Algebraic compressed sensing

Paul Breiding[1]

University Osnabrück, Fachbereich Mathematik/Informatik Albrechtstr. 28a, 49076 Osnabrück, Germany

Fulvio Gesmundo

Saarland University, Saarland Informatics Campus, 66123 Saarbrücken, Germany

Mateusz Michałek[2]

University of Konstanz, Dept. of Mathematics and Statistics, Universitätsstrasse 10, 78457 Konstanz, Germany

Nick Vannieuwenhoven[3]

KU Leuven, Department of Computer Science, Celestijnenlaan 200A, B-3001 Leuven, Belgium;
Leuven.AI, KU Leuven Institute for AI, B-3000 Leuven, Belgium

Abstract

We introduce the broad subclass of algebraic compressed sensing problems, where structured signals are modeled either explicitly or implicitly via polynomials. This includes, for instance, low-rank matrix and tensor recovery. We employ powerful techniques from algebraic geometry to study well-posedness of sufficiently general compressed sensing problems, including existence, local recoverability, global uniqueness, and local smoothness. Our main results are summarized in thirteen questions and answers in algebraic compressed sensing. Most of our answers concerning the minimum number of required measurements for existence, recoverability, and uniqueness of algebraic compressed sensing problems are optimal and depend only on the dimension of the model.

Keywords: Algebraic compressed sensing, recoverability, identifiability

1. Introduction

Compressed sensing investigates inverse problems in which a structured signal in a Euclidean space is to be reconstructed from a smaller number of (typically linear) measurements. Such inverse problems are well posed in the sense of Hadamard if there exists a unique, continuous solution.

The general setting consists of a *model* $\mathcal{X} \subset \mathbb{R}^n$ and a *measurement map*

$$\mu : \mathcal{X} \to \mathbb{R}^s. \tag{IP}$$

Given a *measurement* $y \in \mu(\mathcal{X})$, one is interested in reconstructing a *signal* $x \in \mathcal{X}$ such that $y = \mu(x)$. This setting is schematically depicted in Figure 1.
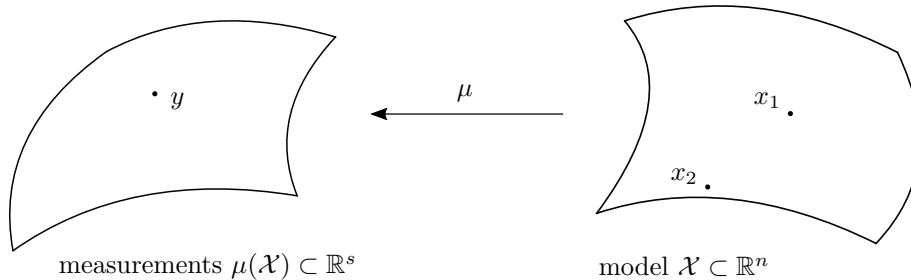
Figure 1: Schematic description of the geometry. On the right is the model $\mathcal{X}$. On the left is the set of measurements $\mu(\mathcal{X})$ taken from $\mathcal{X}$ using the measurement map μ. In the picture we have $\mu(x_1) = y$. We are concerned with the question whether x_1 is the only point in $\mathcal{X}$ mapping to y, or whether there exists another point $x_2 \in \mathcal{X}$ with $\mu(x_2) = y$.

We mainly focus on linear measurements, but we also consider the case where μ is nonlinear. Recall that μ is a *linear measurement map* if

$$\mu = \hat{\mu}|_{\mathcal{X}}, \text{ where } \hat{\mu} : \mathbb{R}^n \to \mathbb{R}^s \text{ is a linear map.}$$

We say that s is the *number of measurements*. The *forward map* μ that takes a signal $x \in \mathcal{X}$ to the observed measurement $y = \mu(x)$ is usually easy to compute (and sometimes is even implemented in hardware [1]). On the other hand, the inverse problem that consists of computing $x \in \mu^{-1}(y)$ on input $y \in \mu(\mathcal{X})$ is generally a more difficult problem, and in fact it is not always possible to solve.

According to Kirsch [2], an inverse problem is *well posed* in the sense of Hadamard at $y \in \mathbb{R}^s$ if the following conditions simultaneously hold:

(E) *Existence:* there exists a solution, i.e., there exists $x \in \mathcal{X}$, with $y = \mu(x)$;

(I) *Identifiability:* the solution is locally unique at y, that is, there is an open neighborhood U of y in $\mu(\mathcal{X})$ such that, for every $y' \in U$, there is a unique $x' \in \mathcal{X}$ with $y' = \mu(x')$;

(C) *Continuity:* the solution $x' = \mu^{-1}(y')$ is continuous as a function of $y' \in U$.

It is argued in [2, Chapter 1] that well-posedness is a practical necessity for numerically solving the inverse problem associated with (IP).

Special cases of the inverse problem (IP) were extensively studied in compressed sensing [3–7]. In the original formulation, $\mathcal{X} \subset \mathbb{R}^n$ is a (special) union of linear subspaces and μ is the restriction to $\mathcal{X}$ of a linear map $\hat{\mu} : \mathbb{R}^n \to \mathbb{R}^s, x \mapsto Ax$ [1, 8]. Later, more general models for $\mathcal{X}$ were investigated, such as smooth manifolds [8, 9]. See Examples 2 to 4 below for further examples of compressed sensing.

Many techniques were developed in compressed sensing to prove well-posedness of a problem, such as mutual coherence, the restricted isometry property (RIP), the nullspace property, or bi-Lipschitz continuity [10]. Arguably, the most popular of these is the RIP [3–5]. However, as we discuss in Section 2, it seems that RIP-based analyses require a number of measurements $s \geq O(d \log n)$ to guarantee well-posedness, where d is the intrinsic dimension of the model $\mathcal{X}$ and n is the dimension of the ambient space. One main contribution of this paper is reducing the number of required samples for (*almost everywhere*) well-posedness to the minimal value for an important class of models $\mathcal{X}$ and measurement maps μ that includes the original union of subspaces and certain smooth manifolds.

This paper studies the well-posedness of *algebraic compressed sensing*; that is, inverse problems where the model $\mathcal{X}$ is explicitly or implicitly defined by polynomials and μ is a polynomial map. Our main tools rely on fundamental techniques from classic algebraic geometry.

We treat the following two complementary settings simultaneously in this work:

2

(IS) *Implicit setting*: $\mathcal{X}$ is defined as the *real vanishing locus* of a real polynomial system of equations. More precisely,

$$\mathcal{X} = \{x \in \mathbb{R}^n : f_1(x) = \cdots = f_k(x) = 0\},$$

where $f_1, \ldots, f_k \in \mathbb{R}[x_1, \ldots, x_n]$ are polynomials with real coefficients in n real variables.

(ES) *Explicit setting*: $\mathcal{X}$ is defined as the *image of a polynomial map*, so that

$$\mathcal{X} = \{x = \phi(t) = (\phi_1(t), \ldots, \phi_n(t)) : t \in \mathbb{R}^\ell\},$$

where $\phi_1, \ldots, \phi_n \in \mathbb{R}[t_1, \ldots, t_\ell]$ are polynomials with real coefficients in ℓ real variables.

Most of our results hold for complex models as well, so for completeness we also cover this case. For convenience and clarity of expression, we assume that $\mathcal{X}$ is *irreducible*: this technical condition is explained in Section 4. Intuitively, it means that $\mathcal{X}$ cannot be expressed as a finite union of such sets. As the *affine space* $\mathbb{R}^\ell$ is irreducible and images of irreducible sets are irreducible, irreducibility always holds in the setting (ES). With some carefulness about the interpretation of genericity, most of our results can be extended to reducible varieties in setting (IS) as well.

In setting (IS) we say that $\mathcal{X}$ is *implicitly defined* by polynomials, and in setting (ES) we say that $\mathcal{X}$ is *explicitly parametrized* by polynomials. In the language of algebraic geometry, a set implicitly defined by polynomials is called a *real algebraic variety*, while sets explicitly parameterized by polynomials are *semi-algebraic sets*. Though note that not all semi-algebraic sets arise in this way.

These two classes of polynomial models cover a wide range of specific cases, many of which were already studied individually in a compressed sensing context. The following examples demonstrate that our assumptions on $\mathcal{X}$ are not very restrictive and the results of this work can be applied in a number of different settings. For an overview of polynomial models in the sciences see also [11].

Example 1 (Classic compressed sensing)**.** The original formulation of compressed sensing assumed that an observed signal $y \in \mathbb{R}^s$ is a k-sparse linear combination of a set of m vectors $v_1, \ldots, v_m$ in $\mathbb{R}^n$. That is, there exist indices i_j, $j = 1, \ldots, k$, so that $y = \hat{\mu}(\lambda_1 v_{i_1} + \lambda_2 v_{i_2} + \cdots + \lambda_k v_{i_k})$, where $\hat{\mu} : \mathbb{R}^n \to \mathbb{R}^s$ is a linear map. This is an algebraic compressed sensing problem where $\mathcal{X} \subset \mathbb{R}^n$ is the union of the $\binom{m}{k}$ linear subspaces of dimension k spanned by the subsets of k vectors from $v_1, \ldots, v_m$. Such a union of subspaces can be formulated as set of solutions of a polynomial system in setting (IS). The sensing map is $\mu = \hat{\mu}|_{\mathcal{X}}$.

Note that $\mathcal{X}$ is not irreducible. Nevertheless, the well-posedness of such compressed sensing problems is well known [10] and it can be verified that for a union of equidimensional varieties, the main result, Theorem 1.1, applies verbatim. ♦

Example 2 (Matrix completion)**.** The algebraic variety of bounded-rank matrices in $\mathbb{R}^{m \times n}$ is a model for compressed sensing that was studied in [12]. When the measurement map is a coordinate projection, the associated inverse problem is also called a *matrix completion problem*, which has many applications, including data analysis and recommender systems [13]. Polynomials that implicitly define $m \times n$ matrices of rank at most r are all the $(r+1) \times (r+1)$ minors. One can also view low-rank matrices in the explicit setting, via the polynomial parameterization given by $\phi : \mathbb{R}^{m \times r} \times \mathbb{R}^{n \times r} \to \mathbb{R}^{m \times n}, (A, B) \mapsto AB^T$. Related low-rank matrix models that our analysis covers include low-rank matrices with sparseness constraints [14] and low-rank plus sparse matrix recovery [15, 16]. The case of measurement maps that are coordinate projections is addressed in [17–22] using a variety of techniques from rigidity theory, algebraic combinatorics, and matroid theory. ♦

Example 3 (Tensor completion)**.** Other examples are various bounded-rank tensor decompositions [23, 24], such as the tensor rank decomposition or canonical polyadic decomposition [25], Tucker decomposition [26], tensor trains [27, 28], hierarchical Tucker [29, 30], and even general tensor network decompositions [31]. All of these examples can be parameterized straightforwardly by polynomials. These decompositions feature in data analysis applications as an a priori model for filling in missing data [23, 32]. The compressed sensing literature already investigated well-posedness conditions for the tensor rank decomposition with sparsity constraints [33], unconstrained tensor rank decomposition [34], and Tucker decomposition, tensor trains, and hierarchical Tucker decompositions were studied jointly in [35]. ♦

Example 4 (Moment varieties). Our final example is from (algebraic) statistics [36]. Multivariate probability distributions are characterized by an infinite sequence of symmetric tensors of increasing degree, namely their *sequence of moments*, or, alternatively, cumulants [37]. Many probability distributions belong to a parameterized family; for example, k-variate Gaussian distributions are parameterized by the mean vector $m \in \mathbb{R}^k$ and the covariance matrix $\Sigma \in \mathbb{R}^{k \times k}$. Pearson [38] observed that for some families a finite sequence of moments suffices to identify the parameters of the distribution from empirical moments. In recent years, it was realized that the set of possible dth order moments of a parameterized family is sometimes defined by polynomials [39–41], thus carving out a *moment variety* $\mathcal{X}$ in the space of dth order symmetric tensors in k variables $S^d \mathbb{R}^k \simeq \mathbb{R}^{\binom{k+d-1}{d}}$. Estimating *all* $\binom{k+d-1}{d}$ entries of the dth order empirical moment in $S^d \mathbb{R}^k$ is computationally challenging. However, since the probability distribution is represented by a point on $\mathcal{X} \subset S^d \mathbb{R}^k$, we could instead compute a linear projection of the empirical moment. This leads to an algebraic compressed sensing problem. Only recently, the well-posedness of this problem was studied by Lindberg, Amendola, and Rodriguez [42]. For mixtures of Gaussians, Guo, Nie and Yang [43] recently presented an efficient numerical algorithm to estimate the parameters from coordinate projections of the third-order moment. ◆

Let

$$d := \dim \mathcal{X}$$

denote the dimension of the model $\mathcal{X}$. The notion of dimension is introduced formally in Section 4.2. This is the formal way to measure the "intrinsic size" of the model $\mathcal{X}$; it coincides with the dimension of the set of non-singular points of $\mathcal{X}$, which is a smooth manifold embedded in $\mathbb{R}^n$ [44]. In the explicit setting (ES), $\dim \mathcal{X}$ is the maximal rank of the Jacobian of the parametrization ϕ at points in $\mathbb{R}^\ell$ [45]; in particular, in this setting, $\dim \mathcal{X}$ is bounded from above by ℓ, which is often called the number of *degrees of freedom* in the parameterization. The latter can be strictly larger than the dimension; the explicit parameterization in Example 2, for example, has $(m+n)r$ degrees of freedom, while the dimension of the variety of bounded-rank matrices is only $(m+n-r)r$ [46, Proposition 12.2]. As we will see, the dimension of the model $\mathcal{X}$ determines the required number of measurements s so that the inverse problem is well posed. Throughout this paper, the dimension d is fixed and the number of measurements s can vary, reflecting the design choice of the measurement map.

The heart of this paper is Section 3. We discuss and explore **thirteen questions and answers** about the well-posedness of algebraic compressed sensing problems. The questions are divided into four themes:

(E) *Existence:* Given $y \in \mathbb{R}^s$ when are there signals $x \in \mathcal{X}$ with $y = \mu(x)$?

(R) *Recoverability*: When are there finitely many points in $\mu^{-1}(y)$ for $y \in \mu(\mathcal{X})$?

(I) *Identifiability*: When is $\mu^{-1}(\mu(x))$ equal to $\{x\}$?

(C) *Continuity:* When is μ^{-1} locally (Lipschitz) continuous at $\mu(x)$?

If (E), (I) and (C) hold at y, the inverse problem is well posed at y by the definition we gave above. The main reason to consider recoverability as well is that it enables localized well-posed algebraic compressed sensing problems. Indeed, if only (E), (R) and (C) hold at y, then the inverse problem can be made well posed by further restricting the variety $\mathcal{X}$ to a neighborhood of one of the elements in the fiber of μ.

The engine that powers our answers in the Q&A section is a fundamental result from commutative algebra called the *Noether Normalization Lemma*; see Theorem 4.12 below. The proof of this result is sketched in Section 4 and we refer to [47] for a comprehensive discussion on the proof and the lemma's deep implications.

Many of our answers use the notion of a "generic point in $\mathcal{X}$". This is a subtle concept, which we carefully explain in Section 4. For the moment, a reader with an affinity to probability may think of a generic point as a point being chosen with probability 1 from a probability measure that is absolutely continuous with respect to the Lebesgue measure; a reader with a background in analysis can interpret the set of generic points as an open dense subset in $\mathcal{X}$. Nevertheless, being a generic point is a strictly stronger property.

4

As a teaser of the questions and answers, we summarize in the next theorem the answers to Questions 1, 2, 7 and 13. In Section 3, we list nine additional questions and answers, which shed light on different aspects of algebraic compressed sensing.

Theorem 1.1. *Let $\mathcal{X}$ be given as in (IS) or (ES), and let μ be a generic linear map $\mu : \mathcal{X} \to \mathbb{R}^s$. Then:*

(E) *the fiber $\mu^{-1}(y) \neq \emptyset$ at $y \in \mathbb{R}^s$ if and only if a particular system of polynomial equations has a solution:*

- *in the implicit setting (IS), this system of equations in the variables $x = (x_1, \ldots, x_n)$ is defined by $f_1(x) = \cdots = f_k(x) = \mu(x) - y = 0$;*
- *in the explicit setting (ES), this system in variables $t = (t_1, \ldots, t_\ell)$ is defined by $\mu(\phi(t)) - y = 0$;*

(R) *if $s \geq d$, the fiber $\mu^{-1}(\mu(x))$ of a generic point $x \in \mathcal{X}$ is finite;*

(I) *if $s \geq d + 1$, the fiber $\mu^{-1}(\mu(x))$ of a generic point $x \in \mathcal{X}$ is equal to $\{x\}$;*

(C) *if $s \geq d + 1$, the fiber $\mu^{-1}(\mu(x))$ of a generic point $x \in \mathcal{X}$ is C^∞ continuous.*

The paper is structured as follows. In the next section, we offer a glimpse of what lies beyond the restricted isometry property for establishing (local) well-posedness of inverse problems like (IP). Thereafter, in Section 3, the main results of this paper are presented as a sequence of thirteen basic questions and answers related to the well-posedness of algebraic compressed sensing problems. Section 4 recalls the main techniques from algebraic geometry that are employed in Section 5 to furnish proofs of the main results from Section 3. Numerical experiments in Section 6 illustrate the minimal requirements on the number of measurements for well-posedness in algebraic compressed sensing. The paper is concluded in Section 7 with an outlook and open problems in algebraic compressed sensing.

