{x} = (\mathbf{v}_j)_{j=1}^N\} \subset V$ be those such that there exists no such extensions $\hat{\mathbf{x}} \in \hat{V} \setminus \hat{Z}$. We have

$$Z = \Big\{\mathbf{x} = (\mathbf{v}_j)_{j=1}^N \in V : \; \hat{\mathbf{x}} = (\mathbf{v}_j)_{j=1}^K \in \hat{Z} \text{ for any } \mathbf{v}_j \in V_j, j > N\Big\}.$$

Since $\hat{Z}$ is variety in $(\mathbb{C}^n)^K$, Z is a variety. Clearly it has $\dim(Z) < \dim(V)$, for otherwise we would have $\dim(\hat{Z}) = \dim(\hat{V})$, which is a contradiction. $\square$

Remark. While the theorem may look abstract as far as the date recovery problem goes, it actually provides a general framework for many applications. One should observe that with the exception of compressive sensing, in all other examples the measurements are in the form of $\mathrm{tr}(A^T X)$ for some matrix A. One can view $\mathrm{tr}(A^T X)$ as a bilinear function in A and X, so the measurements of X are in fact from a bilinear function like $\mathrm{tr}(A^T X)$ by taking suitable samples of A. In the general setting, any linear function $\ell(\mathbf{x})$ on $\mathbb{F}^d$, where $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$, can be expressed in the form of $\ell(\mathbf{x}) = L(\mathbf{v}_0, \mathbf{x})$ for some bilinear function L and sample point $\mathbf{v}_0$. As we move on, this point will become more and more clear.

3. Data recovery on a projective variety

Let $\mathcal{M}$ be a projective variety in $\mathbb{F}^d$ where $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$. Applying Theorem 2.5 and other results in the previous section we can now prove results for $\mathcal{M}$-recovery and almost everywhere $\mathcal{M}$-recovery. First for any $\mathbf{a} = (a_1, \ldots, a_d)^T \in \mathbb{F}^d$ we define $\phi_{\mathbf{a}} : \mathbb{F}^d \longrightarrow \mathbb{F}$ by

$$\phi_{\mathbf{a}}(\mathbf{x}) := \sum_{j=1}^{d} a_j x_j.$$

Any linear function $L(\mathbf{x})$ on $\mathbb{F}^d$ can be written uniquely as $L(\mathbf{x}) = \phi_{\mathbf{a}}(\mathbf{x})$ for some $\mathbf{a} \in \mathbb{F}^d$. Note that $\phi_{\mathbf{a}}(\mathbf{x})$ is a bilinear function of $\mathbf{a}$ and $\mathbf{x}$.

We now examine the $\mathcal{M}$-recovery property from linear samples. We shall first consider the following setup: each linear measurement of $\mathbf{x} \in \mathcal{M}$ is in the form of $\phi_{\mathbf{a}}(\mathbf{x})$ for some $\mathbf{a} \in \mathbb{F}^d$. This offers complete generality. In different problems, $\mathbf{a}$ may be chosen from various special sets. Here we assume they are sampled from projective varieties. Our main results concern the recovery property when the linear samples are generic. First we consider the complex case. Our next two theorems are slightly more general versions of Theorems 4.3 and 4.4 in [18], and we provide proofs here for self-containment.

Theorem 3.1. *Let $\mathcal{M} \subseteq \mathbb{C}^d$ and V_j be projective varieties in $\mathbb{C}^d$, $j = 1, \ldots, N$. Assume that each V_j satisfies. For $\mathcal{A} = (\mathbf{a}_j)_{j=1}^N$ with $\mathbf{a}_j \in V_j$ denote $\mathbf{L}_{\mathcal{A}} = (\phi_{\mathbf{a}_1}, \phi_{\mathbf{a}_2}, \ldots, \phi_{\mathbf{a}_N})^T$.*

(A) *If $N < \dim(\mathcal{M} - \mathcal{M})$ then $\mathbf{L}_{\mathcal{A}}$ does not have the $\mathcal{M}$-recovery property. On the other hand, if $N \geq \dim(\mathcal{M} - \mathcal{M})$ then for a generic $\mathcal{A} \in V_1 \times \cdots \times V_N$ the linear map $\mathbf{L}_{\mathcal{A}}$ has the $\mathcal{M}$-recovery property.*

(B) *If $N < \dim(\mathcal{M})$ then $\mathbf{L}_{\mathcal{A}}$ does not have the almost everywhere $\mathcal{M}$-recovery property. On the other hand, if $N > \dim(\mathcal{M})$ then for a generic $\mathcal{A} \in V_1 \times \cdots \times V_N$ the linear map $\mathbf{L}_{\mathcal{A}}$ has the almost everywhere $\mathcal{M}$-recovery property.*

Proof. Let $K = \dim(\mathcal{M})$. First we prove (B). If $N < \dim(M)$ then $\mathbf{L}_{\mathcal{A}}$ maps smoothly the higher dimensional manifold $\mathcal{M}$ to the lower dimensional one $\mathbb{C}^N$. If $L_{\mathcal{A}}$ is almost everywhere injective, by looking at $\mathcal{M}$ locally we see that there exists a smooth map Φ from a ball B in $\mathbb{C}^K$ to $\mathbb{C}^N$ that is almost everywhere injective. But this is impossible by Lemma 4.2 of [18].

Now for $N > \dim(\mathcal{M}) = K$ let $X \subset \mathbb{C}^d \times \mathbb{C}^d$ be the quasi-projective variety

$$X := \Big\{(\mathbf{v}, \mathbf{u}) \in \mathcal{M} \times \mathcal{M},\ \mathbf{v} \neq \mathbf{u}\Big\}.$$

For each $(\mathbf{v},\mathbf{u}) \in X$ denote $\psi_{(\mathbf{v},\mathbf{u})}(\mathbf{a}) = \phi_{\mathbf{v}-\mathbf{u}}(\mathbf{a})$. As in Theorem 2.4 set

$$Y_{\mathcal{A}} := \Big\{(\mathbf{v},\mathbf{u}) \in X,\ \phi_{\mathbf{a}_j}(\mathbf{v}-\mathbf{u}) = 0 \text{ for } 1 \le j \le N\Big\}. \tag{3.1}$$

Since each V_j is admissible with respect to the maps $\{\psi_{(\mathbf{v},\mathbf{u})} : (\mathbf{v},\mathbf{u}) \in X\}$. By Theorem 2.5 for a generic $\mathcal{A} = (\mathbf{a}_j) \in V_1 \times V_2 \times \cdots \times V_N$ we have $\dim(Y_{\mathcal{A}}) = \dim(X) - N < 2K - K = \dim(\mathcal{M})$. It follows from Theorem 2.4 that $\mathbf{L}_{\mathcal{A}}$ has the almost everywhere $\mathcal{M}$-recovery property.

