J. Eur. Math. Soc. (Online first) DOI 10.4171/JEMS/1082

Anton Bernshteyn · Michelle Delcourt · Anush Tserunyan

Independent sets in algebraic hypergraphs

Received September 14, 2018

Abstract. In this paper we study hypergraphs definable in an algebraically closed field. Our goal is to show, in the spirit of the so-called transference principles in extremal combinatorics, that if a given algebraic hypergraph is "dense" in a certain sense, then a generic low-dimensional subset of its vertices induces a subhypergraph that is also "dense." (For technical reasons, we only consider low-dimensional subsets that are parameterized by rational functions.) Our proof approach is inspired by the hypergraph containers method, developed by Balogh, Morris, and Samotij and independently by Saxton and Thomason (although adapting this method to the algebraic setting presents some unique challenges that do not occur when working with finite hypergraphs). Along the way, we establish a natural generalization of the classical dimension of fibers theorem in algebraic geometry, which is interesting in its own right.

Keywords. Algebraic hypergraph, independent set, containers method, rational parameterization, fiberwise definability, transference principle, dimension of fibers

1. Introduction

An active line of inquiry in combinatorics in recent years has been extending classical results to the so-called *sparse random setting*, where the goal is to show that certain known properties of "dense" combinatorial structures are inherited by their randomly chosen "sparse" substructures. A typical example of this is the *Sparse Szemerédi Theorem*, proved by Conlon and Gowers [7] and independently by Schacht [18]. Say that a finite set $S \subset \mathbb{N}$ is (ε, t)-*Szemerédi*, where $\varepsilon > 0$ and $t \in \mathbb{N}$, if every subset $A \subseteq S$ with $|A| \geqslant \varepsilon|S|$ contains a nonconstant arithmetic progression of length t. The following is a classical theorem of Szemerédi:

Theorem 1.1 (Szemerédi [20]). *For all $\varepsilon > 0$ and $t \in \mathbb{N}$, there exists $n_0 \in \mathbb{N}$ such that for each $n \geqslant n_0$, the set $\{1, \dots, n\}$ is (ε, t)-Szemerédi.*

Anton Bernshteyn: School of Mathematics, Georgia Institute of Technology, Atlanta, GA, USA; bahtoh@gatech.edu

Michelle Delcourt: Department of Mathematics, Ryerson University, Toronto, ON, Canada; mdelcourt@ryerson.ca

Anush Tserunyan: Mathematics and Statistics Department, McGill University, Montreal, QC, Canada; anush.tserunyan@mcgill.ca

Mathematics Subject Classification (2020): 05C65, 05C69, 14A10, 14A25, 05C35, 05C63

Conlon–Gowers and Schacht proved that the property of being (ε, t)-Szemerédi is inherited, with high probability, by a fairly sparse random subset of $\{1, \dots, n\}$:

Theorem 1.2 (Sparse Szemerédi; Conlon–Gowers [7], Schacht [18]). *For all $\varepsilon > 0$ and $t \in \mathbb{N}$, there exist $n_0 \in \mathbb{N}$ and $C > 0$ such that the following holds: For each $n \geqslant n_0$, fix some $p_n \in [Cn^{-\frac{1}{t-1}}, 1]$. Let S_n be the random subset of $\{1, \dots, n\}$ obtained by picking each element $i \in \{1, \dots, n\}$ independently with probability p_n. Then*

$$\lim_{n\to\infty} \mathbb{P}[S_n \text{ is } (\varepsilon, t)\text{-Szemerédi}] = 1.$$

Remark. The lower bound on p_n in Theorem 1.2 in sharp, up to the value of C.

For further examples see, e.g., the survey [6].

The remarkable success of this research program is largely due to the development of powerful general techniques for proving random analogs of combinatorial theorems. One of them is the so-called (*hypergraph*) *containers method*, introduced independently by Balogh, Morris, and Samotij [2] and Saxton and Thomason [17] and based on the previous work of Kleitman and Winston [12] and Sapozhenko [16]. A (*t-uniform*) *hypergraph* on a set V of *vertices* is a family E of t/element subsets of V, called the *edges* of E. A set $I \subseteq V$ of vertices is *E-independent* if it does not include any edge of E as a subset. The starting point of the containers method is the observation that problems in extremal combinatorics often involve independent sets in specific hypergraphs (for instance, Szemerédi's theorem is a statement about independent sets in the t-uniform hypergraph on $\{1, \dots, n\}$ whose edges are the t-term arithmetic progressions). Assuming that the edges of a given hypergraph E are "well-distributed" in a certain technical sense, the containers method puts an upper bound on the number of E-independent sets, and, furthermore, it does so in a very "explicit" manner. Namely, each independent set I gets assigned a *fingerprint* F that, in turn, encodes a *container* C, with the property that $F \subseteq I \subseteq C$, meaning that the total number of independent sets is at most the number of distinct fingerprints times the maximum number of independent sets in an individual container.