2. What lies beyond the restricted isometry property?

The ϵ-*restricted isometry property* (ϵ-RIP) of the measurement map $\mu : \mathcal{X} \to \mathbb{k}^s$ [3–5], where $\mathbb{k} = \mathbb{R}$ or $\mathbb{k} = \mathbb{C}$, is a core concept in compressed sensing. For general models $\mathcal{X} \subset \mathbb{k}^n$, it can be defined as follows [9]: for all $x, y \in \mathcal{X}$ we have

$$0 < 1 - \epsilon \leq \frac{\|\mu(x) - \mu(y)\|_{\mathbb{k}^s}}{\|x - y\|_{\mathbb{k}^n}} \leq 1 + \epsilon < \infty, \tag{RIP}$$

where $\epsilon \in [0, 1)$ is a constant depending on the measurement map μ, and $\|\cdot\|_A$ denotes a norm on the vector space A. The ϵ-RIP implies that the measurement map μ is everywhere identifiable. Maps satisfying the RIP are a special type of *global bi-Lipschitz embeddings*. Such embeddings were studied earlier for general metric spaces in [48–51]. Recall that a map $\mu : \mathcal{X} \to \mathcal{Y}$ between metric spaces $(\mathcal{X}, \mathrm{dist}_\mathcal{X})$ and $(\mathcal{Y}, \mathrm{dist}_\mathcal{Y})$ is *bi-Lipschitz* if there exist constants c, C such that

$$0 < c \leq \frac{\mathrm{dist}_\mathcal{Y}(\mu(x), \mu(y))}{\mathrm{dist}_\mathcal{X}(x, y)} \leq C < \infty$$

for all $x \neq y$ [52]. Geometrically this means that μ is an injective map into $\mathbb{k}^s$ that does not stretch or bend $\mathcal{X}$ too much. For this reason, bi-Lipschitz embeddings are sometimes called *stable embeddings* [8, 9]. A bi-Lipschitz map has a bi-Lipschitz inverse map μ^{-1} with lower and upper Lipschitz constants equal to respectively C^{-1} and c^{-1}. Bi-Lipschitz embeddings are fundamental in compressed sensing because they offer two complementary *global* properties:

(I) the map is everywhere injective, and

(II) the inverse map is Lipschitz continuous.

The first property ensures that for all measurements $y \in \mu(\mathcal{X})$, the inverse problem has a unique solution. The second property guarantees that the inverse problem has a *bounded sensitivity*, i.e., the inverse map has a bounded variation from one point to another, relative to the distance between the points. Indeed, both the ϵ-RIP and the $(1 - \epsilon, C)$-bi-Lipschitz condition offer the *uniform* upper bound

$$\text{dist}_{\mathcal{X}} \left(\mu^{-1}(y_1), \mu^{-1}(y_2) \right) \leq \frac{\text{dist}_{\mathcal{Y}}(y_1, y_2)}{1 - \epsilon} \tag{Lip}$$

for all measurements $y_1, y_2 \in \mu(\mathcal{X})$. Taken together, (I) and (II) thus guarantee that the inverse problem (IP) is well posed.

While properties (I) and (II) are certainly desirable, demanding that the measurement map is *globally* bi-Lipschitz can be challenging to verify theoretically and imposes severe geometric constraints on $\mathcal{X}$ and μ. Moreover, computing the optimal RIP constant for classic compressed sensing of k-sparse vectors is an NP-hard problem [53]. Our viewpoint is that these constraints can sometimes be relaxed. We believe that the more flexible, localized, easily implementable analysis as the one presented in this work can better cater to practical situations.

2.1. How many measurements do we need?

As the intrinsic dimension of the inputs is only $d = \dim \mathcal{X}$, we could a priori hope that $s = d$ measurements would suffice to recover x from $\mu(x)$. However, ϵ-RIP-based analyses of stable recovery of linear maps $\mu : \mathcal{X} \to \Bbbk^s$ in the literature ultimately lead to a required number of samples s that scales at least like $d \cdot \log \left(\frac{n}{\epsilon^2} \right)$ when $\mathcal{X} \subset \Bbbk^n$ is a collection of d-dimensional linear spaces; see, e.g., [1, 8]. For compact d-dimensional manifolds a similar requirement is needed [8, Theorem 3.1].

The ratio between the intrinsic dimension d and the required number of measurements s for recoverability is called the *oversampling rate*. So this rate seems to scale logarithmically in $n\epsilon^{-2}$ in the case where $\mathcal{X}$ is a union of linear spaces. When considering more general models $\mathcal{X}$, insisting on a bi-Lipschitz embedding could lead to even higher oversampling rates. For general Lipschitz manifolds $\mathcal{X}$, the best upper bound we are aware of on the minimal number of measurements s to recover any model $\mathcal{X}$ with (only) a *local*[4] bi-Lipschitz embedding μ is $s \geq d(d + 1)$ [50]. Such a quadratic growth is often too high in applications.

As bi-Lipschitz embeddings preserve much more information about the model $\mathcal{X}$ than mere injective maps, it is not surprising that RIP or bi-Lipschitz conditions lead to an oversampling factor greater than 1. But is this really *necessary* for recoverability? By relaxing the requirement of bi-Lipschitz embeddability, recoverability and identifiability can often be guaranteed with a smaller number of measurements.

This observation is neither new nor controversial. In 1936, Whitney [54] proved that a smooth d-dimensional manifold $\mathcal{X}$ can be smoothly embedded (in particular, the map is injective and its inverse is smooth) into a Euclidean space of dimension $s = 2d$ for some smooth μ. The oversampling rate is thus 2. With such a measurement map, every point is identifiable and μ^{-1} is a local Lipschitz map, but μ^{-1} is not necessarily a (global) Lipschitz map. Birbrair, Fernandes, and Jelonek [55] recently proved that if $\mathcal{X}$ is a closed semi-algebraic set then there exist bi-Lipschitz linear maps μ into $\mathbb{R}^{2d+1}$. Hence, even with bi-Lipschitz embeddings one can get to an oversampling factor as low as $2 + o(1)$ for some models $\mathcal{X}$. Yet, we show that for a generic linear map d measurements suffice for generic recoverability and $d + 1$ measurements suffice for generic identifiabilty; see Theorem 1.1(R) and (I).

2.2. How much does the reconstruction move?

Besides identifiability, the other main use of the RIP is deriving upper bounds on the reconstruction error of specific algorithms [10, Chapter 6]. In most of such analyses, a reconstruction algorithm is called *stable* or *robust* if, in some norm, the reconstruction error $\|x - x^\sharp\|$ is small. Here, $x^\sharp$ is the algorithm's output on input $y = \mu(x) + e$ and e is a small error. For statements about the main algorithms in traditional compressed sensing see also [10, Theorems 6.12, 6.21, 6.25, 6.28].

---

[4]A map f is a local embedding if every point in the domain has a neighborhood such that f restricted to that neighborhood is an embedding.

The concept of *stability* has many meanings in numerical analysis. *Robustness* of a reconstruction algorithm in compressed sensing is different from but closely related to the *forward stability* of an algorithm in numerical linear algebra. Forward stability of an algorithm $\hat{f}$ for a function f is defined in terms of *condition numbers* [56]. Following [57], the condition number $\kappa[f](x)$ of a map $f : \mathcal{X} \to \mathcal{Y}$ between metric spaces $(\mathcal{X}, \mathrm{dist}_{\mathcal{X}})$ and $(\mathcal{Y}, \mathrm{dist}_{\mathcal{Y}})$ at a point $x \in \mathcal{X}$ is

$$\kappa[f](x) = \lim_{\epsilon \to 0} \sup_{\substack{y \in \mathcal{X}, \\ \mathrm{dist}_{\mathcal{X}}(x,y) \leq \epsilon}} \frac{\mathrm{dist}_{\mathcal{Y}}(f(x), f(y))}{\mathrm{dist}_{\mathcal{X}}(x, y)}. \tag{CN}$$

The condition number is an intrinsic property of the function f. It measures how much an infinitesimal perturbation of the input is amplified in the output of the function f. It is a standard way to measure the *sensitivity* to perturbations of a computational problem modeled by the function f [56–58]. If f is a differentiable function, the condition number provides an error bound of the form

$$\mathrm{dist}_{\mathcal{Y}}(f(x), f(y)) \leq \kappa[f](x) \cdot \mathrm{dist}_{\mathcal{X}}(x, y) + o(\mathrm{dist}_{\mathcal{X}}(x, y))$$

and this bound is asymptotically sharp as $\mathrm{dist}_{\mathcal{X}}(x, y) \to 0$. Aside from the foregoing error bound, the condition number is also viewed as a measure of complexity or hardness of computing the output of f [58, 59].

In our setting, the role of f is played by the measurement map μ. An algorithm $\hat{\mu}^{-1}$ for the function μ^{-1} is called *forward stable* [56, 60] if its reconstruction error is "not much larger" than the *backward error* (the difference between $\mu(x)$ and $\mu(x^\sharp)$) multiplied by the condition number. More precisely, a compressed sensing reconstruction algorithm is forward stable if

$$\underbrace{\|x - x^\sharp\|_{\Bbbk^n}}_{\text{output error}} \leq \underbrace{K}_{\text{algorithm-specific}} \cdot \underbrace{\kappa[\mu^{-1}](\mu(x^\sharp))}_{\text{problem-specific}} \cdot \underbrace{\|\mu(x) - \mu(x^\sharp)\|_{\Bbbk^s}}_{\text{backward error}}, \tag{FS}$$

where $1 \leq K < \infty$ is a constant, $\kappa[\mu^{-1}](\mu(x^\sharp))$ is the condition number relative to the distances $\mathrm{dist}_{\mathcal{X}}(x, x') = \|x - x'\|_{\Bbbk^n}$ and $\mathrm{dist}_{\Bbbk^s}(y, y') = \|y - y'\|_{\Bbbk^s}$, and the backward error is $\|\mu(x) - \mu(x^\sharp)\|$. As indicated above, the condition number is a mathematical property of μ^{-1}. The constant K quantifies how much the algorithm $\hat{\mu}^{-1}$ amplifies the inherent sensitivity of the problem μ^{-1}. Note that the algorithm-specific part of the error must be constant, while the problem-specific part depends on the problem instance that is solved. An algorithm is forward stable if K is a constant that is "small" in the eye of the beholder.

Inspecting [10, Theorems 6.12, 6.21, 6.25, 6.28], for example, it is not difficult to prove that all these bounds imply that the corresponding algorithms are forward stable, with an undetermined constant K in (FS). The reason is as follows.

From the vantage point of numerical analysis, the RIP and bi-Lipschitz embeddings guarantee a very strong property: every reconstruction problem instance is *well conditioned*. Informally, an instance of a computational problem locally modeled by a function is called *well conditioned* if the condition number is small, e.g., $\kappa \approx 1$, and *ill conditioned* if the condition number is large, e.g., the inverse of typical backward errors. If μ is a bi-Lipschitz map, then one has a global upper bound on the condition numbers of both μ and μ^{-1}. We include the precise statement of this result for completeness; see [61, Proposition 6.3.10].

Proposition 2.1. *Let $\mu : \mathcal{X} \to \mathcal{Y}$ be a (c, C)-bi-Lipschitz map, then we have*

$$c \leq \kappa[\mu](x) \leq C \quad and \quad \tfrac{1}{C} \leq \kappa[\mu^{-1}](\mu(x)) \leq \tfrac{1}{c}$$

for all x, where the condition number κ is relative to the metrics of $(\mathcal{X}, \mathrm{dist}_{\mathcal{X}})$ and $(\mathcal{Y}, \mathrm{dist}_{\mathcal{Y}})$ in (CN).

A finite condition number κ also implies the existence of the local Lipschitz constant $\kappa + \delta$ for arbitrary $\delta > 0$ of f in a neighborhood of x.

Instead of relying on the RIP or bi-Lipschitz embeddings to guarantee *every* reconstruction problem instance is well conditioned, we propose to rely on the condition number of specific problem instances to

quantify how difficult the reconstruction problem is. For forward, backward, and mixed stable numerical algorithms [60], the bound (FS) then provides a good upper bound on the reconstruction error of a specific problem instance, even if the true solution x is not known.

In the answer to Question 13 below, we show that $\kappa[\mu^{-1}](\mu(x))$ can be computed from the singular values of the derivative of μ. Therefore, it can be approximated numerically for differentiable measurement maps using standard numerical linear algebra algorithms. If an explicit expression of the differential of μ is unknown, one could even apply numerical differentiation techniques to approximate it. Hence, computing the condition number is often significantly easier than determining global lower and upper Lipschitz constants of μ (or μ^{-1}).

By relying on the condition number to furnish reconstruction error bounds (for forward stable algorithms) and moving away from the RIP or bi-Lipschitz embeddings, we sacrifice the *global* upper bound on the condition number of μ^{-1}. That is, we allow that some reconstruction problems are very sensitive to tiny perturbations of the input; hence, they are numerically hard or even impossible to solve [58]. Because of (FS), this also means that these instances should not be expected to be solved with a small forward error [62]. However, this does not mean that most, typical, or average problem instances are difficult to solve accurately!

2.3. Summary

The ϵ-RIP may be difficult or impossible to establish for complicated models $\mathcal{X}$. We therefore believe that trading a global bound on the reconstruction error and global identifiabilty for an algorithmically computable local upper bound and corresponding local identifiabilty (recoverability) can often be a worthwhile tradeoff. This viewpoint is developed further in the rest of the paper.

3. Questions & Answers

We answer several individual questions about recoverability, identifiability, and continuity. Combining these in the appropriate ways establishes the well-posedness of an algebraic compressed sensing problem. The proofs for the statements in this section are given in Section 5. The background in algebraic geometry used in the proofs is introduced in Section 4. The questions and answers, however, can mostly be understood without this background.

Throughout this section, let $\mathcal{X} \subseteq \Bbbk^n$, where $\Bbbk = \mathbb{R}$ or $\Bbbk = \mathbb{C}$, be specified either implicitly as in (IS) or explicitly as in (ES). The definition of these models in the complex case is obtained by replacing $\mathbb{R}$ by $\mathbb{C}$ in both settings. Our answers are general: they do not rely on specific properties of a particular model.

The following discussion centers around the numbers $d, s > 0$. The first is the dimension of $\mathcal{X}$ (real dimension if $\Bbbk = \mathbb{R}$, and complex dimension if $\Bbbk = \mathbb{C}$). This number d is fixed, as the model $\mathcal{X}$ specifying the signal structure is assumed to be fixed beforehand. The number s is the number of measurements and can be varied through the design of the measurement map $\mu : \mathcal{X} \to \Bbbk^s$ in (IP). Of particular interest is the case where μ is the restriction to $\mathcal{X}$ of a linear or an affine linear map $\hat{\mu} : \Bbbk^n \to \Bbbk^s$. An affine linear map has the form $\hat{\mu}(x) = Ax + b$, where $A \in \Bbbk^{s \times n}$ and $b \in \Bbbk^s$. Since shifting by a vector is an isomorphism of $\Bbbk^s$ (as an affine variety), geometric properties of the map $\hat{\mu}$ are directly related to the analogous properties of the *linear map $x \mapsto Ax$*. For this reason, we restrict the analysis to linear maps. Some of the questions below concern the case where μ is a polynomial map.

We first discuss existence. This consists of deciding whether a given $y \in \Bbbk^s$ sits in the image of the model $\mathcal{X}$ under the measurement map μ. Next, we consider recoverability, which consists of determining under which conditions the preimage $\mu^{-1}(\mu(x))$ is finite for a point $x \in \mathcal{X}$. Then, we discuss identifiability, which consists of determining when the preimage is a single point. Finally, we investigate the local continuity of the elements in the fiber of μ.

Many of the statements below are formulated for "generic" data points. This concept is explained formally in Section 4.1 below. In particular, in all the questions below, generic points form an open dense subset of all points.

3.1. Existence

The first question concerns the existence of solutions.

Question 1. *Under what conditions do there exist solutions $y = \mu(x)$ for a given point $y \in \Bbbk^s$?* $\quad\quad$?

Answer 1. First, suppose that $\mathcal{X}$ is the zero locus of the polynomials $f_1, \ldots, f_k$ as in (IS). Then for a fixed measurement $y = (y_1, \ldots, y_s)$ the equality $y = \mu(x)$ holds for some $x \in \mathcal{X}$ if and only if the following system of $k + s$ polynomial equations in the variables $x_1, \ldots, x_n$ has a solution:

$$f_1(x) = \cdots = f_k(x) = y_1 - \mu_1(x) = \cdots = y_s - \mu_s(x) = 0.$$

Now suppose that $\mathcal{X}$ is given as image of a polynomial parameterization as in (ES). Then for a fixed y, the equality $y = \mu(x)$ holds for some $x \in X$ if and only if the following system of s polynomial equations in the variables $t_1, \ldots, t_\ell$ has a solution:

$$y_1 - \mu_1(\phi_1(t), \ldots, \phi_n(t)) = \cdots = y_s - \mu_s(\phi_1(t), \ldots, \phi_n(t)) = 0.$$

The existence of solutions of a polynomial system $g_1 = \cdots = g_N = 0$ as above is characterized as follows.

For $\Bbbk = \mathbb{C}$, the system has solutions if and only if the ideal $\langle g_i : i = 1, \ldots, N \rangle$ does not contain 1. Explicitly, this condition can be tested using Gröbner basis methods [45, 63–66], which are implemented in most computer algebra systems.

For $\Bbbk = \mathbb{R}$, the system has a solution if and only if the *real radical*, in the sense of [67, Section 4.1], does not contain 1. $\quad\quad$ ✓

We note that neither for $\Bbbk = \mathbb{R}$ nor for $\Bbbk = \mathbb{C}$ it is enough to check if some polynomials vanish only at y to guarantee the existence of $x \in \mathcal{X}$ with $y = \mu(x)$. The reason is that a projection of a variety is not necessarily a variety. However, a consequence of Chevalley's Theorem (see Theorem 4.4) is that, in the case $\Bbbk = \mathbb{C}$, there exists a finite number of polynomials $\{h_i(y) : i = 1, \ldots, M\}$, such that the vanishing or non-vanishing of $h_i(y)$ for a fixed y determines whether there is $x \in \mathcal{X}$ with $y = \mu(x)$. Algorithms providing these h_i were implemented in [68, 69]. For $\Bbbk = \mathbb{R}$ one needs to check a finite number of polynomial equations and polynomial *inequalities*; an implementation is provided in the Maple package `RegularChains`.

While the above answer is necessary and sufficient, it is substantially less practical than all of the Q&As that will follow. In fact, deciding whether a solution to a system of polynomial equations exists is an NP-hard decision problem, see, e.g., [70]. Nevertheless, one can resort to numerical methods such as certified homotopy continuation [71, 72], which can find approximate numerical roots. In practice, uncertified homotopy continuation can be a faster alternative with similar accuracy in many cases of practical interest [73–78]. Lairez [79, 80] proved that finding one approximate root of (certain) random polynomial systems has a quasi-linear time complexity in the Blum–Shub–Smale model of computation [81].

3.2. Recoverability

We start with a remarkable, fundamental result about recoverability of generic linear measurement maps. We call a map $\mu : \mathcal{X} \to \Bbbk^s$ *recoverable at a point* $x \in \mathcal{X}$ if the fiber $\mu^{-1}(\mu(x))$ is finite. A map is *everywhere recoverable* if it is recoverable at every point $x \in \mathcal{X}$. A map is *generically recoverable* if it is recoverable for generic points $x \in \mathcal{X}$.