To prove (A), for $N < \dim(\mathcal{M} - \mathcal{M})$ the dimension of the projective variety $Y_{\mathcal{A}}$ given in 3.1 is no less than $\dim(\mathcal{M} - \mathcal{M}) - N > 0$ by Theorem 2.2. Thus $Y_{\mathcal{A}}$ is not empty and $\mathbf{L}_{\mathcal{A}}$ does not have the $\mathcal{M}$-recovery property.

In the case $N \ge \dim(\mathcal{M} - \mathcal{M})$, we apply Theorem 2.5 with $W = X = (\mathcal{M} - \mathcal{M}) \setminus \{0\}$ and $L_j(\mathbf{a}_j, \mathbf{x}) = \phi_{\mathbf{a}_j}(\mathbf{x})$ for all j. Let $V = V_1 \times V_2 \times \cdots \times V_N$. Then for a generic $\mathcal{A} = (\mathbf{a}_j) \in V$ we have $Y_{\mathcal{A}} = \emptyset$. Hence $\mathbf{L}_{\mathcal{A}}$ has the $\mathcal{M}$-recovery property. □

For the real case, the above theorem can be extended. Suppose that $V \subseteq \mathbb{R}^d$ is a real variety. We next introduce a natural extension of V to a variety in $\mathbb{C}^d$. The ideal $I_{\mathbb{R}}(V)$ defining V generates an ideal $I_{\mathbb{C}}(V)$ in $\mathbb{C}^d$, and the variety corresponding to $I_{\mathbb{C}}(V)$ will be our extension, and we denote it by $\bar{V}$. A simple observation is that V is clearly the restriction of $\bar{V}$ to $\mathbb{R}^d$, namely $V = \bar{V}_{\mathbb{R}}$ using the terminology in this paper.

Theorem 3.2. *Let $\mathcal{M}$ and V_j be projective varieties in $\mathbb{R}^d$, $j = 1, \dots, N$. Assume that each $\bar{V}_j$ is admissible with respect to the maps $\{\phi_{\mathbf{v}} : \mathbf{v} \in \bar{\mathcal{M}} - \bar{\mathcal{M}}, \mathbf{v} \ne 0\}$. For $\mathcal{A} = (\mathbf{a}_j)_{j=1}^N$ with $\mathbf{a}_j \in V_j$ denote $\mathbf{L}_{\mathcal{A}} = (\phi_{\mathbf{a}_1}, \phi_{\mathbf{a}_2}, \dots, \phi_{\mathbf{a}_N})^T$. Assume further that $\dim_{\mathbb{R}}(\mathcal{M}) = \dim(\bar{\mathcal{M}})$ and $\dim_{\mathbb{R}}(V_j) = \dim(\bar{V}_j)$ for all j.*

- (A) *If $N < \dim_{\mathbb{R}}(\mathcal{M})$ then $\mathbf{L}_{\mathcal{A}}$ does not have the almost everywhere $\mathcal{M}$-recovery property. On the other hand, if $N > \dim_{\mathbb{R}}(\mathcal{M})$ then a generic $\mathcal{A} = (\mathbf{a}_j)$ in $V_1 \times V_2 \times \cdots \times V_N$ has the almost everywhere $\mathcal{M}$-recovery property.*
- (B) *Assume additionally that $\dim_{\mathbb{R}}(\mathcal{M} - \mathcal{M}) = \dim(\bar{\mathcal{M}} - \bar{\mathcal{M}}) = L$. If $N \ge L$ then a generic a generic $\mathcal{A} = (\mathbf{a}_j)_{j=1}^N$ in $V_1 \times V_2 \times \cdots \times V_N$ has the $\mathcal{M}$-recovery property.*

Proof. Let $V = V_1 \times V_2 \times \cdots \times V_N$. For (A), if $N < \dim(M)$ then the map $\mathbf{L}_{\mathcal{A}}$ cannot be almost everywhere injective from the same argument as in the complex case. If $N > \dim(\mathcal{M})$ we consider $\bar{\mathcal{M}}$ and $\bar{V}_j$. Let $Y_{\mathcal{A}}$ be the same as in Theorem 3.1, but in $\mathbb{R}^d$, and let

$$\bar{X} := \Big\{(\mathbf{v},\mathbf{u}) \in \bar{\mathcal{M}} \times \bar{\mathcal{M}},\ \mathbf{v} \neq \mathbf{u}\Big\},$$
$$\bar{Y}_{\mathcal{A}} := \Big\{(\mathbf{v},\mathbf{u}) \in \bar{X},\ \mathbf{L}_{\mathcal{A}}(\mathbf{x}-\mathbf{y}) = 0\Big\}.$$

By the argument from Theorem 3.1, and use Theorem 2.5 there exists a subvariety $\bar{Z} \subset \bar{V}$ with $\dim \bar{Z} < \dim \bar{V}$ such that for any $\mathcal{A} = (\mathbf{a}_j) \in \bar{V} \setminus \bar{Z}$ we have $\dim(\bar{Y}_{\mathcal{A}}) = \dim(X) - N < 2K - K = \dim(\bar{\mathcal{M}})$. By assumption we have $\dim_{\mathbb{R}} V = \dim \bar{V}$ so the restriction $Z = Z_{\mathbb{R}}$ of $\bar{Z}$ to the reals must have $\dim Z < \dim V$. Furthermore, $\dim_{\mathbb{R}}(Y_{\mathcal{A}}) \leq \dim(\bar{Y}_{\mathcal{A}}) < K$. It follows from Theorem 2.4 that any $\mathcal{A} = (\mathbf{a}_j) \in V \setminus Z$ has the almost everywhere $\mathcal{M}$-recovery property. In other words, a generic $\mathcal{A} = (\mathbf{a}_j)$ in V gives $\mathbf{L}_{\mathcal{A}}$ the $\mathcal{M}$-recovery property. This proves (A).

For (B) we follow the same strategy. Let $V, \bar{V}$ be as in part (A). Since $N \geq \dim(\bar{\mathcal{M}} - \bar{\mathcal{M}})$ it follows from Theorem 3.1 and 2.5 that there exists a variety $\bar{Z} \subset \bar{V}$ with $\dim \bar{Z} < \dim \bar{V}$ such that for any $\mathcal{A} = (\mathbf{a}_j) \in \bar{V} \setminus \bar{Z}$ the map $\mathbf{L}_{\mathcal{A}}$ has the $\bar{\mathcal{M}}$-recovery property. Thus $\mathbf{L}_{\mathcal{A}}$ has the $\mathcal{M}$-recovery property for any $\mathcal{A} = (\mathbf{a}_j) \in V \setminus Z$. Since

$$\dim_{\mathbb{R}}(Z) \leq \dim(\bar{Z}) < \dim(\bar{V}) = \dim_{\mathbb{R}}(V),$$

it follows that a generic $\mathcal{A} = (\mathbf{a}_j) \in V$ has the $\mathcal{M}$-recovery property. $\square$

Given that the admissibility condition plays a key role in our theorem, one may ask whether this condition can be checked rather easily. Indeed, the condition is rather easy to check, and for almost all situations that we encounter, the condition holds. We list some examples. Note that many of the applications of interest involve matrices, so we focus on admissibility in $\mathbb{C}^{p\times q}$. The lemma below is proved in [18], Proposition 4.1.