In this paper we combine the containers method with another recent trend in combinatorics: establishing versions of extremal results for (hyper)graphs that are definable in model/theoretically tame structures; see, e.g., [1, 5, 9, 19, 21] and the references therein for a sample of related work. We shall specifically focus on the case when the ambient structure is an algebraically closed field (although it would be interesting to know if our results could be interpreted and proved in some more general context).

Let $\mathbb{F}$ be an algebraically closed field and let E be a t-uniform hypergraph with vertex set $\mathbb{F}^n$. For convenience, we shall view the edges of E as *ordered* tuples of length t rather than simply t-element sets, so E is a subset of $(\mathbb{F}^n)^t \cong \mathbb{F}^{tn}$. Thus, it makes sense to ask whether E is a *definable* set (in the sense of first-order logic) in the field structure of $\mathbb{F}$. For instance, if E is the set of all t-term arithmetic progressions in $\mathbb{F}^n$, then E is definable, since $(x_1, \dots, x_t) \in E$ if and only if

$$(x_3 - x_2 = x_2 - x_1) \wedge \dots \wedge (x_t - x_{t-1} = x_{t-1} - x_{t-2}) \wedge (x_1 \neq x_2).$$

Given a definable hypergraph E on $\mathbb{F}^n$, we wish to study the properties of the sub-hypergraph of E induced by a "sparse random" *definable* set $X \subseteq \mathbb{F}^n$. It is fairly clear

what the word "sparse" should mean in this context: Since we are working in an algebraically closed field, there is a notion of *dimension* of definable sets, so "sparse" stands for "low-dimensional." It is somewhat less obvious how to interpret the word "random" correctly. The approach we take in this paper is to consider only those subsets of $\mathbb{F}^n$ that are *parameterized* by rational functions. Each such subset is described via a finite tuple of elements of $\mathbb{F}$ – namely the coefficients of the parameterizing polynomials – and hence can be encoded as a vector in the affine space $\mathbb{F}^N$ for some $N \in \mathbb{N}$. This enables us to talk about a subset $X \subseteq \mathbb{F}^n$ whose parameterization is given by *generic* $x \in \mathbb{F}^N$, and that will be our notion of "randomness." (For details, see Section 2.2.)

The motivating question can now be stated in reasonably precise terms:

Question 1.3. Suppose $E \subseteq (\mathbb{F}^n)^t$ is a definable t-uniform hypergraph. What can be said about the subhypergraph $E[X]$ induced by a generic parameterized subvariety $X \subseteq \mathbb{F}^n$ of a given dimension k? In particular, what properties of E would guarantee that every definable $E[X]$-independent set $I \subseteq X$ has dimension less than k?

Here and in what follows we use the words "less" and "greater" to mean "strictly less" and "strictly greater," respectively. Our main result is Theorem 2.7, which gives an answer to Question 1.3 that is inspired by the analogous results for finite hypergraphs. Our proof strategy is to adapt the ideas of the containers method and to control the subhypergraph $E[X]$ using a sequence of small "fingerprints." However, while the standard hypergraph containers method involves the so-called *scythe algorithm* – an iterative procedure that runs through the vertices of the hypergraph – we must use a different approach, since our hypergraphs typically have infinitely many vertices. Similar considerations in a different infinitary setting have previously led the current authors together with Henry Towsner [4] to develop a *nonalgorithmic* proof of the containers theorem for finite hypergraphs. Unfortunately, the argument in [4] is still too "discrete" and not directly applicable in our current framework. Thankfully, there are other tools available in the algebraic context, most notably irreducibility, that allow us to replace induction over the vertex set with induction on the dimension.

A crucial role in our arguments is played by Theorem 4.15 – a certain generalization of the classical dimension of fibers theorem in algebraic geometry, that is interesting in its own right. We state and prove Theorem 4.15 in Section 4; see Section 4.1 for a motivational discussion. The remainder of the proof of Theorem 2.7 proceeds via a sequence of applications of Theorem 4.15 and is presented in Section 5.

2. Main definitions and results

2.1. Basic notation, terminology, and conventions

Integers. We use $\mathbb{N}$ to denote the set of all nonnegative integers. For $n \in \mathbb{N}$, we define $[n] := \{1, \dots, n\}$. By default, the variables $d, i, j, k, m, n, r, s, t$ range over $\mathbb{N}$.

Projections. For a family of sets $(X_i)_{i \in I}$ and $\varnothing \neq S \subseteq I$, let $\mathrm{proj}_S \colon \prod_{i \in I} X_i \to \prod_{i \in S} X_i$ be the projection onto the set S of coordinates. For brevity, given $i \in I$, we write proj_i

instead of $\mathrm{proj}_{\{i\}}$. The Cartesian power X^t of a set X is viewed as a product indexed by $[t]$.