Question 2. *Under what conditions is a generic affine linear map $\mu : \mathcal{X} \to \Bbbk^s$ everywhere recoverable?* ?

Answer 2. A generic affine linear map is everywhere recoverable if and only if the number of measurements is $s \geq d$.

In fact, if $\Bbbk = \mathbb{C}$ and $s = d$, the number of solutions of (IP) at a generic problem instance $y \in \mu(\mathcal{X})$ equals a constant δ, called the *degree of* $\mathcal{X}$, see Section 4.3. In the real case (i.e., $\Bbbk = \mathbb{R}$), the number of real solutions for a generic problem instance is bounded above by that same δ. $\quad\quad$ ✓

The previous Q&A can be extended to certain polynomial maps.

Question 3. *Under what conditions is a measurement map $\mu : \mathcal{X} \to \Bbbk^s$ defined by s generic (non-homogeneous) polynomials of degree p everywhere recoverable?* $\quad\quad$?

Answer 3. A map defined by s generic polynomials of degree equal to p is everywhere recoverable if and only if $s \geq d$.

For $\Bbbk = \mathbb{C}$ and $s = d$, the number of solutions of (IP) at a generic problem instance $y \in \mu(\mathcal{X})$ equals $\delta \cdot p^d$ where δ is the degree of $\mathcal{X}$. If $\Bbbk = \mathbb{R}$, the number of real solutions for a generic problem instance is bounded from above by $\delta \cdot p^d$.   ✓

The answer to Question 3 is implied by applying the *Veronese embedding* to the setting of Question 2. We believe the answer to Question 3 to hold in the more general case where μ is a generic polynomial map with components of possibly different degrees. We expect that a proof should follow by adapting the proof of Noether's Normalization Lemma (Theorem 4.12) to this general setting. We leave this for future work.

Example 5 (Random maps). The space of linear maps and the space of polynomial maps with a fixed degree are finite dimensional vector spaces. One can randomly sample such maps from a probability distribution that is absolutely continuous with respect to the Lebesgue measure. The answers to Questions 2 and 3 guarantee that such randomly sampled maps are recoverable with probability 1. This includes all linear and polynomial maps whose coefficients are sampled identically and independently distributed from any continuous probability distribution, such as `randn` and `rand` in Julia, Matlab, Octave, and Python.

There is a vast body of literature on robust recovery from random measurements. For instance, the case of low-rank matrix recovery from Gaussian measurements is discussed in [82]: the authors show that $10r(m + n)$ Gaussian measurements suffice to robustly recover rank-r matrices in $\mathbb{R}^{m \times n}$ via nuclear norm minimization. However, the dimension of the variety of rank-r matrices is $d = r(m + n - r)$, so the answer to Question 2 guarantees that $s = d$ measurements are enough for recoverability. This provides a significant improvement over [82], albeit with only local guarantees on the robustness of the reconstruction.   ♦

In *data completion* or *imputation*, one tries to recover a structured signal $x \in \mathcal{X}$ from partially observed data, i.e., a subset of the coordinates of $\Bbbk^n$. In this case, the measurement map is a *coordinate projection*. These are specific linear maps for which the foregoing Q&A's do not apply a priori. Even randomly selecting the s coordinates on which to project does not guarantee a generic linear measurement map in the sense of Questions 2 and 3. Nevertheless, for coordinate projections we have the following result.

Question 4. *Under what conditions does there exist a coordinate projection $\hat{\mu} : \Bbbk^n \to \Bbbk^s$ that is generically recoverable when restricted to $\mathcal{X}$?*   ?

Answer 4. There exists a generically recoverable coordinate projection if and only if $s \geq d$.   ✓

The following example shows that *everywhere* recoverability is not possible in general for coordinate projections. More precisely, the only everywhere recoverable coordinate projection is in general the identity map $\hat{\mu} : \Bbbk^n \to \Bbbk^n$.

Example 6. We provide an example of a variety $\mathcal{X} \subseteq \Bbbk^n$ with the property that every coordinate projection $\mu : \mathcal{X} \to \Bbbk^s$, with $s < n$, has at least one infinite fiber. Consider the parameterization $\phi : \Bbbk^2 \to \Bbbk^n$ defined by $\phi(t_1, t_2) = (\phi_1(t_1, t_2), \ldots, \phi_n(t_1, t_2))$, where $\phi_j(t_1, t_2) = t_2 \cdot \prod_{i \neq j}(t_1 - i)$. It is easy to show that the Jacobian matrix of ϕ has rank 2 for every (t_1, t_2) with $t_2 \neq 0$. Let $\mathcal{X} = \overline{\text{im}(\phi)}$, where the overline denotes the closure in the Euclidean topology. This is an algebraic variety of dimension $d = 2$. For $j = 1, \ldots, n$, we have

$$\phi(j, t_2) = t_2(0, \ldots, 0, \underbrace{\prod_{i \neq j}(j - i)}_{j\text{-th entry}}, 0, \ldots, 0).$$

In particular, the lines corresponding to the coordinate axes of $\Bbbk^n$ are contained in $\mathcal{X}$. Now, every coordinate projection maps at least one coordinate axis to 0. Therefore, if $\mu : \mathcal{X} \to \Bbbk^s$ is a coordinate projection with $s < n$, the fiber $\mu^{-1}(y)$ for $y = 0 \in \Bbbk^s$ is always infinite, containing at least one of these lines.   ♦

We stress that the answer to Question 4 only guarantees existence of a generically recoverable coordinate projection. We do not claim that most or even several coordinate projections with $s \geq d$ are generically recoverable. In fact, constructions similar to the one of Example 6 provide examples of algebraic varieties with the property that only some specific coordinate projections are generically recoverable. An example of this behavior is presented next.

Example 7 (Rank-1 matrix recovery). Let $\mathcal{X}$ be the variety of 2×3 matrices of rank at most 1, that is $\mathcal{X} = \phi(\Bbbk^5)$, with

$$\phi : \Bbbk^5 \to \Bbbk^{2 \times 3}, \quad (t_1, \ldots, t_5) \mapsto \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \begin{bmatrix} t_3 & t_4 & t_5 \end{bmatrix}. \tag{1}$$

One can verify that $d = \dim \mathcal{X} = 4$. Let μ be the measurement map defined by restricting the coordinate projection

$$\hat{\mu} : \Bbbk^{2 \times 3} \to \Bbbk^4, \quad \begin{bmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \end{bmatrix} \mapsto (x_1, x_2, x_4, x_5)$$

to $\mathcal{X}$. There is no point $y \in \Bbbk^4$ at which this coordinate projection is recoverable. Indeed, if $A = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \begin{bmatrix} t_3 & t_4 & t_5 \end{bmatrix}$, then $\mu^{-1}(\mu(A)) \supseteq \left\{ \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \begin{bmatrix} t_3 & t_4 & t' \end{bmatrix} : t' \in \Bbbk \right\}$; hence, every fiber has dimension (at least) one.

On the other hand, every coordinate projection that forms a "cross" as discussed in [17], like

$$\tilde{\mu} : \Bbbk^{2 \times 3} \to \Bbbk^4, \quad \begin{bmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \end{bmatrix} \mapsto (x_1, x_2, x_3, x_4),$$

is generically recoverable on $\mathcal{X}$. Note that $\tilde{\mu}$ is not everywhere recoverable: if $x_1 = x_4 = 0$, the fiber $\mu^{-1}(\mu(A))$ is at least 1-dimensional. Like Example 6, there is no everywhere recoverable coordinate projection except for the identity map.

Conditions under which recoverability and identifiability of low-rank matrix completion problems holds were discussed in [17–22], among others. ◆

Given a polynomial measurement map, one can explicitly determine whether it is generically recoverable by computing the rank of the differential of a suitable map. This is discussed next.

Question 5. *Given a specific polynomial map $\mu : \mathcal{X} \to \Bbbk^s$, how can one determine whether μ is generically recoverable?* ?

Answer 5. First consider the implicit setting (IS) and let the variety be given as $\mathcal{X} = \{x \in \Bbbk^n : f_1(x) = \cdots = f_k(x) = 0\}$. Write $f = (f_1, \ldots, f_k)$, which can be regarded as a polynomial map from $\Bbbk^n$ to $\Bbbk^k$. A sufficient condition to guarantee that μ is generically recoverable is that there exists a point $x \in \mathcal{X}$ satisfying

$$\ker(\mathrm{d}_x\mu) \cap \ker(\mathrm{d}_x f) = 0.$$

Explicitly, regarding $\mathrm{d}_x\mu$ as an $s \times n$ matrix and $\mathrm{d}_x f$ as a $k \times n$ matrix, this condition is equivalent to the $(s + k) \times n$ matrix $\begin{bmatrix} \mathrm{d}_x\mu \\ \mathrm{d}_x f \end{bmatrix}$ having rank n.

Now consider the explicit setting (ES), and let $\mathcal{X}$ be parameterized explicitly by $\phi = (\phi_1, \ldots, \phi_n) : \Bbbk^\ell \to \Bbbk^n$. Then a sufficient condition to guarantee that μ is generically recoverable is that there exists a point $t \in \Bbbk^\ell$ such that the Jacobian matrix of $\mu \circ \phi : \Bbbk^\ell \to \Bbbk^s$ at t has rank $d = \dim \mathcal{X}$. ✓

Remark 3.1. There may be several ways to present $\mathcal{X} = \{x \in \Bbbk^n : f_1(x) = \cdots = f_k(x) = 0\}$ and even the ideal $(f_1, \ldots, f_k)$ generated by the f_i's may change. It is always best to choose the f_i's that generate the whole ideal of polynomial functions that vanish on $\mathcal{X}$.

Remark 3.2. We note that the condition that a map $\mu : \mathcal{X} \to \Bbbk^s$ is generically recoverable is equivalent to the fact that the dimension of the image $\mu(\mathcal{X})$ is equal to d.

The second criterion is especially convenient for low-dimensional varieties embedded in a high-dimensional ambient space. In this case, it suffices to compute the rank of an $\ell \times s$ matrix, where ℓ is the number of parameters in the parameterization of $\mathcal{X}$ and s the number of measurements, both of which can be similar in size to $d = \dim \mathcal{X}$. In a numerical setting, computing the rank consists of counting the number of significant singular values.

The following question addresses explicitly the case of coordinate projection, in the setting of parameterized models.

Question 6. *If $\mathcal{X}$ is explicitly parameterized by $\phi : \Bbbk^\ell \to \Bbbk^n$ as in (ES), how can one determine a subset of $s \geq d$ coordinates such that the coordinate projection on these s coordinates is generically recoverable?* ?

Answer 6. The s distinct coordinates $i_1, \ldots, i_s$ should be selected so that the $s \times \ell$ matrix $\left[\frac{\partial \phi_{i_p}}{\partial t_j}\right]_{\substack{p=1,\ldots,s \\ j=1,\ldots,\ell}}$ has rank exactly d for at least one $t \in \Bbbk^\ell$. ✓

We can now use this answer to understand Example 7.

Example 8 (Rank-1 matrix recovery, continued I)**.** Write ϕ_{ij} for the components of the map ϕ of (1). In Example 7, we claimed that the coordinate projection defined by $\hat{\mu}(\left[\begin{smallmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \end{smallmatrix}\right]) = (x_1, x_2, x_4, x_5)$ is not recoverable. We verify that the sufficient condition from Question 5 is not satisfied. Indeed, the Jacobian matrix of $\mu \circ \phi : (t_1, \ldots, t_5) \mapsto (t_1 t_3, t_1 t_4, t_2 t_3, t_2 t_4)$ is

$$
J = \mathrm{d}_t(\mu \circ \phi) = \begin{array}{c} \\ \phi_{1,1} \\ \phi_{1,2} \\ \phi_{2,1} \\ \phi_{2,2} \end{array} \begin{array}{ccccc} \partial_1 & \partial_2 & \partial_3 & \partial_4 & \partial_5 \\ \left[\begin{array}{ccccc} t_3 & 0 & t_1 & 0 & 0 \\ t_4 & 0 & 0 & t_1 & 0 \\ 0 & t_3 & t_2 & 0 & 0 \\ 0 & t_4 & 0 & t_2 & 0 \end{array}\right], \end{array}
$$

which has rank at most 3 for every choice of $(t_1, \ldots, t_5)$ because $t_1 \partial_1 + t_2 \partial_2 = t_3 \partial_3 + t_4 \partial_4$.

On the other hand, the coordinate projection $\tilde{\mu}$ is generically recoverable. In this case $\tilde{\mu} \circ \phi(t_1, \ldots, t_5) = (t_1 t_3, t_1 t_4, t_1 t_5, t_2 t_3)$ and its Jacobian matrix is

$$
\tilde{J} = \mathrm{d}_t(\tilde{\mu} \circ \phi) = \begin{array}{c} \\ \phi_{1,1} \\ \phi_{1,2} \\ \phi_{1,3} \\ \phi_{2,1} \end{array} \begin{array}{ccccc} \partial_1 & \partial_2 & \partial_3 & \partial_4 & \partial_5 \\ \left[\begin{array}{ccccc} t_3 & 0 & t_1 & 0 & 0 \\ t_4 & 0 & 0 & t_1 & 0 \\ t_5 & 0 & 0 & 0 & t_1 \\ 0 & t_3 & t_2 & 0 & 0 \end{array}\right]. \end{array} \qquad (2)
$$

For $t_1 = t_3 = 1$ and $t_2 = t_4 = t_5 = 0$, the $\mathrm{rk}(\tilde{J}) = 4$. Hence, $\tilde{\mu}$ is generically recoverable by Question 5. ◆

3.3. Identifiability

Recall that a map $\mu : \mathcal{X} \to \Bbbk^s$ is *identifiable at a point* $x \in \mathcal{X}$ if the fiber $\mu^{-1}(\mu(x)) = \{x\}$ consists of a single point. We say that μ is *everywhere identifiable* if it is identifiable at every $x \in \mathcal{X}$ and that it is *generically identifiable* if it is identifiable at generic points $x \in \mathcal{X}$.

In this section we discuss three settings. The first one considers identifiability for generic maps and generic points on $\mathcal{X}$. This is the same setting as the one discussed in the case of recoverability in the previous subsection.

Question 7. *Under what conditions is a generic linear map $\mu : \mathcal{X} \to \Bbbk^s$ generically identifiable?* ?

Answer 7. If $\mathcal{X}$ is a linear space, then μ is generically identifiable if and only if $s \geq d$. If $\mathcal{X}$ is not a linear subspace and if $s \geq d + 1$, then μ is generically identifiable. If $\Bbbk = \mathbb{C}$, then the condition $s \geq d + 1$ is also necessary. ✓

The analogous question for polynomial maps has a similar answer.

Question 8. *Under what conditions is a map $\mu : \mathcal{X} \to \Bbbk^s$ defined by s generic polynomials of degree p generically identifiable?* ?

Answer 8. Such a polynomial map is generically identifiable if $s \geq d + 1$. When $\Bbbk = \mathbb{C}$, $d > 0$, and $p > 1$, the condition $s \geq d + 1$ is necessary as well. ✓

Now, we move from questions about generic data to questions about properties that hold *everywhere*. In Question 9 we wonder when a generic linear map is everywhere identifiable and Question 10 asks for the existence of a Euclidean open set of such maps.

A key role in these questions is played by the *set of differences*

$$\Delta(\mathcal{X}) := \{x_1 - x_2 : x_1, x_2 \in \mathcal{X}\} \subseteq \Bbbk^n.$$

This is the image of the map $\mathcal{X} \times \mathcal{X} \to \Bbbk^n, (x_1, x_2) \mapsto x_1 - x_2$. The closure of $\Delta(\mathcal{X})$ is an algebraic variety of dimension at most $2d$. When $\mathcal{X}$ is a cone over $0 \in \Bbbk^n$, then $\Delta(\mathcal{X})$ is the *Minkowski sum* of $\mathcal{X}$ with itself, and its closure is called the *secant variety* of $\mathcal{X}$ [46, 83].

One can immediately observe that for a linear map $\hat{\mu} : \Bbbk^n \to \Bbbk^s$, one has

$$\hat{\mu}(x_1) = \hat{\mu}(x_2) \quad \text{if and only if} \quad \hat{\mu}(x_1 - x_2) = 0.$$

In other words, $\mu = \hat{\mu}|_{\mathcal{X}}$ is everywhere identifiable if and only if the kernel $\ker(\hat{\mu})$ intersects $\Delta(\mathcal{X})$ only at $0 \in \Bbbk^n$.

Question 9. *Under which conditions is a generic linear map $\mu : \mathcal{X} \to \Bbbk^s$ everywhere identifiable?*　　　*?*
Answer 9. A generic linear map μ is everywhere identifiable if $s \geq 1 + \dim \Delta(\mathcal{X})$. In particular, if $s \geq 2d+1$, μ is everywhere identifiable.　　✓

We provide two interesting examples; one where the answer to Question 9 is not optimal, and one where it is.

Example 9 (Nonoptimality of Answer 9 for cones). Suppose $\mathcal{X}$ is a cone over $0 \in \Bbbk^n$: for every $x \in \mathcal{X}$ and every $t \in \Bbbk$, one has $tx \in \mathcal{X}$. Then, $\Delta(\mathcal{X})$ is also a cone over 0; let d' be its dimension. In this case, a generic projection to $\Bbbk^{d'}$ is injective on $\mathcal{X}$, and therefore everywhere identifiable.　　◆

Example 10 (Optimality of Answer 9 for the twisted cubic). Let $\mathcal{X}$ be the curve in $\mathbb{C}^3$ defined as the image of $\phi : \mathbb{C} \to \mathbb{C}^3, t \mapsto (t, t^2, t^3)$. In classic algebraic geometry this is called the *twisted cubic curve* in $\mathbb{C}^3$. The set of differences is
$$\Delta(\mathcal{X}) = \{(t_1 - t_2, t_1^2 - t_2^2, t_1^3 - t_2^3) : t_1, t_2 \in \mathbb{C}\}.$$
In this case, a generic line through 0 intersects $\Delta(\mathcal{X})$ in a point different from 0. Therefore, a generic linear map $\hat{\mu} : \mathbb{C}^3 \to \mathbb{C}^2$ satisfies $\ker(\hat{\mu}) \cap \Delta(\mathcal{X}) \neq 0$. This determines a point (in fact at least two points) $x \in \mathcal{X}$ at which μ is not identifiable. Hence, the bound $s \geq \dim \Delta(\mathcal{X}) + 1 = 3$ of the answer to Question 9 is optimal.