Lemma 3.3. *Let V be one of the following projective varieties in $\mathbb{C}^{p\times q}$. Then V is admissible with respect to any set of nontrivial linear functions on $\mathbb{C}^{p\times q}$:*

(A) $V = \mathcal{M}_{p\times q,s}(\mathbb{C})$, *the set of all $p \times q$ complex matrices of rank s or less, where $1 \leq s \leq \min(p,q)$.*
(B) $q \geq p$ *and V is the set of all scalar multiples of matrices P satisfying $PP^T = I$.*
(C) $p = q$ *and V is the set of all scalar multiples of projection matrices P, i.e. $P^2 = P$.*

We can now apply Theorems 3.1 and 3.2 to various problems, some listed earlier in the introduction, to answer questions concerning the number of measurements needed for data recovery.

Matrix recovery

We have already shown how basic algebraic geometry can be applied to matrix recovery in Theorem 2.3. We can extend it to the more general setting. Recall that $\mathcal{M}_{p\times q,r}(\mathbb{F})$ denotes the set of all matrices in $\mathbb{F}^{p\times q}$ having rank no greater than r, where $\mathbb{F}=\mathbb{C}$ or $\mathbb{R}$.

Theorem 3.4. *Assume that $1\le r\le \frac{1}{2}\min(p,q)$ and let V be a projective variety in $\mathbb{C}^{p\times q}$ that is admissible with respect to all nontrivial linear functions on $\mathbb{C}^{p\times q}$. For $A_1,\dots,A_N\in\mathbb{F}^{p\times q}$, where $\mathbb{F}=\mathbb{C}$ or $\mathbb{R}$, define $\mathbf{L}:\mathbb{F}^{p\times q}\longrightarrow\mathbb{F}^N$ by $\mathbf{L}(X)=(\mathrm{tr}(A_1^TX),\dots,\mathrm{tr}(A_N^TX))$.*

- (A) *If $N<r(p+q)-r^2$ then $\mathbf{L}$ does not have the almost everywhere $\mathcal{M}_{p\times q,r}(\mathbb{F})$-recovery property.*
- (B) *For $\mathbb{F}=\mathbb{C}$, if $N<2r(p+q)-4r^2$ then $\mathbf{L}$ does not have the $\mathcal{M}_{p\times q,r}(\mathbb{C})$-recovery property. This result fails for $\mathbb{F}=\mathbb{R}$.*
- (C) *For $\mathbb{F}=\mathbb{C}$ and generic $A_j\in V$, $\mathbf{L}$ has the $\mathcal{M}_{p\times q,r}(\mathbb{C})$-recovery property if $N\ge 2r(p+q)-4r^2$, and $\mathbf{L}$ has the almost everywhere $\mathcal{M}_{p\times q,r}(\mathbb{C})$-recovery property if $N\ge r(p+q)-r^2$.*
- (D) *For $\mathbb{F}=\mathbb{R}$ and assuming that $\dim V=\dim_{\mathbb{R}} V_{\mathbb{R}}$. For generic $A_j\in V_{\mathbb{R}}$, $\mathbf{L}$ has the $\mathcal{M}_{p\times q,r}(\mathbb{R})$-recovery property if $N\ge 2r(p+q)-4r^2$, and $\mathbf{L}$ has the almost everywhere $\mathcal{M}_{p\times q,r}(\mathbb{R})$-recovery property if $N\ge r(p+q)-r^2$.*

The proof is clearly a straightforward application of Theorems 3.1 and 3.2, and we omit it here.

We are also interested in the cases where the matrix recovery is on subsets of $\mathcal{M}_{p\times q,r}(\mathbb{F})$. For example, the Projection Retrieval problem is a special case of matrix recovery problem, where the measurements are rank one matrices. Similarly we may consider the recovery of a Hermitian matrix X from quadratic measurements $\mathbf{v}_j^*X\mathbf{v}_j$, $j=1,\dots,N$. This, like the Projection Retrieval, is in fact recovery from linear measurements on a manifold, as

$$b_j:=\mathbf{v}_j^*X\mathbf{v}_j=\mathrm{tr}(\mathbf{v}_j^*X\mathbf{v}_j)=\mathrm{tr}(\mathbf{v}_j\mathbf{v}_j^*X).$$

The measurement matrices are rank one Hermitian matrices here. Such problems can be handled similarly. One complication is that although Hermitian matrices are complex, they do not form a complex variety. Thus the theorems we have here on complex recovery cannot be applied directly to the recovery of Hermitian matrices. However, they can be formulated as the affine image of a real projective variety, and from which our theorems can be applied.

Theorem 3.5. *For $\mathcal{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N)$ where $\mathbf{a}_j \in \mathbb{C}^p$ for all j let $\mathbf{L}_\mathcal{A} : \mathbb{C}^{p\times p} \longrightarrow \mathbb{C}^N$ be define by $\mathbf{L}_\mathcal{A}(X) = (\mathbf{a}_1^* X \mathbf{a}_1, \ldots, \mathbf{a}_N^* X \mathbf{a}_N)^T$.*

(A) *Let $\mathcal{M} \subset \mathbb{R}^{p\times p}$ be the set of all real symmetric matrices of rank at most r where $r \le p/2$. For a generic $\mathcal{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_N)$ with $\mathbf{a}_j \in \mathbb{R}^p$, if $N \ge 2pr - 2r^2 + r$ then $\mathbf{L}_\mathcal{A}$ has the $\mathcal{M}$-recovery property. If $N \ge pr - r(r-1)/2 + 1$ then $\mathbf{L}_\mathcal{A}$ has the almost everywhere $\mathcal{M}$-recovery property.*

(B) *Let $\mathcal{M} \subset \mathbb{C}^{p\times p}$ be the set of all Hermitian matrices of rank at most r where $r \le p/2$. For a generic $\mathcal{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_N)$ with $\mathbf{a}_j \in \mathbb{C}^p$, if $N \ge 4pr - 4r^2$ then $\mathbf{L}_\mathcal{A}$ has the $\mathcal{M}$-recovery property. If $N \ge 2pr - r^2 + 1$ then $\mathbf{L}_\mathcal{A}$ has the almost everywhere $\mathcal{M}$-recovery property.*

Proof. Part (A) follows from Theorem 3.2 and the fact that the projective variety of all rank s complex symmetric matrices in $\mathbb{C}^{p\times p}$ has dimension $ps - s(s-1)/2$, which is also the real dimension of $\mathcal{M}$.