There can be special projections with $s < \dim \Delta(\mathcal{X}) + 1$ taking $\mathcal{X}$ injectively to $\mathbb{C}^s$, even if generic projections are not injective. For instance, in the case of the twisted cubic, the projection $(x_1, x_2, x_3) \mapsto x_1$ maps $\mathcal{X}$ injectively to $\mathbb{C}^1$.　　◆

The precise meaning of *generic* is given in terms of open subsets in the Zariski topology, as explained in Section 4. Over the complex numbers, these open sets are dense, with respect to the classic Euclidean topology, in their Zariski closure. As a consequence, requiring a property to hold generically is usually a strong condition. It is therefore natural to wonder whether one can relax the genericity condition, requiring that a property holds on a set containing a Euclidean open subset. Over the real numbers, it is often possible to determine Euclidean open subsets of linear projections $\hat{\mu} : \mathbb{R}^n \to \mathbb{R}^s$ which are identifiable (over the reals) even though the corresponding extension $\hat{\mu}_{\mathbb{C}} : \mathbb{C}^n \to \mathbb{C}^s$ is not. However, the existence of such open sets of projections depends on the geometry of the model, and not only on its dimension, as pointed out in the next question.

Question 10. *Under what conditions does there exist a Euclidean open set of everywhere identifiable linear maps $\mu : \mathcal{X} \to \mathbb{R}^s$?*　　　*?*
Answer 10. There is an open set of everywhere identifiable linear maps if we have $s \geq 1 + \dim \Delta(\mathcal{X})$. In particular, this holds if $s \geq 2d + 1$.　　✓

We provide an example of a model $\mathcal{X}$ for which there exists an open set of identifiable linear maps $\mu : \mathcal{X} \to \mathbb{R}^s$ for $s < 1 + \dim(\Delta(\mathcal{X}))$.

Example 11 (The twisted cubic reconsidered)**.** Let $\mathcal{X}$ be the twisted cubic from Example 10. Its set of differences over the reals is

$$\Delta(\mathcal{X}) = \{(t_1 - t_2, t_1^2 - t_2^2, t_1^3 - t_2^3) : t_1, t_2 \in \mathbb{R}\}.$$

In particular, if $(x_1, x_2, x_3) \in \Delta(\mathcal{X})$, then x_1 and x_3 have the same sign. As a consequence, any linear map $\hat{\mu} : \mathbb{R}^3 \to \mathbb{R}^2$ whose kernel is the line $\ker \hat{\mu} = \langle (1, \alpha, \beta) \rangle$ for some $\beta < 0$ satisfies $\ker(\hat{\mu}) \cap \Delta(\mathcal{X}) = 0$, so that μ is everywhere identifiable. This is a Euclidean open set of linear maps. ◆

Finally, we return to the case of coordinate projections from data completion applications. We stress that it is not possible in general to guarantee identifiability results in this special setting. This is highlighted in the following example.

Example 12. Let $\mathcal{X}$ be an algebraic variety defined implicitly as the vanishing locus of polynomials $f_1, \ldots, f_k$ and assume $\mathcal{X}$ does not lie in a coordinate hyperplane. Suppose that each variable x_i appears in $f_1, \ldots, f_k$ only with even powers. In this case, if $(\xi_1, \ldots, \xi_n)$ is a point of $\mathcal{X}$, then any point of the form $(\pm\xi_1, \ldots, \pm\xi_n)$ belongs to $\mathcal{X}$ as well. Consequently, if $\hat{\mu} : \mathbb{k}^n \to \mathbb{k}^s$ is a coordinate projection different from the identity, then $\mu = \hat{\mu}|_{\mathcal{X}}$ is not generically identifiable. For instance, if $x = (\xi_1, \ldots, \xi_n) \in \mathcal{X}$ is a generic point, and $\hat{\mu}$ is the map that projects away the last coordinate, then $\mu^{-1}(\mu(x)) = \{(\xi_1, \ldots, \xi_{n-1}, \pm\xi_n)\}$. Note that $\xi_n \neq 0$ because of the genericity condition and the fact $\mathcal{X}$ does not lie in a coordinate hyperplane.

A minimal example is the following:

$$\mathcal{X} = \left\{ (x_1, x_2, x_3) \in \mathbb{k}^3 : x_1^2 + x_2^2 + x_3^2 - 1 = x_1^2 x_2^2 - x_3^4 - 1 = 0 \right\}.$$

In this case, the variety $\mathcal{X} \subseteq \mathbb{k}^3$ has dimension 1 and does not admit a generically identifiable coordinate projection. On the other hand, a generic projection to $\mathbb{k}^2$ is generically identifiable. ◆

A consequence of Example 12 is that in general the only everywhere identifiable coordinate projection is the identity map. This also follows from the analogous result on recoverability, as seen in Example 6.

Example 13 (The rational normal curve)**.** The 1-dimensional rational normal curve is again an interesting example: $\mathcal{X} = \{(t, t^2, t^3, \ldots, t^n) : t \in \mathbb{k}\} \subset \mathbb{k}^n$. The answer to Question 9 ensures there is an identifiable projection to $1 + \dim \Delta(\mathcal{X}) = 3$ coordinates. For example, projecting to the first three coordinates (t, t^2, t^3) enables us to recover t (from the first element alone). In fact, this shows once more that the answer to Question 9 is not optimal: a projection to the first coordinate of $\mathbb{k}^n$ is already everywhere identifiable.

On the other hand, there are several generically nonidentifiable coordinate projections, even using much more than $1 + \dim \Delta(\mathcal{X}) = 3$ coordinates. Consider for example projecting to any subset of $3 \leq s \leq \frac{n}{2}$ even coordinates. This leaves at least two possibilities for the reconstruction: If $(t_0, t_0^2, t_0^3, \ldots, t_0^n) \in \mathcal{X}$ is one solution, then necessarily $(-t_0, t_0^2, -t_0^3, \ldots, (-t_0)^n)$ is another solution. ◆

We conclude the discussion of identifiability emphasizing the intricacy of the questions in this section, even in the mathematically friendlier setting of complex projective geometry. For example, it is an open problem whether a d-dimensional complex projective variety may be algebraically injected to a projective space of dimension $2d$ using arbitrary algebraic maps; see [84–86] and [87, Question 3]. A weaker negative result is known: there are algebraic curves $\mathcal{X}$ in a complex 3-dimensional projective space that cannot be injectively projected to the projective plane [88]. In our affine setting this corresponds to an algebraic surface ($d = 2$), inside a 4-dimensional space ($n = 4$) so that for any projection to a plane ($s = 2$) there exist two lines inside the surface that are mapped to the same line.

Finally, we point out that the structure described in Example 12 is very special and in most cases we expect a generically identifiable coordinate projection $\mathbb{k}^n \to \mathbb{k}^s$ to exist for $s = d + 1$, as in the case of generic projections. We sketch an argument we believe could be formalized to guarantee the existence of such projection. Recall that from the answer to Question 4, there exists a generically recoverable coordinate projection to d coordinates. Let $\mu : \mathcal{X} \to \mathbb{k}^d$ be such projection. A classic construction in algebraic geometry is the *second fiber power of $\mathcal{X}$ over μ*, denoted $\mathcal{X}^{\mu[2]}$. Informally, this is an algebraic variety containing pairs of points having the same image under μ. The variety $\mathcal{X}^{\mu[2]}$ is reducible; its irreducible components of

14

highest dimension have dimension d and there are at least two of them. One can prove that if the number of d-dimensional components of $\mathcal{X}^{\mu[2]}$ is exactly two, then there is a coordinate projection $\mu' : \mathcal{X} \to \Bbbk^{d+1}$ which is generically identifiable; in fact, μ' factors μ, in the sense that μ projects away all the coordinates projected away by μ' plus an additional one.

3.4. Continuity

The answers from the previous subsections give several conditions under which the measurement map is locally or globally injective. In particular, under these conditions, the measurement map admits a (local) inverse. The final requirement for an inverse problem to be well posed is that this (local) inverse map is at least continuous. In fact, much stronger properties hold.

First, we recall terminology to succinctly ask the questions. We refer to [89, 90] for the theory and definitions. A map between manifolds is called *smooth* if it can be differentiated arbitrarily many times; that is, if it is C^∞ on the local charts defining the manifolds. Smooth maps are continuous and locally Lipschitz continuous.

Let $\mathcal{M}$ be a smooth $\Bbbk$-manifold, and let $\mu : \mathcal{M} \to \Bbbk^s$ be a map. The map μ is a *global diffeomorphism onto its image* if $\mu(\mathcal{M})$ is a smooth manifold, μ is smooth and injective, and μ^{-1} is smooth [90]. The map μ is a *local diffeomorphism onto its image* at $x \in \mathcal{M}$ if there is a Euclidean open neighborhood $U_x \subseteq \mathcal{M}$ of x such that the restriction $\mu|_{U_x}$ is a global diffeomorphism onto its image and $\mu(U_x)$ is open in $\mu(\mathcal{M})$ [90].

We investigate which measurement maps μ are local or global diffeomorphisms.

Question 11. *Under what conditions does a generic linear map $\mu : \mathcal{X} \to \Bbbk^s$ have the property that there exists a Zariski open manifold $\mathcal{X}' \subseteq \mathcal{X}$ such that $\mu(\mathcal{X}')$ is a smoothly embedded submanifold of $\Bbbk^s$ and $\mu|_{\mathcal{X}'}$ is a local diffeomorphism onto its image at every point $x \in \mathcal{X}'$?*   *?*

Answer 11. There exists such a manifold $\mathcal{X}'$ if $s \geq d$. Moreover, the bound is sharp, independently of the variety $\mathcal{X}$.   ✓

Question 12. *Under what conditions does a generic linear map $\mu : \mathcal{X} \to \Bbbk^s$ have the property that there exists a Zariski open manifold $\mathcal{X}' \subset \mathcal{X}$ such that $\mu(\mathcal{X}')$ is a smoothly embedded submanifold of $\Bbbk^s$ and $\mu|_{\mathcal{X}'}$ is a global diffeomorphism onto its image?*   *?*

Answer 12. There exists such a manifold $\mathcal{X}'$ if $s \geq d + 1$. If $\mathcal{X}$ is not an affine space, and if $\Bbbk = \mathbb{C}$, the bound is sharp independently of the variety $\mathcal{X}$.   ✓

The fact that in both these questions $\mu(\mathcal{X}')$ is a smoothly embedded submanifold of $\Bbbk^s$ is important when solving inverse problems using methods from Riemannian optimization [91].

The proofs to the answers of Questions 11 and 12 reveal the structure of $\mathcal{X}'$. In the setting of Question 11, the manifold $\mathcal{X}'$ has the property that $\mathcal{X} \setminus \mathcal{X}'$ contains the singular locus of $\mathcal{X}$, the locus of points $x \in \mathcal{X}$ with $(T_x \mathcal{X}) \cap \ker(\mu) \neq 0$, and the preimage of the singular locus of $\mu(\mathcal{X})$. In the setting of Question 12, $\mathcal{X} \setminus \mathcal{X}'$ additionally contains the set of points $x \in \mathcal{X}$ on which μ is not injective.

Example 14 (Rank-1 matrix recovery, continued II). Let $\mathcal{X}$ be the variety of 2×3 matrices of rank one, parameterized as in (1). In Example 8, we proved that the coordinate projection $\tilde{\mu}$ defined by $\tilde{\mu}\left(\left[\begin{smallmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \end{smallmatrix}\right]\right) = (x_1, x_2, x_3, x_4)$ is generically recoverable. Let $\mathcal{X}' = \mathcal{X} \cap \{x_1 \neq 0\}$, which is Zariski open in $\mathcal{X}$. Then, a point $y = (x_1, \ldots, x_4) \in \mu(\mathcal{X}) \subseteq \Bbbk^4$ has a unique preimage

$$ x = (\mu|_{\mathcal{X}'})^{-1}(x_1, \ldots, x_4) = \begin{bmatrix} x_1 & x_2 & x_3 \\ x_4 & x_4 x_1^{-1} x_2 & x_4 x_1^{-1} x_3 \end{bmatrix}. $$

This shows that $\mu|_{\mathcal{X}'}$ is smooth and injective. The argument of Example 8 shows that $\mathrm{d}_x \tilde{\mu} : T_x \mathcal{X} \to \Bbbk^4$ is a linear isomorphism for every $x \in \mathcal{X}'$. Consequently, the image $\tilde{\mu}(\mathcal{X}')$ is an immersed submanifold of $\Bbbk^4$ by [90, Proposition 5.18] and $\tilde{\mu}|_{\mathcal{X}'}$ is a global diffeomorphism between $\mathcal{X}'$ and $\tilde{\mu}(\mathcal{X}')$.

Note that even though $s = 4 = \dim \mathcal{X}$, this example does not contradict the sharpness result in the answer to Question 12 because the considered coordinate projection is not generic.   ♦

The inverse problem (IP) is well posed for generic $x \in \mathcal{X}$ and generic choice of linear projection $\mu : \mathcal{X} \to \Bbbk^s$ with $s \geq d+1$ because of the answer to Question 12. We argued in Section 2 that computing a global Lipschitz constant for μ^{-1} is a difficult problem. Instead we proposed to compute the smallest local Lipschitz constant, that is, the condition number of μ^{-1}.

Question 13. *How does one compute the condition number of the inverse problem* (IP) *at a smooth point* $x \in \mathcal{X}$? ?

Answer 13. The condition number with respect to the Euclidean metrics on both domain and codomain at a smooth point $x \in \mathcal{X}$ is $\kappa(x) = \frac{1}{\sigma_d(\mathrm{d}_x\mu)}$, where $\sigma_d(\mathrm{d}_x\mu)$ is the dth largest singular value of the derivative of $\mu : \mathcal{X} \to \Bbbk^s$ at x.

More concretely, let $M \in \Bbbk^{s \times n}$ be the matrix representing the linear map $\hat{\mu} : \Bbbk^n \to \Bbbk^s$ in the ambient space. Let $Q \in \Bbbk^{n \times d}$ be a matrix whose columns form an orthonormal basis of the tangent space $T_x\mathcal{X}$. Then $\kappa(x) = \frac{1}{\sigma_d(MQ)}$. ✓

Example 15 (Rank-1 matrix recovery, continued III)**.** Let $\mathcal{X}$ be the variety 2×3 matrices of rank one. The tangent space of $\mathcal{X} = \mathrm{im}(\phi)$ at a point $X = \phi(t)$ is spanned by the columns of the Jacobian matrix of ϕ,

$$
J = \mathrm{d}_t\phi = \begin{array}{c} \\ \phi_{1,1} \\ \phi_{1,2} \\ \phi_{1,3} \\ \phi_{2,1} \\ \phi_{2,2} \\ \phi_{2,3} \end{array} \begin{array}{ccccc} \partial_1 & \partial_2 & \partial_3 & \partial_4 & \partial_5 \\ \begin{bmatrix} t_3 & 0 & t_1 & 0 & 0 \\ t_4 & 0 & 0 & t_1 & 0 \\ t_5 & 0 & 0 & 0 & t_1 \\ 0 & t_3 & t_2 & 0 & 0 \\ 0 & t_4 & 0 & t_2 & 0 \\ 0 & t_5 & 0 & 0 & t_2 \end{bmatrix} \end{array}.
$$

Using, for example, a rank-revealing QR-decomposition or the singular value decomposition of J, one determines a matrix Q whose columns form an orthonormal basis of the column span of J. This is the matrix Q required in Answer 13.

For instance, consider $X = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{bmatrix}$, and $t = (1, 2, 1, 1, 1) \in \phi^{-1}(X)$. Then, the matrix Q can be computed numerically via a (truncated) singular value decomposition of J:

$$
Q \approx \begin{bmatrix} -0.258199 & -0.329793 & -0.156748 & 0.516398 \\ -0.258199 & 0.300644 & -0.207235 & 0.516398 \\ -0.258199 & 0.029148 & 0.363983 & 0.516398 \\ -0.516398 & -0.659586 & -0.313497 & -0.258199 \\ -0.516398 & 0.601289 & -0.414469 & -0.258199 \\ -0.516398 & 0.058297 & 0.727966 & -0.258199 \end{bmatrix}.
$$

Let M be the matrix representing the coordinate projection $\tilde{\mu}$ of Example 7. Then, the condition number of the inverse problem at the (locally unique) solution X is $\sigma_4(MQ)^{-1} \approx 3.8730$.

On the other hand, the coordinate projection $\hat{\mu}$ of Example 7 is not recoverable, or equivalently not locally identifiable. The fibers of $\hat{\mu}$ have positive dimension, so that from (CN) the condition number is infinite. This can be verified numerically. Composing Q with the matrix of representing $\hat{\mu}$, we obtain a matrix whose singular values, rounded to two significant digits, are 1.0, 1.0, 0.82, and $6.2 \cdot 10^{-17}$. The numerically computed condition number is thus $1.6 \cdot 10^{16}$. ◆

The last question assumed that we are given as input a data point $y \in \mu(\mathcal{X})$. If the data point is $y \notin \mu(\mathcal{X})$ and if one seeks to solve the least squares minimization problem $\min_{x \in \mathcal{X}} \frac{1}{2}\|\mu(x) - y\|^2$, then the condition number of this alternative problem is determined both by the condition number of μ^{-1} and the *curvature* of $\mu(\mathcal{X})$ at $\mu(x)$; see [92] for more precise statements. For the case where $\mathcal{X}$ is the variety of low-rank matrices, the condition number of the associated least-squares problem was worked out in detail in [93].

4. The main technical tools: Classic algebraic geometry

In this section, we provide the relevant definitions and preliminaries from classic algebraic geometry, in the real and complex settings. We use them in Section 5 to prove the correctness of the answers given in Section 3. We refer to [45, 46] for a general introduction and to [47, 94] for a more precise account.