Part (B) is a bit more complicated because $\mathcal{M}$ is not a projective variety. However, we use a technique that works also for other problems involving Hermitian matrices. This technique is first used in [11]. Consider the map $\varphi : \mathbb{C}^{p\times p} \to \mathbb{C}^{p\times p}$ defined by

$$\varphi(A) = \frac{1}{2}(A + A^T) + \frac{i}{2}(A - A^T).$$

Then φ is a isomorphism on $\mathbb{C}^{p\times p}$ that maps $\mathbb{R}^{p\times p}$ one-to-one to the set of all Hermitian matrices in $\mathbb{C}^{p\times p}$. Let

$$\bar{\mathcal{N}} = \Big\{ A \in \mathbb{C}^{p\times p} : \operatorname{rank}(\varphi(A)) \le r \Big\}$$
$$\mathcal{N} = \Big\{ A \in \mathbb{R}^{p\times p} : \operatorname{rank}(\varphi(A)) \le r \Big\}.$$

Then $\mathcal{N} = \bar{\mathcal{N}}_\mathbb{R}$. Observe that $\mathcal{M} = \varphi(\mathcal{N})$. Define $\tilde{\mathbf{L}}_\mathcal{A}(X) := \mathbf{L}_\mathcal{A}(\varphi(X))$. We only need to show that $\tilde{\mathbf{L}}_\mathcal{A}$ has the $\mathcal{N}$-recovery property if $N \ge 4pr - 4r^2$, and almost everywhere $\mathcal{N}$-recovery property if $N \ge 2pr - r^2 + 1$. It is known that $\dim_\mathbb{R}(\mathcal{M}) = 2pr - r^2$ and $\dim_\mathbb{R}(\mathcal{M} - \mathcal{M}) = 4pr - 4r^2$. Thus $\dim_\mathbb{R}(\mathcal{N}) = 2pr - r^2$ and $\dim_\mathbb{R}(\mathcal{N} - \mathcal{N}) = 4pr - 4r^2$. Furthermore, $\dim(\bar{\mathcal{N}}) = \dim_\mathbb{R}(\mathcal{N})$ and $\dim(\bar{\mathcal{N}} - \bar{\mathcal{N}}) = \dim_\mathbb{R}(\mathcal{N} - \mathcal{N})$, see [18]. Because the projective variety of rank one or less is admissible with respect to all linear functions on $\mathbb{C}^{p\times p}$, Theorems 3.1 and 3.2 now imply that for generic $\mathcal{A}$, $\tilde{\mathbf{L}}_\mathcal{A}$ has the $\mathcal{N}$-recovery property if $N \ge 4pr - 4r^2$ and it has the almost everywhere $\mathcal{N}$-recovery property if $N \ge 2pr - r^2 + 1$. The theorem follows. $\square$

The above theorem can be applied immediately to phase retrieval to yield the following

Corollary 3.6. *Let $\{\mathbf{f}_j\}_{j=1}^N$ be a generic set of vectors in $\mathbb{F}^d$, where $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$. Then $\{\mathbf{f}_j\}_{j=1}^N$ have the phase retrieval property in $\mathbb{F}^d$ if $N \geq 2d-1$ for $\mathbb{F} = \mathbb{R}$, or if $N \geq 4d-4$ for $\mathbb{F} = \mathbb{C}$.*

The proof is a straightforward conclusion from Theorem 3.5 with $r = 1$. These results are well known. While for $\mathbb{F} = \mathbb{R}$ it is a rather straightforward to prove using basic linear algebra [4], the case $\mathbb{F} = \mathbb{C}$ is in fact quite nontrivial and was first proved recently in [7]. The above results also extend to generalized phase retrieval under the admissibility conditions [11], and in particular for fusion frame phase retrieval [8, 11].

Projection retrieval

Projection Retrieval is a special case of the matrix recovery problem in which the matrix we try to recover is an orthogonal projection. This has been studied in recent years in e.g. [14]. Here we consider a slightly more general setting where we try to recover a matrix Q from the measurements $\|P\mathbf{a}_j\|^2$, $j = 1, \ldots, N$, knowing that $Q = aP$ where $a > 0$ and P is an orthogonal projection matrix, namely $P = P^*$ and $P^2 = P$. In other words, we try to recover a scalar multiple of an orthogonal projection instead of just an orthogonal projection as in the original Projection Retrieval problem. We shall focus on the real case. The complex case is slightly more tedious, but can be handled with the same techniques used to prove the complex case of Theorem 3.5.

Theorem 3.7. *Let $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N$ be generic vectors in $\mathbb{R}^d$. If $N \geq 2r(d-r)+2$ then every $Q = aP$ where $a > 0$ and P is an orthogonal projection matrix of rank $1 \leq r < d$ can be recovered from $\{\|Q\mathbf{a}_j\|\}_{j=1}^N$. If $N \geq r(d-r)+2$ then almost every such $Q = aP$ can be recovered from $\{\|Q\mathbf{a}_j\|\}_{j=1}^N$.*

Proof. We first observe that $\|Q\mathbf{a}_j\|^2 = \mathbf{a}_j^T Q^T Q \mathbf{a}_j = \mathrm{tr}(A_j^T X)$, where $A_j = \mathbf{a}_j \mathbf{a}_j^T$ and $X = a^2 P$. Thus proving the theorem is equivalent to proving that $X = a^2 P$ is uniquely determined by $\mathrm{tr}(A_j^T X)$, $j = 1, \ldots, N$. Denote

$$\mathcal{M} := \{aP \in \mathbb{C}^{d\times d} : \ a \in \mathbb{C}, P^T = P, P^2 = P, \mathrm{rank}(P) = r, 1 \leq r < d\}.$$

Note that $\mathcal{M}_{\mathbb{R}}$ is precisely the set of all scalar multiples of real orthogonal projections of rank r. Furthermore, $\mathcal{M}$ is a projective variety. One can easily see that $\mathcal{M}$ consists of all matrices Q in $\mathbb{C}^{d\times d}$ satisfying

$$Q^T = Q, \quad Q^2 = \frac{1}{r}\mathrm{tr}(Q)Q.$$

We know that $\dim(\mathcal{M}) = r(d-r)+1$, which is evident from counting degree of freedoms. Furthermore, we also have $\dim_{\mathbb{R}}(\mathcal{M}_{\mathbb{R}}) = r(d-r)+1$. As with Theorem 3.5 the required admissibility condition is also met for us to apply Theorem 3.2. It follows that for $N \geq \dim(\mathcal{M}-\mathcal{M})$ we can recover $X = a^2P$ from $\{\|Q\mathbf{a}_j\|^2\}_{j=1}^N$. But $\dim(\mathcal{M}-\mathcal{M}) \leq 2\dim(\mathcal{M}) = 2r(d-r)+2$. Of course we can now recover $Q = aP$ from X (since $a > 0$). The first part of the theorem now follows. The second part follows immediately from the fact that $\dim(\mathcal{M}) = r(d-r)+1$. ☐

Remark. Right now we do not have a precise result for $\dim(\mathcal{M}-\mathcal{M})$, and as a result the first conclusion in the theorem may not be sharp. We can prove, however, that $\dim(\mathcal{M}-\mathcal{M}) = 2r(d-r)+\delta$ with $\delta = 1$ or 2 (we omit the details in this paper). Also, the original Projection Retrieval problem poses an additional challenge that the set of orthogonal projections is not a projective variety. We leave these questions as open problems for interested researchers.

4. Computational aspect of data recovery on a manifold

For some special manifolds, the data recovery problem can be solved successfully by convex programmings using very few linear measurements. For examples, the low-rank matrix recovery and phase retrieval can be done by nuclear norm minimization [22, 23], and compressed sensing uses ℓ_1 norm minimization to reconstruct the sparse signal [24]. From convex geometry and sparse representation points of view, [25] gives a unified convex optimization, called atomic norm minimization. However, the atomic norm minimization is sometimes computationally intractable, as it is NP-hard for many data recovery problems on manifold (e.g. [26]). Moreover, the convex optimization framework does not utilize the structure of the low-dimensional manifold.

Here we provide a more general computational framework, by considering the fact that the data are on a low-dimensional manifold. Since X satisfies $\mathbf{L}(X) = \mathbf{b}$ and $X \in \mathcal{M}$, it is natural to recast the recover of X into the following constrained least squares problem

$$\min_{Z\in\mathbb{F}^d} \frac{1}{2}\|\mathbf{L}(Z)-\mathbf{b}\|_2^2, \quad \text{s.t.} \quad Z \in \mathcal{M}. \tag{4.1}$$

In other words, we minimize the least square error of the linear measurements on the manifold. Obviously, the underlying true data X is a global minimizer of (4.1), and any global minimizer of (4.1) is a solution of the linear equation on $\mathcal{M}$. Therefore, in the case that $\mathbf{L}(X) = \mathbf{b}$ has a unique solution on $\mathcal{M}$, the recovery of X is equivalent to finding the unique global minimizer of (4.1).

To better exploit the structure of the manifold, we employ numerical optimization algorithms on manifold [27, 28] to solve (4.1). For this purpose, we assume the manifold $\mathcal{M}$ is smooth, so that the tangent of $\mathcal{M}$ is well defined. For any $Z \in \mathcal{M}$, denote $\mathcal{T}_Z$ the tangent of $\mathcal{M}$ at Z. We endow the tangent space a Riemannian metric the standard Euclidean metric. We consider the gradient descent algorithm on the Riemannian manifold.

Let $f(Z)$ be the objective function in (4.1), i.e.,

$$f(Z) = \frac{1}{2}\|\mathbf{L}(Z) - \mathbf{b}\|^2$$

Due to the Euclidean embedding, the gradient of f at Z on $\mathcal{M}$ is given by

$$\nabla_{\mathcal{M}} f(Z) = \mathcal{P}_{\mathcal{T}_Z}\mathbf{L}^*(\mathbf{L}(Z) - \mathbf{b}),$$

where $\mathcal{P}_{\mathcal{T}_Z}$ is the projection onto the tangent space $\mathcal{T}_Z$. With this, the Riemannian gradient descent (RGrad) applied to (4.1) is

$$\begin{cases} G_l = \mathcal{P}_{\mathcal{T}_{Z_l}}\mathbf{L}^*(\mathbf{L}(Z_l) - \mathbf{b}), \\ Z_{l+1} = \mathcal{R}_{\mathcal{M}}(Z_l - \alpha_l G_l), \end{cases} \tag{4.2}$$

where $\mathcal{R}_{\mathcal{M}}$ is the retraction onto the manifold $\mathcal{M}$ and α_l is a step size. Since f is quadratic, we may choose α_k the steepest descent step size. More precisely, we may define $\alpha_l = \arg\min_\alpha f(Z_l - \alpha G_l)$, which has a closed form

$$\alpha_l = \frac{\langle \mathbf{L}(Z_l) - \mathbf{b}, \mathbf{L}(G_l)\rangle}{\|\mathbf{L}(G_l)\|^2} = \frac{\|G_l\|_2^2}{\|\mathbf{L}(G_l)\|^2}. \tag{4.3}$$

The RGrad algorithm can be accelerated by conjugate gradient (CG) algorithms on the Riemannian manifold. Instead of the gradient direction, the Riemannian CG algorithm uses a linear combination of the gradient and the previous updating direction, projected onto the tangent space, to update the current iteration. We omit the details here and interested readers may consult [27, 29, 30].

To apply Riemannian optimization algorithms to get practical data recovery algorithms on manifold, there are still several issues unsolved and we need to tune the algorithms. The problem (4.1) is a non-convex optimization. The convergence of a non-convex numerical solver to a global minimum is generally not guaranteed. How to find a good initialization for RGrad to achieve a global minimum? How many linear measurements are sufficient to find the correct solution X on the manifold?

In the rest of this section, we will apply Riemannian optimization algorithms to some example problems of data recovery on manifold. We will discuss how to tune them to get efficient algorithms, and we will also address how many linear measurements are sufficient for the successful recovery of X.

Example 1: matrix recovery

The unknown data $X \in \mathbb{F}^{p\times q}$ lies on $\mathcal{M}_{p\times q,r}$, the manifold of all matrices with rank not larger than r. The linear measurements $\mathbf{L}$ is defined by

$$L_j(X) = \mathrm{tr}(A_j^T X), \quad j = 1, \dots, N,$$

where $A_j \in \mathbb{F}^{p\times q}$ are measurement matrices. However, the manifold $\mathcal{M}_{p\times q,r}$ is not smooth at matrices whose rank is strictly smaller than r. To apply the Riemannian optimization algorithms, instead of the manifold $\mathcal{M}_{p\times q,r}$, we find X on the manifold of matrices with rank exactly r, i.e.,

$$\mathcal{M}^{\mathrm{E}}_{p\times q,r} := \Big\{Q \in \mathbb{F}^{q\times p} : \mathrm{rank}(Q) = r\Big\}.$$

Note that the dimension of $\mathcal{M}_{p\times q,r} \setminus \mathcal{M}^{\mathrm{E}}_{p\times q,r}$ is strictly smaller than that of $\mathcal{M}_{p\times q,r}$. Therefore, $\mathcal{M}_{p\times q,r} \setminus \mathcal{M}^{\mathrm{E}}_{p\times q,r}$ is measure 0 on $\mathcal{M}_{p\times q,r}$ and therefore neglectable.