Let $\Bbbk = \mathbb{R}$ or $\Bbbk = \mathbb{C}$. An *algebraic variety* is a subset $\mathcal{X} \subseteq \Bbbk^n$ defined as the vanishing set of a set of polynomials $f_1, \ldots, f_k$ in n variables $x_1, \ldots, x_n$ with coefficients in $\Bbbk$. The *ideal*

$$I^{\Bbbk}(\mathcal{X}) := \{f \in \Bbbk[x_1, \ldots, x_n] : f(x) = 0 \text{ for all } x \in \mathcal{X}\}$$

is the set of polynomials vanishing on $\mathcal{X}$. Hilbert's basis theorem states that $I^{\Bbbk}(\mathcal{X})$ is finitely generated, say $I^{\Bbbk}(\mathcal{X}) = \langle g_1, \ldots, g_r \rangle$. Depending on whether $\Bbbk = \mathbb{R}$ or $\Bbbk = \mathbb{C}$, we say that $\mathcal{X}$ is a real or complex algebraic variety, and that $I^{\Bbbk}(\mathcal{X})$ is its real or complex ideal.

A variety $\mathcal{X}$ is *irreducible* if it cannot be expressed as the union of two proper subvarieties. Every variety $\mathcal{X}$ is an irredundant finite union of irreducible subvarieties, called *irreducible components* of $\mathcal{X}$. For an arbitrary subset $\mathcal{X} \subseteq \Bbbk^n$, we say that $\mathcal{X}$ is irreducible if its Zariski closure $\overline{\mathcal{X}} \subseteq \Bbbk^n$ (see below) is irreducible.

A *semi-algebraic* set in $\mathbb{R}^n$ is a finite union of subsets of algebraic varieties satisfying a set of (strict) polynomial inequalities. In particular, every algebraic variety is a semi-algebraic set whose set of defining inequalities is empty.

4.1. Topology on algebraic varieties

The *Euclidean topology* of $\Bbbk^n$ induces a topology on every subset S of $\Bbbk^n$. We refer to this subspace topology on S as the Euclidean topology on S. The set of subvarieties of a variety $\mathcal{X}$ is closed under taking arbitrary intersections and finite unions; hence, algebraic subvarieties of $\mathcal{X}$ define the closed subsets of a topology on $\mathcal{X}$. This is the *Zariski topology* on $\mathcal{X}$.

The Zariski topology is coarser than the Euclidean topology: if $\mathcal{X}$ is irreducible and U is a Zariski open subset of $\mathcal{X}$, then U is open and dense in $\mathcal{X}$ with respect to the Euclidean topology. The subset $\mathcal{X} \setminus U$ is a proper subvariety of $\mathcal{X}$ and it has measure zero with respect to the measure on $\mathcal{X}$ induced by the Lebesgue measure on $\Bbbk^n$. Hence, a Zariski open subset U also has full measure on $\mathcal{X}$.

Let $\mathcal{X}$ be a semi-algebraic set in $\mathbb{R}^n$ and let $\overline{\mathcal{X}}$ be its closure in the Zariski topology of $\mathbb{R}^n$. Let $\mathcal{X}^{\circ}$ be the Euclidean interior of $\mathcal{X}$ in $\overline{\mathcal{X}}$. Then $\mathcal{X}^{\circ}$ is semi-algebraic as well and $\overline{\mathcal{X}^{\circ}} = \overline{\mathcal{X}}$ [67, Proposition 2.2.2].

It is common in algebraic geometry, as we have done in the foregoing sections, to state results *generically* or in terms of *generic points*.

Definition 4.1 (Genericity). Let $\mathcal{X} \subseteq \Bbbk^n$ be an irreducible algebraic variety. A property $\mathcal{P}$ holds *generically on $\mathcal{X}$* if it holds for all points in a Zariski open subset $U \subseteq \mathcal{X}$. The points of U are called *generic points* (with respect to the property $\mathcal{P}$). The set $\Sigma_{\mathcal{P}} = \mathcal{X} \setminus U$ is a proper subvariety of $\mathcal{X}$, called the *discriminant* of the property $\mathcal{P}$.

The discriminant $\Sigma_{\mathcal{P}}$ of a property can be tedious to describe explicitly. Nevertheless, if a property $\mathcal{P}$ holds generically, then the discriminant $\Sigma_{\mathcal{P}}$ exists and it has Lebesgue measure zero.

The space of linear maps $\Bbbk^n \to \Bbbk^s$ is a vector space isomorphic to $\Bbbk^{ns}$. Similarly, the space of polynomial maps of a fixed degree pattern forms a vector space. Thus, one can state results in terms of generic linear maps, or generic polynomial maps, meaning that the results hold for generic points in the corresponding vector space.

4.2. Dimension and tangent space of an algebraic variety

Let $\mathcal{X} \subseteq \Bbbk^n$ be an irreducible variety with ideal $I^{\Bbbk}(\mathcal{X}) = \langle g_1, \ldots, g_r \rangle$. Then, there is a Zariski open subset $\mathcal{X}' \subseteq \mathcal{X}$ such that the rank of the Jacobian matrix of $g = (g_1, \ldots, g_r)$ is constant on $\mathcal{X}'$. This set $\mathcal{X}'$ is called the set of *smooth* (or *non-singular*) points of $\mathcal{X}$. The dimension of $\mathcal{X}$ is defined as

$$\dim \mathcal{X} = n - \operatorname{rk}(\mathrm{d}_x g) \quad \text{for any } x \in \mathcal{X}'.$$

If $\Bbbk = \mathbb{R}$ (respectively, $\Bbbk = \mathbb{C}$), then $\mathcal{X}'$ is a smooth real (respectively, complex) embedded submanifold of $\Bbbk^n$ and its dimension as a real (respectively, complex) manifold is $\dim \mathcal{X}$.

For $x \in \mathcal{X}'$, the linear subspace

$$T_x \mathcal{X} = \ker(\mathrm{d}_x g)$$

of $\Bbbk^n$ is the tangent space to $\mathcal{X}$ at x. Note that $\dim T_x \mathcal{X} = \dim \mathcal{X}$.

4.3. Degree of a variety

Let $\mathcal{X} \subset \mathbb{C}^n$ be a d-dimensional algebraic variety. A generic affine subspace of dimension $n - d$ intersects $\mathcal{X}$ in finitely many points by Bertini's Theorem (see, e.g., [46, Section 18] or [95, Theorem 0.5]). This number of points does not depend on the chosen subspace and is called the *degree* of $\mathcal{X}$.

The degree of a variety is a property of its embedding in $\mathbb{C}^n$. An immediate consequence of Bezout's Theorem [96, Theorem 7.7] is the following, describing the degree of the image of a variety under the Veronese embedding ν_r from (3).

Theorem 4.2. *Let $\mathcal{X} \subset \mathbb{C}^n$ be an algebraic variety of dimension d and degree δ. Then, $\nu_r(\mathcal{X})$ is an algebraic variety of dimension d and degree $r^d \cdot \delta$.*

The only irreducible varieties of degree one are affine subspaces.

4.4. Complexification of real algebraic varieties

Every real variety defines a complex variety via a process called complexification. Let $\mathcal{X} \subset \mathbb{R}^n$ be a real algebraic variety. The *complexification* $\mathcal{X}^{\mathbb{C}}$ of $\mathcal{X}$ is

$$\mathcal{X}^{\mathbb{C}} := \{x \in \mathbb{C}^n : f(x) = 0 \text{ for all } f \in I^{\mathbb{R}}(\mathcal{X})\}.$$

There are classic results relating the algebraic and geometric properties of a real algebraic variety to the ones of its complexification.

The following is a rephrasing of [97, Lemmata 6 and 8]:

Theorem 4.3. *Let $\mathcal{X} \subseteq \mathbb{R}^n$ be a real algebraic variety and let $\mathcal{X}^{\mathbb{C}}$ be its complexification. Regard $\mathbb{R}^n$ as a (real) subspace of $\mathbb{C}^n$. Then,*

1. *$\mathcal{X}^{\mathbb{C}}$ is the closure in the (complex) Zariski topology of $\mathcal{X}$;*

2. *the set of real polynomials in $I^{\mathbb{C}}(\mathcal{X}^{\mathbb{C}})$ is $I^{\mathbb{R}}(\mathcal{X})$;*

3. *the real dimension of $\mathcal{X}$ equals the complex dimension of $\mathcal{X}^{\mathbb{C}}$.*

Theorem 4.3 implies that a generic point in a real algebraic variety $\mathcal{X}$ is also a generic point of the complexification $\mathcal{X}^{\mathbb{C}}$. More precisely, let $\mathcal{X}$ be a real variety, and let $U \subseteq \mathcal{X}^{\mathbb{C}}$ be the Zariski open subset on which a certain property $\mathcal{P}$ holds. Then $U \cap \mathcal{X}$ is non-empty and Zariski open in $\mathcal{X}$. In particular, Theorem 4.3 guarantees that results in complex algebraic geometry that hold for generic elements of a complex algebraic variety can be interpreted in the real setting.

4.5. Maps between varieties

For two varieties $\mathcal{X} \subseteq \Bbbk^\ell$ and $\mathcal{Y} \subseteq \Bbbk^n$, let

$$\phi : \mathcal{X} \to \mathcal{Y}$$

be a map and write $\phi(t) = (\phi_1(t), \ldots, \phi_n(t))$. We say that ϕ is a *regular map* or a *morphism of varieties* if, for every $j = 1, \ldots, n$, ϕ_j is (the restriction to $\mathcal{X}$ of) a polynomial in ℓ variables $t_1, \ldots, t_\ell$, which can be regarded as coordinates on $\Bbbk^\ell$.

The image of a morphism of varieties is not necessarily a variety. However, the following fundamental results characterize images of algebraic varieties in the complex and in the real case respectively.

18

Theorem 4.4 (Chevalley's Theorem)**.** *Let $\mathcal{X} \subseteq \mathbb{C}^\ell$ be a variety and let $\phi : \mathcal{X} \to \mathbb{C}^n$ be a morphism of varieties. Then $\mathcal{Y} = \phi(\mathcal{X})$ is locally closed in the Zariski topology. Hence, $\mathcal{Y}$ contains a Zariski open subset of its Zariski closure. Moreover, the closures of $\mathcal{Y}$ in the Zariski and the Euclidean topologies coincide.*

Proof. We refer to [98, Theorem 4.19]. $\qquad\square$

In the real setting, the analog result is the Tarski–Seidenberg Theorem.

Theorem 4.5 (Tarski–Seidenberg Theorem)**.** *Let $\mathcal{X} \subseteq \mathbb{R}^\ell$ be a semi-algebraic set and let $\phi : \overline{\mathcal{X}} \to \mathbb{R}^n$ be a morphism of varieties, where $\overline{\mathcal{X}}$ denotes the Zariski closure of $\mathcal{X}$ in $\mathbb{R}^\ell$. Then $\phi(\mathcal{X})$ is a semi-algebraic subset of $\mathbb{R}^n$.*

Proof. We refer to [67, Section 2.2]. $\qquad\square$

In the complex case, the dimension of the (Zariski closure of the) image of an algebraic variety is controlled by the following result.

Theorem 4.6 (Dimension of the fibers)**.** *Let $\mathcal{X}, \mathcal{Y}$ be complex irreducible varieties and let $\phi : \mathcal{X} \to \mathcal{Y}$ be a morphism such that $\phi(\mathcal{X})$ is (Zariski) dense in $\mathcal{Y}$. Then for every $y \in \phi(\mathcal{X})$,*

$$\dim \phi^{-1}(y) \geq \dim \mathcal{X} - \dim \mathcal{Y},$$

and there is a Zariski open subset of $\mathcal{Y}$ for which equality holds.

Proof. The proof relies on rank conditions of the differential of the map ϕ. We refer to [46, Theorem 11.22] and [99, Theorem 1.25] for a complete proof. $\qquad\square$

Given a morphism of varieties $\phi : \mathbb{k}^\ell \to \mathbb{k}^n$, we say that ϕ is a *parameterization* of $\mathrm{im}(\phi) \subseteq \mathbb{k}^n$. This is the setting (ES). In this case, for every $t \in \mathbb{k}^\ell$, one has a *Jacobian map* $\mathrm{d}_t\phi : \mathbb{k}^\ell \to \mathbb{k}^n$ defined by the differential of ϕ at t, that is

$$\mathrm{d}_t\phi : \mathbb{k}^\ell \to \mathbb{k}^n, \quad v \mapsto \left[\frac{\partial}{\partial t_j} \phi_i(t) \right]_{\substack{i=1,\ldots,n \\ j=1,\ldots,\ell}} \cdot v,$$

where $\frac{\partial}{\partial t_j} \phi_i(t)$ denotes the partial derivative of the polynomial ϕ_i with respect to the variable t_j, evaluated at t. For $r \geq 0$, the set $V_r^\phi = \{ t \in \mathbb{k}^\ell : \mathrm{rk}\,(\mathrm{d}_t\phi) \leq r \}$ is Zariski closed, and its defining equations are given by the $(r+1) \times (r+1)$ minors of the Jacobian matrix, regarded as polynomials in the coordinates $t_1, \ldots, t_\ell$ of $\mathbb{k}^\ell$. Let $r_0 = \min\{r : V_r^\phi = \mathbb{k}^\ell\}$. Then r_0 is called the generic rank of ϕ and it is attained on the Zariski open subset $V_{r_0}^\phi \setminus V_{r_0-1}^\phi$.

In this setting, we provide the following theorem.

Theorem 4.7. *Let $\phi : \mathbb{k}^\ell \to \mathbb{k}^n$ be an algebraic parameterization having generic rank r_0. Let $\mathcal{X} = \overline{\phi(\mathbb{k}^\ell)}$ be the closure in the Zariski topology of the image of ϕ. Then $\mathcal{X}$ is an algebraic variety of dimension r_0.*

Moreover, if $x = \phi(t)$ is a smooth point of $\mathcal{X}$ with $t \in V_{r_0}^\phi$, then $T_x\mathcal{X} = \mathrm{im}(\mathrm{d}_t\phi)$.

Proof. Restricting to smooth points of the image, for $\mathbb{k} = \mathbb{R}$, the last statement follows from the Global Rank Theorem [90, Theorem 4.14 a]. As complex manifolds are in particular real manifolds the last statement also holds for $\mathbb{k} = \mathbb{C}$.

The statement about the dimension of $\mathcal{X}$ holds for a smooth point $x = \phi(t)$ of $\mathcal{X}$ with $t \in V_{r_0}^\phi$ because

$$\dim \mathcal{X} = \dim T_x\mathcal{X} = \dim \mathrm{im}(\mathrm{d}_t\phi) = r_0.$$

$\qquad\square$

Note that one can always choose generic points in the image of a parameterization like (ES). More precisely, if $\phi : \Bbbk^\ell \to \Bbbk^n$ is a polynomial parametrization of a subset of a variety $\mathcal{X} = \overline{\mathrm{im}(\phi)}$, and $\mathcal{P}$ is a property that holds generically on $\mathcal{X}$, then the property $\mathcal{P}$ also holds almost everywhere on the model $\mathrm{im}(\phi)$. This allows us to study the generic properties of the model $\mathrm{im}(\phi)$ through the closed algebraic variety $\mathcal{X}$.

A special role is played by the Veronese embedding $\nu_r : \Bbbk^n \to \Bbbk^{N(n,r)}$ where $N(n,r) = \binom{n+r}{r}$, which is

$$\begin{aligned} \nu_r : \Bbbk^n &\to \Bbbk^{N(n,r)}, \\ \mathbf{x} &\mapsto (\mathbf{x}^\alpha : \alpha \in \mathbb{N}^n, |\alpha| \leq r); \end{aligned} \tag{3}$$

that is, the parameterization defined by all the monomials of degree at most r in the coordinates $\mathbf{x} = (x_1, \ldots, x_n)$ of $\Bbbk^n$. The properties of the Veronese embedding are summarized in the next immediate result.

Lemma 4.8. *The map $\nu_r : \Bbbk^n \to \Bbbk^{N(n,r)}$ is an embedding. That is, ν_r is injective and its differential is injective at every point $x \in \Bbbk^n$. Moreover, if $f : \Bbbk^n \to \Bbbk$ is a polynomial of degree r, then there exists a linear map $L_f : \Bbbk^{N(n,r)} \to \Bbbk$ such that we have $f = L_f \circ \nu_r$.*

Since polynomial maps are differentiable, a geometric version of the implicit function theorem holds.

Theorem 4.9. *Let $\mathcal{X}, \mathcal{Y}$ be irreducible varieties, with $\dim \mathcal{X} = d$ and $\dim \mathcal{Y} = m$. Let $\phi : \mathcal{X} \to \mathcal{Y}$ be a morphism of varieties with $\mathcal{Y} = \overline{\phi(\mathcal{X})}$. Then, there exist Zariski open subsets $U_\mathcal{X} \subseteq \mathcal{X}$ and $U_\mathcal{Y} \subseteq \mathcal{Y}$ such that $U_\mathcal{X}$ and $U_\mathcal{Y}$ are smooth manifolds, $\phi|_{U_\mathcal{X}} : U_\mathcal{X} \to U_\mathcal{Y}$ is a smooth map and for every $y \in U_\mathcal{Y}$, the fiber $\phi^{-1}(y)$ is a smooth manifold of dimension $d - m$.*

Proof. Let $U'_\mathcal{X} \subset \mathcal{X}$ be the subset of smooth points on which ϕ is regular; that is,

$$U'_\mathcal{X} = \{x \in \mathcal{X} : x \text{ is smooth in } \mathcal{X} \text{ and } \mathrm{rk}\,(\mathrm{d}_x\phi) = m\}.$$

By Sard's Theorem [90, Theorem 6.8], $U'_\mathcal{X}$ is non-empty and it is Zariski open because it is the intersection of two Zariski open subsets. Let $U_\mathcal{Y}$ be the subset of $\phi(U'_\mathcal{X})$ consisting of smooth points of $\mathcal{Y}$. Then $U_\mathcal{Y}$ is Zariski open in $\mathcal{Y}$. Let $U_\mathcal{X} = \phi^{-1}(U_\mathcal{Y})$. Since ϕ is a regular map, $U_\mathcal{X}$ is Zariski open in $\mathcal{X}$.

The restriction $\phi|_{U_\mathcal{X}} : U_\mathcal{X} \to U_\mathcal{Y}$ is a smooth map of manifolds, and for every $x \in U_\mathcal{X}$, $\mathrm{rk}\,(\mathrm{d}_x\phi) = m$. By [90, Theorem 5.22], we conclude that the fibers of $\phi|_{U_\mathcal{X}}$ are smooth manifolds of dimension $d - m$. $\qquad\square$

A consequence of Theorem 4.9 is that the rank of the differential at a point controls the local injectivity of the map. We have the following result.