The rank-r manifold $\mathcal{M}^{\mathrm{E}}_{p\times q,r}$ is smooth and has a very nice structure embedded in $\mathbb{F}^{p\times q}$. Its tangent space at $Z \in \mathcal{M}^{\mathrm{E}}_{p\times q,r}$ is given by

$$\mathcal{T}_Z = \{UA^* + BV^* : A \in \mathbb{F}^{q\times r}, B \in \mathbb{F}^{p\times r}\},$$

where $U \in \mathbb{F}^{p\times r}$ and $V \in \mathbb{F}^{q\times r}$ are left and right singular vector matrices respectively in the compact singular value decomposition (SVD) $Z = U\Sigma V$. It is easy to check that the orthogonal projection onto the tangent space is given by,

$$\mathcal{P}_{\mathcal{T}_Z} Y = UU^*Y + YVV^* - UU^*YVV^*, \quad \forall\, Y \in \mathbb{F}^{p\times q}.$$

We need to find a retraction operator $\mathcal{R}_{\mathcal{M}}$, whose role is to retract a matrix back to the rank-r manifold. There are several choices of such an operator. We choose $\mathcal{H}_r$, the projection onto $\mathcal{M}^{\mathrm{E}}_{p\times q,r}$ or the r-truncated SVD, as the retraction $\mathcal{R}_{\mathcal{M}}$. More precisely, for any $W \in \mathbb{F}^{p\times q}$,

$$\mathcal{H}_r(W) = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^*, \quad \text{where } Y = \sum_{i=1}^{\min\{p,q\}} \sigma_i \mathbf{u}_i \mathbf{v}_i^* \text{ is the SVD.}$$

For the initial guess, we use the standard spectral method. In particular, we assume the measurement matrices A_j, $j = 1, \ldots, N$, have i.i.d. random entries with mean 0 and variance $1/N$, and a straightforward calculation implies

$$\mathrm{Exp}(\mathbf{L}^*(\mathbf{b})) = \mathrm{Exp}\left(\sum_{j=1}^{N} \mathrm{tr}(A_j^T X) A_j\right) = X, \tag{4.4}$$

where $\mathrm{Exp}(\cdot)$ denotes the expectation. Therefore, it is reasonable to choose

$$Z_0 = \mathcal{H}_r(\mathbf{L}^*(\mathbf{b})). \tag{4.5}$$

The purpose of $\mathcal{H}_r$ is to set $\mathrm{rank}(X_0) = r$ while not escaping too far away from $\mathbf{L}^*(\mathbf{b})$, so that Z_0 will be close to the under truth solution X if $\mathbf{L}^*(\mathbf{b})$ has a good concentration around its expectation.

The full RGrad algorithm for low-rank matrix recovery is shown in Algorithm 1. To the efficient implementation of Algorithm 1, the structure of the tangent space $\mathcal{T}_{Z_l}$ can be further exploited. In particular, by using the fact that the matrices in $\mathcal{T}_{Z_l}$ have rank at most $2r$, the SVD of size $p \times q$ in the evaluation of $\mathcal{H}_r$ can be reduced to two QR decompositions of size $p \times r$ and $q \times r$ respectively and one SVD of size $2r \times 2r$, which significantly save

Algorithm 1 Riemannian Gradient Descent (RGrad) for Low-Rank Matrix Recovery

1: Initialize $Z_0 = \mathcal{H}_r(\mathbf{L}^*(\mathbf{b}))$.
2: **for** $l = 0, 1, \ldots$. **do**
3: $\quad G_l = \mathcal{P}_{\mathcal{T}_{Z_l}} \mathbf{L}^*(\mathbf{L}(Z_l) - \mathbf{b})$.
4: $\quad$ Choose α_l.
5: $\quad Z_{l+1} = \mathcal{H}_r(Z_l - \alpha_l G_l)$.
6: **end for**

the computational cost per step. We omit the details, and interested readers are referred to [29, 30].

The following theorem is an immediate corollary of [29, Theorem 2.1], and it shows that Algorithm 1 converges linearly to the true solution X with dominant probability, if $N \geq O((p+q)r\log(\kappa\sqrt{r}))$, where κ is the condition number of X. Compared to the bound in Theorem 3.4, the minimum measurement for RGrad to work is the same order as the least measurement for $\mathcal{M}$-recovery up to a logarithmic factor.

Theorem 4.1 (A corollary of [29, Theorem 2.1])**.** *Consider the real case. Assume the entries of A_j, $j = 1, \ldots, N$, are i.i.d. Gaussian with mean* 0 *and variance* $1/N$*. For any $\rho \in (0,1)$, there exist positive universal constants c_0, c_1, c_2, such that: for any $X \in \mathbb{R}^{p\times q}$ with rank r and condition number κ, the sequence Z_l generated by Algorithm 1 with α_l as in (4.3) satisfies*

$$\|Z_l - X\|_F \leq \rho^l \|Z_0 - X\|_F$$

with probability at least $1 - c_1 e^{-c_2 N}$*, provided*

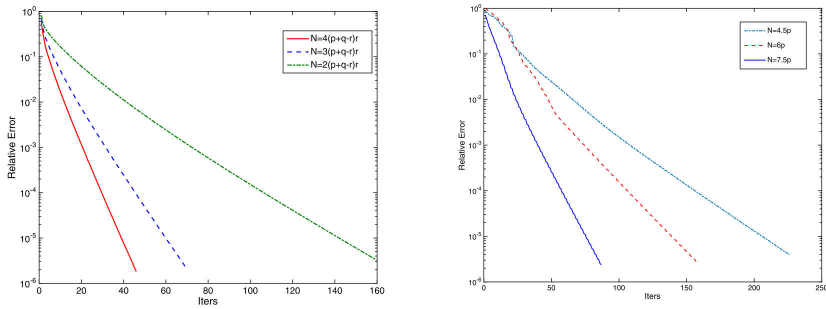
$$N \geq c_0(p+q)r\log(\kappa\sqrt{r}).$$

A numerical experiment is performed. We choose N to be 2, 3, and 4 times of the dimension of the manifold respectively. The convergence curves of Algorithm 1 with stepsize (4.3) is shown in Figure 1(a). We see that, the more measurements, the faster convergence of the algorithm.

Example 2: phase retrieval

By introducing $X = \mathbf{x}\mathbf{x}^*$, the recovery of $\mathbf{x} \in \mathbb{F}^p$ from phaseless measurements $|\langle \mathbf{x}, \mathbf{f}_j\rangle|^2 = b_j$ for $j = 1, \ldots, N$ as in (1.1) can be reformulated as a problem of finding a real symmetric or a Hermitian rank-1 solution of $\mathbf{L}(X) = \mathbf{b}$ with $L_j(X) = \mathrm{tr}(\mathbf{f}_j\mathbf{f}_j^* X)$ for $j = 1, \ldots, N$. By a simple calculation, if we start with a symmetric/Hermitian initial guess, the Riemannian gradient descent on the symmetric/Hermitian rank-1 manifold is exactly the same as the one on the standard rank-1 manifold. Therefore, the RGrad algorithm for low-rank matrix recovery can be applied equally to the phase retrieval problem, where $p = q$ and $r = 1$.

However, all the measurement matrices in $\mathbf{L}$ are of rank 1, and the direct application of RGrad algorithm may need a large number of measurements theoretically. To overcome this, we shall slightly modify the RGrad algorithm. For simplicity, we discuss only the complex case. We assume that



(a) Low-rank matrix recovery (b) Phase retrieval

Figure 1: Convergence of the RGrad algorithm.

$\mathbf{f}_j$, $j = 1, \ldots, N$, follow a complex Gaussian model, i.e., the real and complex parts of $\mathbf{f}_j$ are all real random Gaussian vectors with expectation 0 and variance $I/2$. With this random model, the measurement matrices $\mathbf{f}_j\mathbf{f}_j^*$, $j = 1, \ldots, N$, are outer products of Gaussian random vectors, which have a heavier tail distribution. This makes the phase retrieval problem more difficult to solve than the standard low-rank matrix recovery problems. To eliminate the effects caused by the heavy tail, our idea is to drop out those ill-posed measurements in the initialization and each iteration.