Corollary 4.10. *Let $\mathcal{X} \subseteq \Bbbk^n$ be an irreducible variety. Let $\hat{\mu} : \Bbbk^n \to \Bbbk^s$ be a morphism of varieties and let $x \in \mathcal{X}$. If $\ker(\mathrm{d}_x\hat{\mu}) \cap T_x\mathcal{X} = 0$, then there exists a Euclidean open neighborhood $U_x \subseteq \mathcal{X}$ of x such that $\hat{\mu}|_{U_x} : U_x \to \Bbbk^s$ is a diffeomorphism onto its image.*

Proof. Let $\psi : \mathcal{X} \to \Bbbk^s$ be the restriction of $\hat{\mu}$ to $\mathcal{X}$. Then the kernel of its derivative is $\ker(\mathrm{d}_x\psi) = \ker(\mathrm{d}_x\hat{\mu}) \cap T_x\mathcal{X} = 0$, so that $\mathrm{rk}\,(\mathrm{d}_x\psi) = \dim \mathcal{X}$. This implies x is a smooth point of $\mathcal{X}$. Hence, an open neighborhood of x is a $\Bbbk$-manifold. Since $\mathrm{rk}\,(\mathrm{d}_x\psi) = \dim \mathcal{X}$, the assumptions of the inverse function theorem for manifolds are satisfied. For $\Bbbk = \mathbb{R}$, the claim follows from [90, Theorem 4.5].

For $\Bbbk = \mathbb{C}$, the proof of [90, Theorem 4.5] applies by replacing "smooth manifold" with "complex manifold" [100, Chapter IV], "smooth map" with "holomorphic function" [100, Chapter I], "diffeomorphism" with "biholomorphism" [100, Chapter I], and using the inverse function theorem for holomorphic functions [100, Chapter I, Theorem 7.5] instead of the one for reals. $\qquad\square$

Finally, we need the theorem of birationality to a hypersurface. It is a strong result on generic identifiability that does not depend on the geometry of the variety.

Theorem 4.11 (Birationality to a hypersurface). *Let $\mathcal{X} \subseteq \mathbb{C}^n$ be a d-dimensional variety. Let $\hat{\mu} : \mathbb{C}^n \to \mathbb{C}^{d+1}$ be a generic projection. Then $\mu = \hat{\mu}|_\mathcal{X}$ is generically one-to-one, i.e., a generic point of the image of μ has a unique preimage in $\mathcal{X}$.*

Proof. We refer to [99, Theorem 1.8 and Remark 1.2]. $\qquad\square$

20

4.6. The Noether Normalization Lemma

The main results about the required number of measurements s for recoverability are consequences of the following fundamental theorem in commutative algebra.

Theorem 4.12 (Noether's Normalization Lemma)**.** *Let $\Bbbk$ be an infinite field and $\mathcal{X} \subset \Bbbk^n$ be a variety of dimension d and degree δ. Let $\hat{\mu} : \Bbbk^n \to \Bbbk^d$ be a generic affine linear projection and let $\mu = \hat{\mu}|_{\mathcal{X}}$ be its restriction to $\mathcal{X}$. Then $\mu : \mathcal{X} \to \Bbbk^d$ is a finite map. In particular, every fiber of μ is finite.*

Proof. The proof of this theorem is classic. We provide a sketch of the proof following [101, Chapter 10].

The usual algebraic formulation of Noether's normalization lemma is as follows. Let R be a finitely generated algebra over $\Bbbk$ with generators $x_1, \ldots, x_n$. Then, there exists an integer d and algebraically independent elements $r_1, \ldots, r_d \in R$ such that R is a finite module over $\Bbbk[r_1, \ldots, r_d]$. Moreover, $r_1, \ldots, r_d$ can be chosen to be linear combinations of the generators $x_1, \ldots, x_n$. This statement is proved in [101, Proposition 10.5].

The geometric statement is obtained by taking the affine coordinate ring of $\mathcal{X}$, i.e., $R = \Bbbk[\mathcal{X}] = \Bbbk[\Bbbk^n]/I(\mathcal{X})$, which is finitely generated with generators $x_1, \ldots, x_n$, the (equivalence classes modulo $I(\mathcal{X})$ of the) coordinates of $\Bbbk^n$. The algebraic statement guarantees that there exist linear combinations $r_1, \ldots, r_d$ of the x_j's such that $\Bbbk[\mathcal{X}]$ is a finitely generated module over $\Bbbk[r_1, \ldots, r_d]$. In fact, the proof [101, Proposition 10.5] and [101, Lemma 10.3] show that a generic choice of linear combinations satisfies the statement. The inclusion map $\Bbbk[r_1, \ldots, r_d] \to R$ defines an algebraic map $\mu : \mathcal{X} \to \Bbbk^d$ as $\Bbbk[r_1, \ldots, r_d]$ can be regarded as the coordinate ring of the affine space $\Bbbk^d$. The fact that the r_j's are linear combinations of the x_i's guarantees that μ is the restriction of a linear map $\hat{\mu} : \Bbbk^n \to \Bbbk^d$. $\qquad\square$

Noether's Normalization Lemma is sharp in the following sense:

1. For any affine linear projection $\hat{\mu} : \Bbbk^n \to \Bbbk^{d-1}$, the fiber of $\mu = \hat{\mu}|_{\mathcal{X}}$ over $y \in \Bbbk^{d-1}$ may be infinite. Further, if $\Bbbk$ is algebraically closed, then the generic fiber of μ is infinite.

2. It is possible that special projections $\hat{\mu} : \Bbbk^n \to \Bbbk^d$ map $\mathcal{X}$ to a proper subvariety of $\Bbbk^d$. In this case, the fiber of $\mu = \hat{\mu}|_{\mathcal{X}}$ at a generic point of $\Bbbk^d$ is empty and any fiber over the image of μ is infinite.

An example when case (2) holds is a projection of any product of (positive dimensional) varieties $\mathcal{X}_1 \times \mathcal{X}_2 \subset \Bbbk^{n_1} \times \Bbbk^{n_2} = \Bbbk^{n_1 + n_2}$ to $\Bbbk^{n_1}$.

Remark 4.13. The finiteness condition of the map $\mu : \mathcal{X} \to \Bbbk^d$ in the statement of the Noether normalization lemma has stronger implications than the fact that the fibers are finite. In fact, Theorem 4.12 implies that for every coordinate x_j of $\Bbbk^n$, there exists a fixed univariate polynomial $P_j \in \Bbbk[y_1, \ldots, y_d][x_j]$, (i.e., the coefficients of P_j are polynomials on $\Bbbk^d$ depending on $\mathcal{X}$ and on μ) with the property that for every $y \in \Bbbk^d$, the jth coordinates of the points of $\mu^{-1}(y)$ are the roots of $P_j(y, -) \in \Bbbk[x_j]$. Once we find the polynomials P_j, the jth coordinates of the points in the fiber of μ thus vary like the roots of P_j as $y \in \Bbbk^d$ varies. This provides an additional viewpoint on the continuity of the solutions of the inverse problem (IP), which we investigated in Section 3.4.

5. Proofs for the Q&A section

In this section, we give complete proofs of the correctness of the answers provided in Section 3. The proofs in this section exploit the fact that a generic point in $\mathcal{X}$ is a generic point in the complexification $\mathcal{X}^{\mathbb{C}}$ of $\mathcal{X}$, and that $d = \dim \mathcal{X}$ equals the complex dimension of $\mathcal{X}^{\mathbb{C}}$ (this is true irrespective of $\mathcal{X}$ being real or complex). For models given in implicit form (IS) this is implied by Theorem 4.3. For models in explicit form (ES) we can use Theorem 4.7.

5.1. The answer to Question 1

The restatement in terms of a system of polynomial equations is straightforward. The fact that any system of polynomial equations over $\mathbb{C}$ has a solution if and only if the ideal generated by the polynomials does not contain 1 is one of the forms of Hilbert's Nullstellensatz [47, Corollary 1.6]. Similarly, the statement over the reals can be found in [67, Corollary 4.1.8]. The fact that checking if an ideal contains 1 may be done using Gröbner bases follows from the definition of Gröbner basis. □

5.2. The answer to Question 2

If $s \geq d$, Theorem 4.12 guarantees that a generic projection has finite fibers. The inequality is sharp, both in the real and in the complex case. Indeed, suppose $s \leq d-1$ and distinguish two cases.

First, if $\mathbb{k} = \mathbb{C}$, by Theorem 4.7, there exists a Zariski open subset $U \subset \mu(\mathcal{X})$ such that for any $y \in U$ we have $\dim \mu^{-1}(\mu(y)) + \dim \mu(\mathcal{X}) = \dim \mathcal{X}$. In particular, since $\dim \mu(\mathcal{X}) \leq s < d$, we have $\dim \mu^{-1}(\mu(x)) > 0$. This implies that the generic fibers are infinite.

Second, if $\mathbb{k} = \mathbb{R}$, the map $\mu : \mathcal{X} \to \mathbb{R}^s$ is a morphism of varieties. Therefore, by Theorem 4.9, its generic fibers contain manifolds of dimension equal to $\dim \mathcal{X} - \dim(\operatorname{im}(\mu)) \geq \dim \mathcal{X} - s$. In particular, if $s < d$, generic fibers are infinite. □

5.3. The answer to Question 3

By Lemma 4.8 and Theorem 4.2, this reduces to Question 2. □

5.4. The answer to Question 4

The condition $s \geq d$ is necessary as in the answer to Question 2. We prove that a coordinate projection exists for $s \geq d$.

Assume that $s = d$. Let $x \in \mathcal{X}$ be a smooth point. Let $v_1, \ldots, v_d$ be a spanning set for $T_x\mathcal{X}$, the tangent space to $\mathcal{X}$ at x. Let M be the matrix whose rows are the vectors v_i. Hence, M is a $d \times n$ matrix with $\operatorname{rk}(M) = d$. In particular, M contains a $d \times d$ submatrix N with $\operatorname{rk}(N) = d$. This submatrix is obtained by selecting d columns of M, or, equivalently, d coordinates in $\mathbb{k}^n$. We claim that the projection $\hat{\mu}$ to these d coordinates is generically recoverable when restricted to $\mathcal{X}$.

Since $\hat{\mu}$ is linear, the differential $\mathrm{d}_x\hat{\mu}$ coincides with $\hat{\mu}$ itself. Now, its image $\operatorname{im}(\mathrm{d}_x\hat{\mu} : \mathbb{k}^n \to \mathbb{k}^d)$ restricted to $T_x\mathcal{X}$ is the column span of the matrix N, which coincides with the whole $\mathbb{k}^d$. Thus, $\mathrm{d}_x\hat{\mu}$ is of full rank and $T_x\mathcal{X} \cap \ker(\mathrm{d}_x\hat{\mu}) = 0$. By Corollary 4.10, there is a Euclidean open neighborhood $\tilde{U}$ of $x \in \mathcal{X}$ such that $\mu = \hat{\mu}|_\mathcal{X}$ is a diffeomorphism of $\tilde{U}$ onto its image.

We distinguish two subvarieties $\mathcal{S}_1, \mathcal{S}_2$ of $\mathcal{X}$. Let $\mathcal{S}_1$ be the variety of singular points of $\mathcal{X}$, which is a proper subvariety of $\mathcal{X}$. Let $\mathcal{S}_2$ be the set consisting of the points $x \in \mathcal{X}$ such that $T_x\mathcal{X} \cap \ker(\mathrm{d}_x\hat{\mu}) \neq 0$. This is a closed condition. Since $\tilde{U} \subseteq \mathcal{X}\backslash\mathcal{S}_2$, we deduce that $\mathcal{S}_2$ is a proper subvariety of $\mathcal{X}$. Let $U = \mathcal{X}\backslash(\mathcal{S}_1 \cup \mathcal{S}_2)$, which is Zariski open in $\mathcal{X}$.

We can conclude that the fiber over every point $y \in \mathbb{k}^d$ with $y \in \mu(U)$ is finite. Indeed, suppose that it has positive dimension. Then there exists an $x \in \mu^{-1}(y)$ that is smooth, $T_x\mathcal{X} \cap \ker(\mathrm{d}_x\hat{\mu}) = 0$, and every Euclidean neighborhood of x contains infinitely many points of $\mu^{-1}(y)$. But this contradicts Corollary 4.10. □

Remark 5.1. We note that a central point of the previous proof is a simple fact from linear algebra, that if we have a full rank r rectangular matrix, then we can choose a square submatrix of rank r. For a matroidal perspective on this proof we refer to [22, 102].

5.5. The answer to Question 5

In the implicit setting (IS) our model is given by $\mathcal{X} = \{x \in \mathbb{k}^n : f_1(x) = \cdots = f_k(x) = 0\}$. Let $x \in \mathcal{X}$ be a point satisfying the condition described in the answer. We note that this is an open condition, i.e., if it holds for some x, then there is a Zariski open set $U \subset \mathcal{X}$ such that the condition holds for all $x \in U$.

Let us consider the Zariski open set $\mathcal{X}^{\mathrm{sm}} \subset \mathcal{X}$ of smooth points of X. Taking $x \in U \cap \mathcal{X}^{\mathrm{sm}}$ we see that $\ker(\mathrm{d}_x\mu) \cap \ker(\mathrm{d}_xf) = 0$ implies that the restriction of $\mathrm{d}_x\mu$ to $T_x\mathcal{X}$ is injective. By Corollary 4.10, we deduce

that μ is injective in a Euclidean neighborhood of x. The same argument as in the answer to Question 4 shows that μ is generically recoverable.

In the explicit setting (ES), let $\mathcal{X} = \overline{\mathrm{im}(\phi)}$ where $\phi : \Bbbk^\ell \to \Bbbk^n$ is a polynomial parameterization. The map $\mu \circ \phi$ parameterizes the image $\mu(\mathcal{X})$. If $t^* \in \Bbbk^\ell$ is a point with $\mathrm{rk}\left(\mathrm{d}_{t^*}(\mu \circ \phi)\right) = d$, then $\mathrm{rk}\left(\mathrm{d}_t(\mu \circ \phi)\right) \geq d$ for generic $t \in \Bbbk^\ell$ by semicontinuity. By Theorem 4.7, we obtain $\dim \mu(\mathcal{X}) \geq d$. Since $\dim \mathcal{X} = d$, $\mu : \mathcal{X} \to \Bbbk^s$ maps $\mathcal{X}$ onto a (possibly semi-algebraic) set of the same dimension. By Theorem 4.9, the generic fiber of μ is 0-dimensional, so that μ is generically recoverable. $\qquad\square$

5.6. The answer to Question 6

The answer to Question 5 entails that the coordinate projection should satisfy the condition that the differential of $\mu \circ \phi$ has rank d at some point $t \in \Bbbk^\ell$. The matrix $\left[\frac{\partial \phi_{i_p}}{\partial t_j}\right]_{\substack{p=1,\ldots,s \\ j=1,\ldots,\ell}}$ is precisely the matrix representation of this differential. $\qquad\square$

5.7. The answer to Question 7

For any field, if $\mathcal{X}$ is an affine space then a generic projection to $\Bbbk^d$ is identifiable everywhere.

Consider the case $\Bbbk = \mathbb{C}$. For $s \geq d+1$ the answer follows from Theorem 4.11. Assume now that a generic projection $\hat{\mu}$ from $\mathbb{C}^n$ to $\mathbb{C}^d$ is generically identifiable. Let $\mu = \hat{\mu}|_{\mathcal{X}}$. If $\dim \mu(\mathcal{X}) < d$ then, as in the proof for Question 2, the generic fibers would be of positive dimension. Hence, we may assume that $\mu(\mathcal{X})$ is dense in $\mathbb{C}^d$ by Theorem 4.4. The kernel of $\hat{\mu}$ is a generic linear subspace in $\mathbb{C}^n \supset \mathcal{X}$. Hence, the fiber of $\hat{\mu}$ over a generic point of $\mu(\mathcal{X})$ is a generic affine subspace of $\mathbb{C}^n$ of codimension d. Thus, it must intersect $\mathcal{X}$ in δ distinct points, where δ is the degree of $\mathcal{X}$. Since μ is generically identifiable, this means that $\mathcal{X}$ must have degree one, i.e., it is an affine subspace (see the discussion at the end of Section 4.3).

For $\Bbbk = \mathbb{R}$ and $s \geq d+1$ we may take the complexification of the variety $\mathcal{X}$ and see that a generic complex, hence, a fortiori real, fiber is unique. To see that $s = d$ measurements do not suffice, we may take $\mathcal{X}$ to be a sphere. Generic points in the image $\mu(\mathcal{X})$ of a generic projection of the sphere have 2 elements in the preimage.

If $s < d$, as in the proof for Question 2, generic fibers have positive dimension. $\qquad\square$

5.8. The answer to Question 8

As in the proof for the answer to Question 3 we conclude by applying Lemma 4.8.

We note that for $d > 0$ and $p > 1$ the Veronese map strictly increases the degree, thus we may exclude the case of the affine subspace in Question 7. $\qquad\square$

5.9. The answer to Question 9

Let $\hat{\mu} : \Bbbk^n \to \Bbbk^s$ be the linear map that restricts to μ on $\mathcal{X}$. For distinct $x_1, x_2 \in \mathcal{X}$ we have that $\hat{\mu}(x_1) = \hat{\mu}(x_2)$ is equivalent to

$$\hat{\mu}(x_1 - x_2) = 0 \quad \Leftrightarrow \quad \exists p \in \Delta(\mathcal{X}) \setminus \{0\} : \hat{\mu}(p) = 0 \quad \Leftrightarrow \quad \ker(\hat{\mu}) \cap \Delta(\mathcal{X}) \neq 0.$$

Thus, the identifiability property depends only on the kernel of $\hat{\mu}$. The kernel of a generic $\hat{\mu}$ is a generic linear subspace of codimension s. A generic linear subspace of codimension s intersects $\Delta(\mathcal{X})$ only in zero if $s > \dim \Delta(\mathcal{X})$ (cf. Section 4.3), which proves the first equality in the answer. As $\Delta(\mathcal{X})$ may be realized as the image of $\mathcal{X} \times \mathcal{X}$ we must necessarily have the inequality $\dim \Delta(\mathcal{X}) \leq 2d$. $\qquad\square$

5.10. The answer to Question 10

Since Zariski open subsets are also Euclidean open, the answer to Question 9 concludes the proof. $\qquad\square$

5.11. The answer to Question 11

If $s < d$, then the answer to Question 2 shows that μ cannot be a local diffeomorphism, because it is not recoverable generically.

Let $s \geq d$. We denote the set of singular points in $\mu(\mathcal{X})$ by $\mathcal{S}$, so that $\mu(\mathcal{X}) \setminus \mathcal{S}$ is a smooth embedded submanifold of $\Bbbk^s$ of dimension equal to the dimension of $\mu(\mathcal{X})$.

The preimage $\mu^{-1}(\mathcal{S}) \cap \mathcal{X}$ is a proper subvariety of $\mathcal{X}$. As in the answer to Question 4, we consider two additional subvarieties of $\mathcal{X}$. The first $\mathcal{S}_1$ is the set of singular points, which is a proper subvariety. The second one $\mathcal{S}_2$ is the set of points $x \in \mathcal{X}$ such that $(T_x\mathcal{X}) \cap \ker(\mathrm{d}_x\hat{\mu}) \neq 0$. Since $\dim \ker(\hat{\mu}) = n - s \leq n - d$, the variety $\mathcal{S}_2$ is a proper subvariety. Then, we define $\mathcal{X}' := \mathcal{X} \setminus (\mu^{-1}(\mathcal{S}) \cup \mathcal{S}_1 \cup \mathcal{S}_2)$. This is a Zariski open subset of $\mathcal{X}$ contained in the smooth locus of $\mathcal{X}$. Hence, $\mathcal{X}'$ is a smooth embedded submanifold of $\Bbbk^s$. By construction, $\mu(\mathcal{S}_1) \cup \mu(\mathcal{S}_2)$ is of lower dimension in $\mu(\mathcal{X})$. Therefore, it is also of lower dimension in the manifold $\mu(\mathcal{X}) \setminus \mathcal{S}$. This implies that $\mu(\mathcal{X}')$ is an open subset of a smooth embedded submanifold of $\Bbbk^s$, hence it is itself a smooth embedded submanifold of $\Bbbk^s$.

The derivative $\mathrm{d}_x(\mu|_{\mathcal{X}'})$ is an isomorphism of the tangent spaces $T_x\mathcal{X}'$ and $T_{\mu(x)}\mu(\mathcal{X}')$. By the inverse function theorem for manifolds, $\mu|_{\mathcal{X}'}$ is a local diffeomorphism onto its image. $\square$

Remark 5.2. Note, however, in general, for $s = d$ there will be no Zariski open neighbourhoods that would give isomorphic algebraic varieties. This means that the inverse map will be analytic but not algebraic.

5.12. The answer to Question 12

For $s \geq d + 1$, the answer to Question 7 implies that μ is identifiable for generic $x \in \mathcal{X}$. Let $\mathcal{S}, \mathcal{S}_1, \mathcal{S}_2$ be as in the answer to Question 11 and let $\mathcal{S}_3 \subset \mathcal{X}$ be subvariety of points where μ is not identifiable. Let $\mathcal{X}' := \mathcal{X} \setminus (\mu^{-1}(\mathcal{S}) \cup \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3)$. As in the answer to Question 11 we have that $\mathcal{X}'$ and $\mu(\mathcal{X}')$ are smooth embedded submanifolds of $\Bbbk^s$ and that $\mu|_{\mathcal{X}'}$ is a local diffeomorphism. Since $\mu|_{\mathcal{X}'}$ is also injective, it is a global diffeomorphism.

It remains to prove that if $\Bbbk = \mathbb{C}$, $s = d$ and $\mathcal{X}$ is not an affine space then we do not have a global diffeomorphism. This follows from the answer to Question 7. $\square$

5.13. The answer to Question 13

Recall the definitions of the singular locus $\mathcal{S}_1 \subset \mathcal{X}$ and the set of points $\mathcal{S}_2 \subset \mathcal{X}$ for which $(T_x\mathcal{X}) \cap \ker(\mathrm{d}_x\hat{\mu}) \neq 0$ from the answer to Question 11. The smooth locus is $\mathcal{X}^{\mathrm{sm}} := \mathcal{X} \setminus \mathcal{S}_1$. We also set $\mathcal{X}' := \mathcal{X}^{\mathrm{sm}} \setminus \mathcal{S}_2$. As in the answer to Question 11, both $\mathcal{X}^{\mathrm{sm}}$ and $\mathcal{X}'$ are Zariski open submanifolds of $\mathcal{X}$. We show the asserted characterization of the condition number for smooth points $x \in \mathcal{X}^{\mathrm{sm}}$. We distinguish two cases.

The first case is when $x \in \mathcal{X}'$. In this case we have $(T_x\mathcal{X}) \cap \ker(\mathrm{d}_x\hat{\mu}) = 0$, which implies that the derivative $\mathrm{d}_x\mu$ is of full rank. This means that $\mu|_{\mathcal{X}'} : \mathcal{X}' \to \Bbbk^s$ is an immersion. By [90, Theorem 4.25] (for $\Bbbk = \mathbb{C}$ the same argument works), there exists an open neighbourhood $V_x \subset \mathcal{X}'$ of x such that $\mu(V_x)$ is an embedded submanifold of $\Bbbk^s$ [90, Section 5]. Hence, there exists a local inverse map $\mu_x^{-1} : \mu(V_x) \to V_x$. Let $y = \mu(x)$. The condition number at x is given by (CN):

$$\kappa(x) = \lim_{\epsilon \to 0} \sup_{y' \in \mu(V_x), \|y - y'\| \leq \epsilon} \frac{\|\mu_x^{-1}(y) - \mu_x^{-1}(y')\|}{\|y - y'\|},$$

where $\|\cdot\|$ is the Euclidean norm on both $\Bbbk^n$ and $\Bbbk^s$. Since μ_x^{-1} is smooth (holomorphic for $\Bbbk = \mathbb{C}$) and $\mu(V_x)$ is an embedded manifold, we can use [57, Theorems 3 and 4] to obtain

$$\kappa(x) = \max_{u \in T_y\mu(V_x), \|u\|=1} \|(\mathrm{d}_y\mu_x^{-1})\, u\| = \sigma_1(\mathrm{d}_y\mu_x^{-1}),$$

where σ_1 denotes the largest singular value. Because of the inverse function theorem [90, Theorem 4.5] (or its complex analogue), $\mathrm{d}_y\mu_x^{-1} = (\mathrm{d}_x\mu)^{-1}$. Hence, we have $\kappa(x) = \frac{1}{\sigma_d(\mathrm{d}_x\mu)}$ because $\dim V_x = \dim \mu(V_x) = d$ by the answer to Question 11. The reformulation in terms of matrices is immediate.

The second case is when $x \in \mathcal{X}^{\mathrm{sm}} \setminus \mathcal{X}'$ so that $(T_x\mathcal{X}) \cap \ker(\mathrm{d}_x\mu) \neq 0$. Then, $\mathrm{d}_x\mu$ has a nontrivial kernel, which implies that there is a smooth (holomorphic for $\Bbbk = \mathbb{C}$) curve $\gamma(t) \subset \mathcal{X}^{\mathrm{sm}}$ with $\gamma(0) = x$ and $\frac{\mathrm{d}}{\mathrm{d}t}\mu(\gamma(t))|_{t=0} = 0$. We can bound the condition number as $\kappa(x) \geq \lim_{t \to 0} \frac{\|x - \gamma(t)\|}{\|y - \mu(\gamma(t))\|} = \infty$. On the other hand, $(T_x\mathcal{X}) \cap \ker \mathrm{d}_x\mu \neq 0$ implies $\sigma_d(\mathrm{d}_x\mu) = 0$. Therefore, $\kappa(x) = \frac{1}{\sigma_d(\mathrm{d}_x\mu)}$ also in this case. $\square$

24

6. A numerical experiment

We confirm Theorem 1.1 experimentally in a few examples.[5] We showcase numerical homotopy continuation methods from HomotopyContinuation.jl [73] to accurately solve small algebraic compressed sensing problem instances. Recall that homotopy methods are numerical techniques for computing all isolated solutions of a system of polynomial equations. Additional experiments confirming Theorem 1.1 with alternative techniques were independently conducted by Krämer [104].

The structure we consider is that of tensors with a low-rank tensor rank decomposition [25]. Specifically, let $\mathcal{T}_2$ be the set of tensors in $\mathbb{R}^4 \otimes \mathbb{R}^3 \otimes \mathbb{R}^2 \simeq \mathbb{R}^{4 \times 3 \times 2}$ of rank bounded by 2. This set $\mathcal{T}_2$ can be parameterized explicitly as in (ES) with

$$\phi'(a_1', a_2', b_1', b_2', c_1', c_2') = a_1' \otimes b_1' \otimes c_1' + a_2' \otimes b_2' \otimes c_2'$$
$$= [(a_1')_i (b_1')_j (c_1')_k]_{i,j,k=1}^{4,3,2} + [(a_2')_i (b_2')_j (c_2')_k]_{i,j,k=1}^{4,3,2},$$

where $a_1', a_2' \in \mathbb{R}^4$, $b_1', b_2' \in \mathbb{R}^3$, and $c_1', c_2' \in \mathbb{R}^2$.

We want to compute isolated solutions of the polynomial system in Theorem 1.1 with numerical homotopy continuation methods to find alternative decompositions into rank-1 terms, depending on the measurement map μ. This requires a parameterization ϕ from a Euclidean space of dimension $d = \dim \mathcal{T}_2$ into $\mathcal{T}_2$. However, it is known that this is not always possible. In general $\mathcal{T}_2$ needs to be covered with multiple charts $\mathbb{R}^d \to \mathcal{T}_2$.

The dimension of $\mathcal{T}_2$ is known to be $d = 2(4 + 3 + 2 - 2) = 14$ [105] but the domain of ϕ' has dimension $2(4 + 3 + 2) = 18$, so clearly ϕ' overparameterizes $\mathcal{T}_2$. We restrict ourselves to parameterizing the Zariski open subset $\mathcal{X}$ of $\mathcal{T}_2$ given by the image of

$$\phi(a_1, a_2, b_1, b_2, c_1, c_2) = \begin{bmatrix} a_1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} b_1 \\ 1 \end{bmatrix} \otimes c_1 + \begin{bmatrix} a_2 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} b_2 \\ 1 \end{bmatrix} \otimes c_2,$$

where now $a_1, a_2 \in \mathbb{R}^3$, $b_1, b_2 \in \mathbb{R}^2$, and $c_1, c_2 \in \mathbb{R}^2$. This $\mathcal{X}$ is the Zariski open subset of tensors of rank bounded by 2 that have at least one decomposition where all vectors in the first and second factors $\mathbb{R}^4$ and $\mathbb{R}^3$ are nonzero in the last element.

Given a measurement map $\hat{\mu} : \mathbb{R}^{4 \times 3 \times 2} \to \mathbb{R}^s$ and a measurement $y = \mu(T)$ of a rank-2 tensor $T \in \mathcal{X}$, we can construct the polynomial system from Theorem 1.1,

$$y - \mu(\phi(a_1, a_2, b_1, b_2, c_1, c_2)) = 0, \tag{PS}$$

and solve for the variables $a_1, a_2, b_1, b_2, c_1, c_2$.

We investigate the interesting case of random coordinate projections, as the behavior for generic linear projections is completely described in the Q&A section. The minimum number of measurements for generic recoverability is $s = d$ by Theorem 1.1. So let us start with this many measurements. We consider the following incomplete $4 \times 3 \times 2$ tensor of rank 2 (the fact that it has rank at most 2 will be implied by Theorem 1.1 after we have solved the system (PS) and found a solution):

$$y = \left[\begin{array}{ccc|ccc} -32 & ? & -24 & ? & 10 & ? \\ 72 & ? & ? & -57 & 27 & ? \\ -104 & 40 & ? & -11 & 1 & ? \\ ? & 16 & 0 & -1 & ? & -7 \end{array} \right],$$

where the left submatrix is the first 4×3 slice of $y \in \mathbb{R}^{4 \times 3 \times 2}$ and the right submatrix is the second slice. Solving the system (PS), we find 4 isolated, nonsingular, real solutions $a_1^i, a_2^i, b_1^i, b_2^i, c_1^i, c_2^i$, $i = 1, \ldots, 4$. One can show that the parameterization ϕ is generically 2-to-1: swapping the variables $a_1 \leftrightarrow a_2$, $b_1 \leftrightarrow b_2$, and

$c_1 \leftrightarrow c_2$ corresponds to summing the rank-1 terms in a different order, which produces the same result. Consequently, only 2 of these 4 solutions represent distinct points of $\mathcal{X} \subset \mathcal{T}_2$. These distinct completed tensors, with imputed entries highlighted and rounded to 3 decimals, are

$$
T_1 = \left[\begin{array}{ccc|ccc}
-32 & \mathbf{8} & -24 & \mathbf{-26} & 10 & \mathbf{-2} \\
72 & \mathbf{-40} & \mathbf{-5} & -57 & 27 & \mathbf{21} \\
-104 & 40 & \mathbf{-8} & -11 & 1 & \mathbf{-17} \\
-40 & 16 & 0 & -1 & \mathbf{-1} & -7
\end{array}\right], \text{ and}
$$

$$
T_2 = \left[\begin{array}{ccc|ccc}
-32 & \mathbf{411.518} & -24 & \mathbf{-26} & 10 & \mathbf{-182.054} \\
72 & \mathbf{1112.908} & \mathbf{1075.769} & -57 & 27 & \mathbf{-396.574} \\
-104 & 40 & \mathbf{-728.216} & -11 & 1 & \mathbf{-78.642} \\
\mathbf{-1.114} & 16 & 0 & -1 & \mathbf{0.389} & -7
\end{array}\right],
$$

and their (unique) decompositions into rank-1 terms, rounded to 3 decimals, are

$$
T_1 = \begin{bmatrix} -1 \\ -6 \\ 2 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 8 \\ -4 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \\ 3 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} -8 \\ -3 \end{bmatrix}, \text{ and}
$$

$$
T_2 = \begin{bmatrix} 25.925 \\ 60.344 \\ 8.736 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0.139 \\ -0.001 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} -116.229 \\ -9.791 \end{bmatrix} + \begin{bmatrix} 25.719 \\ 69.6 \\ 2.471 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0.129 \\ 0.137 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 116.229 \\ 2.791 \end{bmatrix}.
$$

We verified that the differential of $\mu \circ \phi$ at all 4 solutions for the variables $a_1, a_2, b_1, b_2, c_1, c_2$ was of (numerical) rank equal to $d = 14 = \dim \mathcal{T}_2$. Hence, these solutions are locally continuous by the inverse function theorem. We computed the (absolute) condition number of the inverse problem (IP) at the solution T_1, respectively T_2, and found approximately 11.642, respectively 1907.954.

It is interesting to observe that both T_1 and T_2 coincide, up to 14 significant digits, in the unobserved coordinate $(1, 1, 2)$ with value -26. This implies that augmenting μ with a projection to this coordinate (so it projects to $s = d + 1$ coordinates) will not eliminate either completed tensor, even though this number of measurements generically suffices for generic identifiability by Theorem 1.1! On the other hand, additionally projecting to one of the other coordinates eliminates either T_1 or T_2, so a unique solution remains with $s = d + 1$ even for such a coordinate projection (which are not generic linear maps in the sense of Theorem 1.1).

7. Conclusions

We investigated the class of algebraic compressed sensing problems where linear or polynomial measurements $\mu : \mathcal{X} \to \Bbbk^s$ are made of a model $\mathcal{X}$, defined implicitly or explicitly by polynomials. Our focus was on theoretical questions of existence, recoverability, identifiability, and continuity. We showed how standard results from classic algebraic geometry and differential geometry can be applied to give almost complete answers to these questions in the case of *generic* algebraic compressed sensing problem instances. The numerical experiments illustrated that all reconstructions in small-scale problem instances without noise can be found using numerical homotopy continuation methods.

We hope that the present work can be a starting point for further study of algebraic compressed sensing problems. There are several exciting challenges and directions one may follow. Indeed, we already mentioned (i) proving that a map $\mu : \mathcal{X} \to \Bbbk^s$ given by s polynomials $f_1, \ldots, f_s$, where each f_i is generic of degree a_i, is everywhere recoverable if and only if $s \geq d$; (ii) deciding if every smooth, complex projective curve may be injected by a morphism of algebraic varieties to a projective plane; and (iii) providing conditions when there exists a generically identifiable coordinate projection into a $(d + 1)$-dimensional space.

An interesting observation is that one can regard the set of coordinate projections as a 0-dimensional reducible variety in the space of linear forms. This naturally leads to the following question: what are the algebraic varieties $\mathcal{Y}$ in the space of linear forms with the property that for every $\mathcal{X}$, $\dim \mathcal{X} + 1$ many (generic) measurements from $\mathcal{Y}$ are sufficient to obtain identifiability? We saw that the entire space of linear forms satisfies this condition and that the set of coordinate projections does not. We conjecture that a sufficient condition is that $\mathcal{Y}$ is irreducible, not linearly degenerate, and of positive dimension.

Alternative characterizations or sufficient conditions that certify existence, recoverability, and identifiability of *specific* algebraic compressed sensing problems is another area that merits further study. This likely requires a toolset that extends beyond the one we used here. The case of coordinate projections of low-rank matrices, for example, has been investigated using rigidity theory [18, 106], algebraic combinatorics [19], and matroid theory [21, 22].

Lastly, we focused on theoretical questions of algebraic compressed sensing problems, showing that many generic problem instances are well posed even with very few measurements. We did not investigate how these problems should be solved algorithmically in practice, where one must take into account measurement errors and the propagation of numerical roundoff errors. It would be interesting to investigate whether numerical homotopy continuation, or other numerical methods for solving systems of polynomial equations, can stably solve algebraic compressed sensing problems, not just for generic instances, but for all instances.

Acknowledgements

References

[1] M. F. Duarte, Y. C. Eldar, Structured compressed sensing: From theory to applications, IEEE Trans. Signal Process. 59 (9) (2011) 4053–4085. doi:10.1109/TSP.2011.2161982.

[2] A. Kirsch, An Introduction to the Mathematical Theory of Inverse Problems, 2nd Edition, Springer, 2011.

[3] D. L. Donoho, Compressed sensing, IEEE Trans. Inform. Theory 52 (4) (2006) 1289–1306. doi:10.1109/TIT.2006.871582.

[4] E. J. Candès, T. Tao, Decoding by linear programming, IEEE Trans. Inform. Theory 51 (12) (2005) 4203–4215. doi:10.1109/TIT.2005.858979.