For the initialization, we cannot choose the initial guess (4.5), because (4.4) does not hold true due to the rank-1 measurement matrices. We use the initialization presented in [31]. The expectation of $b_j = |\langle \mathbf{x}, \mathbf{f}_j \rangle|^2$ is $\|\mathbf{x}\|_2^2$, which can be approximated well by $\|\mathbf{b}\|_1/N$. Therefore, we use only those measurements in the set

$$\Omega_0 = \left\{ j : \sqrt{b_j} \leq \beta_0 \sqrt{\|\mathbf{b}\|_1/N} \right\} \tag{4.6}$$

for a predefined constant $\beta_0 > 0$. In other words, we use only those measurements that do not deviate too much from their expectations. Define $Y = \sum_{j \in \Omega} b_j \mathbf{f}_j \mathbf{f}_j^*$, and let $\mathbf{u}$ be its leading unit eigenvector. Following [31], the leading eigenvector of the expectation of Y is parallel to $\mathbf{x}$. This, together with $\|\mathbf{x}\|_2^2 \approx \|\mathbf{b}\|_1/N$, gives us the initialization

$Z_0 = \mathbf{z}_0 \mathbf{z}_0^*$,

where $\mathbf{z}_0 = \sqrt{\|\mathbf{b}\|_1/N} \cdot \mathbf{u}$ and $\mathbf{u}$ is the unit leading eigenvector of Y.

By induction, the positive defniteness of Z_l is preserved. At each iteration l, given $Z_l = \mathbf{z}_l \mathbf{z}_l^*$, we use again only those well-posed measurements adapted to $\mathbf{z}_l$ and $\mathbf{x}$. We set

$$\Omega_l = \Omega_{l1} \cap \Omega_{l2} \cap \Omega_0, \tag{4.7}$$

where

$$\Omega_{l1} = \{j : |\langle \mathbf{z}_l, \mathbf{f}_j \rangle| \leq \beta_1 \|\mathbf{z}_l\|\}$$

and

$$\Omega_{l2} = \left\{ j : \left| b_j - |\langle \mathbf{z}_l, \mathbf{f}_j \rangle|^2 \right| \leq \frac{\beta_2}{N} \|\mathbf{b} - \mathbf{L}(Z_l)\|_1 \frac{|\langle \mathbf{z}_l, \mathbf{f}_j \rangle| + \sqrt{b_j}}{\|\mathbf{z}_l\|} \right\}.$$

The set Ω_{l1} is to enforce $|\langle \mathbf{z}_l, \mathbf{f}_j \rangle|^2$ not too far away from its expectation, and Ω_{l2} is to remove the tail of $|\langle \mathbf{z}_l - \mathbf{x}, \mathbf{f}_j \rangle|^2$. The new iterate is produced by using only those measurements on Ω_l. The complete algorithm is shown in Algorithm 2.

Algorithm 2 Riemannian Gradient Descent (RGrad) for Phase Retrieval

1: Define Ω_0 by (4.6)
2: Initialize $\mathbf{z}_0 = \sqrt{\|\mathbf{b}\|_1/N} \cdot \mathbf{u}$ and $Z_0 = \mathbf{z}_0\mathbf{z}_0^*$, where $\mathbf{u}$ is the leading unit eigenvector of $Y = \sum_{j \in \Omega_0} b_j \mathbf{f}_j \mathbf{f}_j^*$.
3: **for** $l = 0, 1, \ldots$. **do**
4: Z_l is given in the form of $Z_l = \mathbf{z}_l \mathbf{z}_l^*$.
5: Define Ω_l by (4.7).
6: $G_l = \mathcal{P}_{\mathcal{T}_{Z_l}} \left(\sum_{j \in \Omega_j} (|\langle \mathbf{z}_l, \mathbf{f}_j \rangle|^2 - b_j) \mathbf{f}_j \mathbf{f}_j^* \right)$.
7: Choose α_l.
8: $Z_{l+1} = \mathcal{H}_1(Z_l - \alpha_l G_l)$.
9: **end for**

The following theorem is the main theorem in the forthcoming paper [32]. It shows that Algorithm 2 converges linearly to the true solution as long as N is larger than $O(p)$. This bound is the same order as the the minimum number of measurements required in Theorem 3.5 with $r = 1$ for $\mathcal{M}$-recovery, hence it is in the optimal order.

Theorem 4.2 ([32])**.** *Assume the entries of $\mathbf{f}_j$, $j = 1, \ldots, N$, are i.i.d. complex Gaussian with mean 0 and variance 1. For any $\mathbf{x} \in \mathbb{C}^p$, there exist positive constants β_0, β_1, β_2, c_0, c_1, and c_2 such that: if $N \geq c_0 p$, then Z_l generated by the Algorithm 2 step size $\alpha_l = \frac{1}{2N}$ satisfies*

$$\|Z_l - \mathbf{x}\mathbf{x}^*\|_F \leq \left(\frac{1}{2}\right)^l \|Z_0 - \mathbf{x}\mathbf{x}^*\|_F$$

with probability at least $1 - c_1 e^{-c_2 N}$.

In Figure 1(b), we demonstrate the convergence of Algorithm 2. We choose $N = 4.5p$, $N = 6p$, and $N = 7.5p$ respectively. Again we see that increasing the number of measurements accelerate the convergence of the algorithm.

References

[1] E. J. Candès, J. Romberg, and T. Tao, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure Appl. Math., 59(8)(2006) 1207–1223. MR2230846

[2] E. J. Candès, J. Romberg, T. Tao, *Robust uncertainty principles: Exact signal reconstruction from highly incomplte frequency information*, IEEE Trans. Inform. Theory, 52(2)(2006)489–509. MR2236170

[3] D. L. Donoho, *Compressed sensing*, IEEE Transactions on information Theory, 52(4)(2006), 1289–1306. MR2241189

[4] Radu Balan, Pete Casazza, and Dan Edidin. On signal reconstruction without phase. *Applied and Computational Harmonic Analysis*, 20(3):345–356, 2006. MR2224902

[5] Bernhard G. Bodmann and Nathaniel Hammen. Stable phase retrieval with low-redundancy frames. *Advances in computational mathematics*, 41(2):317–331, 2015. MR3337494

[6] E. J. Candes, T. Strohmer, and V. Voroninski. Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics*, 66(8):1241–1274, 2013. MR3069958

[7] Aldo Conca, Dan Edidin, Milena Hering, and Cynthia Vinzant. An algebraic characterization of injectivity in phase retrieval. *Applied and Computational Harmonic Analysis*, 38(2):346–356, 2015. MR3303679

[8] Dan Edidin. Projections and phase retrieval. *Applied and Computational Harmonic Analysis*, 2015. MR3591243

[9] Matthew Fickus, Dustin G. Mixon, Aaron A. Nelson, and Yang Wang. Phase retrieval from very few measurements. *Linear Algebra and its Applications*, 449:475–499, 2014. MR3191879

[10] Teiko Heinosaari, Luca Mazzarella, and Michael M. Wolf. Quantum tomography under prior information. *Communications in Mathematical Physics*, 318(2):355–374, 2013. MR3020161

[11] Yang Wang and Zhiqiang Xu. Generalized phase retrieval: measurement number, matrix recovery and beyond. *Applied and Computational Harmonic Analysis*, 2017.