[5] E. J. Candès, T. Tao, Near-optimal signal recovery from random projections: Universal encoding strategies?, IEEE Trans. Inform. Theory 52 (12) (2006) 5406–5425. doi:10.1109/TIT.2006.885507.

[6] E. J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, IEEE Trans. Inform. Theory 52 (2) (2006) 489–509. doi:10.1109/TIT.2005.862083.

[7] E. J. Candès, J. K. Romberg, T. Tao, Stable signal recovery from incomplete and inaccurate measurements, Comm. Pure Appl. Math. 59 (8) (2006) 1207–1223. doi:10.1002/cpa.20124.

[8] R. G. Baraniuk, M. B. Wakin, Random projections of smooth manifolds, Found. Comput. Math. 9 (2009) 51–77. doi:10.1007/s10208-007-9011-z.

[9] R. G. Baraniuk, V. Cevher, M. B. Wakin, Low-dimensional models for dimensionality reduction and signal recovery: A geometric perspective, Proc. IEEE 98 (6) (2010) 959–971. doi:10.1109/JPROC.2009.2038076.

[10] S. Foucart, H. Rauhut, A Mathematical Introduction to Compressive Sensing, Birkhäuser, New York, NY, 2013. doi:10.1007/978-0-8176-4948-7.

[11] P. Breiding, T. O. Çelik, T. Duff, A. Heaton, A. Maraj, A.-L. Sattelberger, L. Venturello, O. Yürük, Nonlinear algebra and applications, Numer. Algebra Control Optim. 13 (1) (2023) 81–116. doi:10.3934/naco.2021045.

[12] M. Fazel, E. J. Candes, B. Recht, P. Parrilo, Compressed sensing and robust recovery of low rank matrices, in: 2008 42nd Asilomar Conference on Signals, Systems and Computers, 2008, pp. 1043–1047. doi:10.1109/ACSSC.2008.5074571.

[13] L. T. Nguyen, J. Kim, B. Shim, Low-rank matrix completion: A contemporary survey, IEEE Access 7 (2019) 94215–94237. doi:10.1109/ACCESS.2019.2928130.

[14] M. Golbabaee, P. Vandergheynst, Hyperspectral image compressed sensing via low-rank and joint-sparse matrix recovery, in: 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2012, pp. 2741–2744. doi:10.1109/ICASSP.2012.6288484.

[15] V. Chandrasekaran, S. Sanghavi, P. A. Parrilo, A. S. Willsky, Sparse and low-rank matrix decompositions, IFAC Proceedings Volumes 42 (10) (2009) 1493–1498, 15th IFAC Symposium on System Identification. doi:10.3182/20090706-3-FR-2004.00249.

[16] J. Tanner, S. Vary, Compressed sensing of low-rank plus sparse matrices, arXiv:2007.09457 (2020).

[17] M. Bebendorf, S. Rjasanow, Matrix Compression for the Radiation Heat Transfer in Exhaust Pipes, 2000, pp. 183–192. doi:10.1007/978-3-662-04015-7_20.

[18] A. Singer, M. Cucuringu, Uniqueness of low-rank matrix completion by rigidity theory, SIAM J. Matrix Anal. Appl. 31 (4) (2010) 1621–1641. doi:10.1137/090750688.

[19] F. J. Király, L. Theran, R. Tomioka, The algebraic combinatorial approach for low-rank matrix completion, J. Mach. Learn. Res. 16 (2015) 1391–1436.

[20] F. Gesmundo, J. Hauenstein, C. Ikenmeyer, J. M. Landsberg, Complexity of Linear Circuits and Geometry, Found. Comp. Math. 16 (2016) 599–635.

[21] M. C. Tsakiris, Low-rank matrix completion theory via Plücker coordinates, arXiv:2004.12430 (2021).

[22] M. C. Tsakiris, Results on the algebraic matroid of the determinantal variety, arXiv:2002.058082v6 (2021).

[23] T. G. Kolda, B. W. Bader, Tensor decompositions and applications, SIAM Rev. 51 (3) (2009) 455–500. doi:10.1137/07070111x.

[24] L. Grasedyck, D. Kressner, C. Tobler, A literature survey of low-rank tensor approximation techniques, GAMM Mitteilungen 36 (1) (2013) 53–78. doi:10.1002/gamm.201310004.

[25] F. L. Hitchcock, The expression of a tensor or a polyadic as a sum of products, J. Math. Phys. 6 (1) (1927) 164–189. doi:10.1002/sapm192761164.

[26] L. R. Tucker, Some mathematical notes on three-mode factor analysis, Psychometrika 31 (3) (1966) 279–311. doi:10.1007/bf02289464.

[27] G. Vidal, Efficient classical simulation of slightly entangled quantum computations, Phys. Rev. Letters 91 (14) (2003) 147902. doi:10.1103/physrevlett.91.147902.

[28] I. V. Oseledets, Tensor-train decomposition, SIAM J. Sci. Comput. 33 (5) (2011) 2295–2317. doi:10.1137/090752286.

[29] W. Hackbusch, S. Kühn, A new scheme for the tensor representation, J. Fourier Anal. Appl. 15 (5) (2009) 706–722. doi:10.1007/s00041-009-9094-9.

[30] L. Grasedyck, Hierarchical singular value decomposition of tensors, SIAM J. Matrix Anal. Appl. 31 (4) (2010) 2029–2054. doi:10.1137/090764189.

[31] K. Ye, L.-H. Lim, Tensor network ranks, arXiv:1801.02662 (2018).

[32] E. E. Papalexakis, C. Faloutsos, N. D. Sidiropoulos, Tensors for data mining and data fusion: Models, applications, and scalable algorithms, ACM Trans. Intell. Syst. Technol. 8 (2) (2017) 16:1–16:44. doi:10.1145/2915921.

[33] N. D. Sidiropoulos, A. Kyrillidis, Multi-way compressed sensing for sparse low-rank tensors, IEEE Signal Process. Letters 19 (11) (2012) 757–760. doi:10.1109/LSP.2012.2210872.

[34] H. Rauhut, Z. Stojanac, Tensor theta norms and low rank recovery, Numer. Algorithms (2020). doi:10.1007/s11075-020-01029-x.

[35] H. Rauhut, R. Schneider, Z. Stojanac, Low rank tensor recovery via iterative hard thresholding, Linear Algebra Appl. 523 (2017) 220–262. doi:10.1016/j.laa.2017.02.028.

[36] S. Sullivant, Algebraic Statistics, Vol. 194 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2018.

[37] J. E. Kolassa, Series Approximation Methods in Statistics, 3rd Edition, Springer, 2006.

[38] K. Pearson, Contributions to the mathematical theory of evolution, Phil. Trans. Roy. Soc. London A 185 (1894) 71–110.

[39] D. Agostini, C. Améndola, K. Ranestad, Moment identifiability of homoscedastic Gaussian mixtures, Found. Comput. Math. 21 (2021) 695–724. doi:10.1007/s10208-020-09469-6.

[40] C. Améndola, J. C. Faugères, B. Sturmfels, Moment varieties of Gaussian mixtures, J. Algebr. Stat. 7 (2016) 14–28. doi:10.18409/jas.v7i1.42.

[41] K. Kohn, B. Shapiro, B. Sturmfels, Moment varieties of measures on polytopes, Ann. Scuola Norm. Superiore Pisa 21 (2020) 739–770. doi:10.2422/2036-2145.201808_003.

[42] J. Lindberg, C. Améndola, J. I. Rodriguez, Estimating Gaussian mixtures using sparse polynomial moment systems, arXiv:2106.15675 (2021).

[43] B. Guo, J. Nie, Z. Yang, Learning diagonal gaussian mixture models and incomplete tensor decompositions, Viet. J. Math. 50 (2) (2021) 421–446. doi:10.1007/s10013-021-00534-3.

[44] R. Benedetti, J.-J. Risler, Real Algebraic and Semi-algebraic Sets, Actualités Mathématiques. [Current Mathematical Topics], Hermann, Paris, 1990.

[45] D. Cox, J. Little, D. O'Shea, Ideals, Varieties, and Algorithms, 4th Edition, Springer, 2015.

[46] J. Harris, Algebraic Geometry, A First Course, Vol. 133 of Graduate Text in Mathematics, Springer-Verlag, 1992.

[47] D. Eisenbud, Commutative Algebra: with a View Toward Algebraic Geometry, Vol. 150 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1995.

[48] P. Assouad, Plongements lipschitziens dans $\mathbb{R}^n$, Bull. Soc. Math. France 111 (1983) 429–448.

[49] W. B. Johnson, J. Lindenstrauss, Extensions of Lipschitz mappings into a Hilbert space, Contemp. Math. 26 (1984) 189–206. doi:10.1090/conm/026/737400.

[50] J. Luukkainen, J. Väisälä, Elements of Lipschitz topology, Ann. Acad. Sci. Fenn. Math. 3 (1977) 85–122.

doi:10.5186/aasfm.1977.0315.

[51] H. Movahedi-Lankerani, Minimal Lipschitz Embeddings, Ph.D. thesis, Pennsylvania State University (1990).

[52] D. Burago, Y. Burago, S. Ivanov, A Course in Metric Geometry, Vol. 33 of Graduate Series in Mathematics, American Mathematical Society, Providence, Rhode Island, 2001.

[53] A. M. Tillmann, M. E. Pfetsch, The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing, IEEE Trans. Inform. Theory 60 (2) (2014) 1248–1259. doi:10.1109/TIT.2013.2290112.

[54] H. Whitney, Differentiable manifolds, Ann. Math. 37 (3) (1936) 645–680. doi:10.2307/1968482.

[55] L. Birbrair, A. Fernandes, Z. Jelonek, On the extension of bi-Lipschitz mappings, Sel. Math. New Ser. 27 (15) (2021). doi:10.1007/s00029-021-00625-6.

[56] L. N. Trefethen, D. Bau, Numerical Linear Algebra, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997. doi:10.1137/1.9780898719574.

[57] J. R. Rice, A theory of condition, SIAM J. Numer. Anal. 3 (2) (1966) 287–310. doi:10.1137/0703023.

[58] P. Bürgisser, F. Cucker, Condition: The Geometry of Numerical Algorithms, Springer, Heidelberg, 2013.

[59] L. Blum, F. Cucker, M. Shub, S. Smale, Complexity and Real Computation, Springer-Verlag, New York, 1998. doi:10.1007/978-1-4612-0701-6.

[60] N. J. Higham, Accuracy and Stability of Numerical Algorithms, SIAM, Philadelphia, PA, 1996.

[61] J. Humpherys, T. J. Jarvis, E. J. Evans, Foundations of Applied Mathematics, SIAM, Philadelphia, PA, 2017.

[62] D. Armentano, Stochastic perturbations and smooth condition numbers, J. Complexity 26 (2) (2010) 161–171. doi:10.1016/j.jco.2010.01.003.

[63] B. Sturmfels, What is . . . a Gröbner basis?, Notices Amer. Math. Soc. 52 (10) (2005).

[64] G. M. Greuel, G. Pfister, O. Bachmann, C. Lossen, H. Schönemann, A Singular Introduction to Commutative Algebra, Springer, 2002.

[65] B. Sturmfels, Gröbner Bases and Convex Polytopes, Vol. 8 of University Lecture Series, American Mathematical Society, 1996.

[66] B. Sturmfels, Solving Systems of Polynomial Equations, no. 97 in CBMS Regional Conferences Series, American Mathematical Society, 2002.

[67] J. Bochnak, M. Coste, M. Roy, Real Algebraic Geometry, Springer–Verlag, 1998.

[68] C. Harris, M. Michałek, E. C. Sertöz, Computing images of polynomial maps, Adv. Comput. Math. 45 (5) (2019) 2845–2865. doi:10.1007/s10444-019-09715-8.

[69] M. Barakat, M. Lange-Hegermann, An algorithmic approach to Chevalley's Theorem on images of rational morphisms between affine varieties, Math. Comput. 91 (333) (2021) 451–490. doi:10.1090/mcom/3632.

[70] A. S. Fraenkel, Y. Yesha, Complexity of solving algebraic equations, Inform. Process. Lett. 10 (4–5) (1980) 178–179. doi:10.1016/0020-0190(80)90134-9.

[71] C. Beltrán, A. Leykin, Certified numerical homotopy tracking, Exp. Math. 21 (1) (2012) 69–83. doi:10.1080/10586458.2011.606184.

[72] J. D. Hauenstein, F. Sottile, Algorithm 921: AlphaCertified: Certifying solutions to polynomial systems, ACM Trans. Math. Softw. 38 (4) (2012). doi:10.1145/2331130.2331136.

[73] P. Breiding, S. Timme, HomotopyContinuation.jl: A Package for Homotopy Continuation in Julia, in: International Congress on Mathematical Software, Springer, 2018, pp. 458–465.

[74] T. L. Lee, T. Y. Li, C. H. Tsai, HOM4PS-2.0: a software package for solving polynomial systems by the polyhedral homotopy continuation method, Computing 83 (2) (2008) 109. doi:10.1007/s00607-008-0015-6.

[75] D. J. Bates, J. D. Hauenstein, A. J. Sommese, C. W. Wampler, Bertini: Software for Numerical Algebraic Geometry, Available at bertini.nd.edu with permanent doi: dx.doi.org/10.7274/R0H41PB5.

[76] A. Leykin, Numerical Algebraic Geometry for Macaulay2, The Journal of Software for Algebra and Geometry: Macaulay2 3 (2011) 5–10.

[77] J. Verschelde, Algorithm 795: PHCpack: a general-purpose solver for polynomial systems by homotopy continuation, ACM Trans. Math. Softw. 25 (2) (1999) 251–276. doi:10.1145/317275.317286.

[78] S. Telen, M. Van Barel, J. Verschelde, A robust numerical path tracking algorithm for polynomial homotopy continuation, SIAM J. Sci. Comput. 42 (6) (2020) A3610–A3637. doi:10.1137/19m1288036.

[79] P. Lairez, A deterministic algorithm to compute approximate roots of polynomial systems in polynomial average time, Found. Comput. Math. 17 (5) (2017) 1265–1292. doi:10.1007/s10208-016-9319-7.

[80] P. Lairez, Rigid continuation paths I. Quasilinear average complexity for solving polynomial systems, J. AMS 33 (2020) 487–523. doi:10.1090/jams/938.

[81] L. Blum, M. Shub, S. Smale, On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines, Bull. Amer. Math. Soc. 21 (1) (1989) 1–46.

[82] M. Kabanava, R. Kueng, H. Rauhut, U. Terstiege, Stable low-rank matrix recovery via null space properties, Inf. Inference 5 (4) (2016) 405–441. doi:10.1093/imaiai/iaw014.

[83] F. Russo, On the geometry of some special projective varieties, Lecture Notes of the Unione Matematica Italiana, Springer, 2016.

[84] E. Ballico, E. Ventura, Injective linear series of algebraic curves on quadrics, Ann. Univ. Ferrara 66 (2) (2020) 231–254. doi:10.1007/s11565-020-00343-5.

[85] P. Görlach, Projective Geometry, Toric Algebra and Tropical Computations, Ph.D. thesis (2020).

[86] P. Görlach, Injection dimensions of projective varieties, arXiv:1905.11306 (2019).

[87] M. Michałek, Can local cohomology prevent injections?, Singularities and Homological Aspects of Commutative Algebra,

Oberwolfach Reports 16 (1) (2020) 407–464, edited by: A. Conca and S. D. Cutkosky and S. B. Iyengar.

[88] R. Piene, Cuspidal projections of space curves, Math. Ann. 256 (1) (1981) 95–119. doi:10.1007/bf01450947.

[89] M. Spivak, Calculus on Manifolds, Addison–Wesley, New York, 1965.

[90] J. M. Lee, Introduction to Smooth Manifolds, 2nd Edition, Springer, New York, USA, 2013.

[91] P.-A. Absil, R. Mahony, R. Sepulchre, Optimization Algorithms on Matrix Manifolds, Princeton University Press, 2008. doi:10.1515/9781400830244.

[92] P. Breiding, N. Vannieuwenhoven, The condition number of Riemannian approximation problems, SIAM J. Optim. (31) (2021) 1049–1077. doi:10.1137/20m1323527.

[93] P. Breiding, N. Vannieuwenhoven, Sensitivity of low-rank matrix recovery, Numerische Mathematik 152 (4) (2022) 725–759. doi:10.1007/s00211-022-01327-7.

[94] P. A. Griffiths, J. Harris, Principles of Algebraic Geometry, Wiley Classics Library, John Wiley & Sons Inc., New York, 1994, reprint of the 1978 original.

[95] D. Eisenbud, J. Harris, 3264 and All That: A Second Course in Algebraic Geometry, Cambridge University Press, 2016.

[96] R. Hartshorne, Algebraic Geometry, Vol. 52, Springer Science & Business Media, 2013.

[97] H. Whitney, Elementary structure of real algebraic varieties, Ann. Math. 66 (3) (1957). doi:10.2307/1969908.

[98] M. Michałek, B. Sturmfels, Invitation to Nonlinear Algebra, Vol. 211 of Graduate Studies in Mathematics, American Mathematical Society, 2021.

[99] I. R. Shafarevich, Basic Algebraic Geometry 1, Varieties in Projective Space, 3rd Edition, Springer-Verlag, 2013.

[100] K. Fritzsche, H. Grauert, From Holomorphic Functions to Complex Manifolds, Vol. 213 of Graduate Texts in Mathematics, Springer-Verlag, New York, 2002.

[101] A. Gathmann, Commutative Algebra, class Notes – TU Kaiserslautern 2013/2014 (2014).

[102] Z. Rosen, J. Sidman, L. Theran, Algebraic matroids in action, The American Mathematical Monthly 127 (3) (2020) 199–216.

[103] P. Breiding, F. Gesmundo, M. Michałek, N. Vannieuwenhoven, Low rank tensor recovery, `https://www.JuliaHomotopyContinuation.org/examples/tensor-recovery/`.

[104] S. Krämer, Asymptotic log-det sum-of-ranks minimization via tensor (alternating) iteratively reweighted least squares, arXiv:2106.15201 (2021).

[105] H. Abo, G. Ottaviani, C. Peterson, Induction for secant varieties of Segre varieties, Trans. Amer. Math. Soc. 361 (2009) 767–792.

[106] J. Cruickshank, F. Mohammadi, A. Nixon, S.-I. Tanigawa, Rigidity of hypergraphs under algebraic constraints, in preparation (2023).