[12] Jameson Cahill, Peter G. Casazza, Jesse Peterson, and Lindsey Woodland. Phase retrieval by projections. *arXiv preprint arXiv:1305.6226*, 2013. MR3529968

[13] G. E. Pfander, Sampling of operators, J. Fourier Anal. Appl., 19:612–650, 2013. MR3048592

[14] Fickus M., Mixon D. G. Projection retrieval: Theory and algorithms, Sampling Theory and Applications (SampTA), 2015 International Conference on. IEEE, 2015: 183–186.

[15] Zhao, Bo, et al. “Low rank matrix recovery for real-time cardiac MRI.” Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on. IEEE, 2010.

[16] Chen, Chih-Fan, Chia-Po Wei, and Yu-Chiang Frank Wang. “Low-rank matrix recovery with structural incoherence for robust face recognition.” Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012. MR3234185

[17] Zhang, H., He, W., Zhang, L., Shen, H., & Yuan, Q. Hyperspectral image restoration using low-rank matrix recovery. IEEE Transactions on Geoscience and Remote Sensing, 52(8):4729–4743, 2014.

[18] Yi Rong Yang Wang and Zhiqiang Xu. Almost Everywhere Matrix Recovery. *arXiv preprint arXiv:1707.09112*, 2017.

[19] Joe Harris. *Algebraic geometry: a first course*, volume 133. Springer Science & Business Media, 2013. MR1182558

[20] Jacek Bochnak, Michel Coste, and Marie-Françoise Roy. *Real algebraic geometry*, volume 36. Springer Science & Business Media, 2013. MR1659509

[21] Zhiqiang Xu, The minimal measurement number for low-rank matrix recovery, Appl. Comput. Harmon. Anal. (2017), http://dx.doi.org/10.1016/j.acha.2017.01.005

[22] Emmanuel J. Candes and Benjamin Recht. Exact matrix completion via convex optimization. *Foundations of Computational mathematics*, 9(6):717–772, 2009. MR2565240

[23] Jian-Feng Cai, Emmanuel J. Candes, and Zuowei Shen. A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, 20(4):1956–1982, 2010. MR2600248

[24] Emmanuel J. Candes and Michael B. Wakin. An introduction to compressive sampling. *IEEE signal processing magazine*, 25(2):21–30, 2008. MR2275736

[25] Venkat Chandrasekaran, Benjamin Recht, Pablo A. Parrilo, and Alan S. Willsky. The convex geometry of linear inverse problems. *Foundations of Computational mathematics*, 12(6):805–849, 2012. MR2989474

[26] Haixia Liu, Jian-Feng Cai, and Yang Wang. Subspace clustering by (k, k)-sparse matrix factorization. *Inverse Problems and Imaging*, 11(3):539–551, 2017. MR3642256

[27] P.-A. Absil, Robert Mahony, and Rodolphe Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009. MR2364186

[28] Constantin Udriste. *Convex functions and optimization methods on Riemannian manifolds*, volume 297. Springer Science & Business Media, 1994. MR1326607

[29] Ke Wei, Jian-Feng Cai, Tony F. Chan, and Shingyu Leung. Guarantees of Riemannian optimization for low rank matrix recovery. *SIAM Journal on Matrix Analysis and Applications*, 37(3):1198–1222, 2016.

[30] Ke Wei, Jian-Feng Cai, Tony F. Chan, and Shingyu Leung. Guarantees of Riemannian optimization for low rank matrix completion. *arXiv preprint arXiv:1603.06610*, 2016. MR3543156

[31] Yuxin Chen and Emmanuel Candes. Solving random quadratic systems of equations is nearly as easy as solving linear systems. *Advances in Neural Information Processing Systems (NIPS)*, pp. 739–747, 2015.

[32] Jian-Feng Cai and Ke Wei. Phase retrieval via Riemannian optimization: theory and algorithms. *in preparation*, 2017. MR3332993

[33] Saeid Bahmanpour, Jameson Cahill, Peter G. Casazza, John Jasper, and Lindsey M Woodland. Phase retrieval and norm retrieval. *arXiv preprint arXiv:1409.8266*, 2014. MR3441730

[34] E. J. Candes, Y. Eldar, T. Strohmer, and V. Voroninski. Phase retrieval via matrix completion. *SIAM Journal on Imaging Sciences*, 6(1):199–225, 2013. MR3032952

[35] Matthew Fickus and Dustin G. Mixon. Projection Retrieval: Theory and Algorithms. *Sampling Theory and Applications (SampTA), 2015 International Conference on. IEEE*, 2015: 183–186.

[36] S. T. Flammia, A. Silberfarb, C. M. Caves, Minimal informationally complete measurements for pure states. *Foundations of Physics*, 35:1985–2006, 2005. MR2192944

[37] J. Finkelstein, Pure-state informationally complete and 'really' complete measurements. *Physical Review A*, 70(5), 2004. MR2109108

[38] Michael Kech and Michael Wolf. Quantum tomography of semialgebraic sets with constrained measurements. *arXiv preprint arXiv:1507.00903*, 2015. MR3579060

[39] Dani Kogan, Yonina C. Eldar and Dan Oron. On The 2D Phase Retrieval Problem. *arXiv preprint arXiv:1605.08487*, 2016. MR3583013

[40] Dustin Mixon. Phase transitions in phase retrieval. *Excursions in Harmonic Analysis*, Volume 4. Springer International Publishing, 123–147, 2015. MR3411094

[41] Cynthia Vinzant. A small frame and a certificate of its injectivity. *arXiv preprint arXiv:1502.04656*, 2015.

Jian-Feng Cai
Department of Mathematics
Hong Kong University of Science and Technology
Clear Water Bay, Kowloon
Hong Kong
E-mail address: jfcai@ust.hk

Yi Rong
Department of Mathematics
Hong Kong University of Science and Technology
Clear Water Bay, Kowloon
Hong Kong
E-mail address: yrong@ust.hk

Yang Wang
Department of Mathematics
Hong Kong University of Science and Technology
Clear Water Bay, Kowloon
Hong Kong
E-mail address: yangwang@ust.hk

Zhiqiang Xu
LSEC, Inst. Comp. Math.
Academy of Mathematics and System Science
Chinese Academy of Sciences
Beijing, 100190
China
School of Mathematical Sciences
University of Chinese Academy of Sciences
Beijing 100049
China
E-mail address: xuzq@lsec.cc.ac.cn

Received September 15, 2017