Definable sets. Throughout, we work in a fixed algebraically closed field $\mathbb{F}$. The word "definable" always means "definable in $(\mathbb{F}, +, \cdot)$ with parameters." We mostly work in affine spaces $\mathbb{F}^n$, but sometimes we shall also use the projective n-space $\mathbb{P}^n$ over $\mathbb{F}$. The only topology we refer to is the Zariski topology (on $\mathbb{F}^n$, $\mathbb{P}^n$, etc.). The closure of a set $X \subseteq \mathbb{F}^n$ (in the Zariski topology on $\mathbb{F}^n$) is denoted by $\overline{X}$. We say that a definable set $X \subseteq \mathbb{F}^n$ is *irreducible* if $\overline{X}$ is an irreducible variety. An (*irreducible*) *component* of a definable set X is any set of the form $C \cap X$, where $C \subseteq \overline{X}$ is an irreducible component of $\overline{X}$. The *dimension* $\dim X$ of a definable set X is equal to $\dim \overline{X}$, the Krull dimension of the variety $\overline{X}$ (which coincides with the Morley rank of X, see [13, Section 6]). By convention, $\dim \varnothing := -1$. Further algebraic-geometric and model-theoretic preliminaries are reviewed in Section 3.

Genericity. For a definable set $X \subseteq \mathbb{F}^n$ and a property $\mathfrak{P}$ of elements of $\mathbb{F}^n$, we say that a *generic point* $x \in X$ *satisfies* $\mathfrak{P}$, in symbols $\forall^* x \in X\ (\mathfrak{P}(x))$, if the set

$$\mathfrak{P}(X) := \{x \in X : \mathfrak{P}(x)\}$$

is definable and $\overline{\mathfrak{P}(X)} = \overline{X}$. This definition is equivalent to an (apparently stronger) requirement that $\mathfrak{P}(X)$ is a definable set that contains a dense relatively open subset of $\overline{X}$. Another convenient way to phrase this is that generic $x \in X$ satisfies $\mathfrak{P}$ if and only if the set $\mathfrak{P}(X)$ is definable and has codimension 0 in every irreducible component of X. See Section 3.4 for some basic properties of the $\forall^*$ quantifier.

2.2. *Rational maps*

By a *rational map* from $\mathbb{F}^k$ to $\mathbb{F}^n$ we mean an expression of the form

$$f(x_1, \dots, x_k) = \left(\frac{p_1(x_1, \dots, x_k)}{p_0(x_1, \dots, x_k)}, \dots, \frac{p_n(x_1, \dots, x_k)}{p_0(x_1, \dots, x_k)} \right), \tag{2.1}$$

where $p_0, p_1, \dots, p_n \in \mathbb{F}[x_1, \dots, x_k]$ and p_0 is not identically zero. For convenience, we identify each rational map f as in (2.1) with the tuple of polynomials $(p_0, p_1, \dots, p_n)$ rather than with the corresponding partial function $\mathbb{F}^k \dashrightarrow \mathbb{F}^n$ (for instance, multiplying every polynomial in (2.1) by the same element of $\mathbb{F} \setminus \{0, 1\}$ produces a different rational map). The set of all rational maps from $\mathbb{F}^k$ to $\mathbb{F}^n$ is denoted by $\mathcal{R}(k, n)$.

Given a nonzero polynomial $q \in \mathbb{F}[x_1, \dots, x_k]$, we write $\mathcal{R}_d(k, n; q)$ to indicate for the set of all rational maps f of the form (2.1) with

$$p_0 = q \qquad \text{and} \qquad \deg(p_i) \leqslant d \quad \text{for all } 1 \leqslant i \leqslant n.$$

In particular, $\mathcal{R}_d(k, n; 1)$ is the set of all *polynomial maps* from $\mathbb{F}^k$ to $\mathbb{F}^n$ of degree at most d. For brevity, let $\mathcal{L}(k, n) := \mathcal{R}_1(k, n; 1)$ denote the set of all *affine maps* from $\mathbb{F}^k$ to $\mathbb{F}^n$. We will also use the shortcut $\mathcal{L}_{<k}(n) := \bigcup_{i<k} \mathcal{L}(i, n)$.

A rational map in $\mathcal{R}_d(k,n;q)$ can be described by a tuple of $N := n\binom{k+d}{d}$ elements of $\mathbb{F}$, namely the coefficients of the corresponding polynomials $p_1,\dots,p_n$ in (2.1). As mentioned in the introduction, we use this observation to identify $\mathcal{R}_d(k,n;q)$ with the space $\mathbb{F}^N$, so it makes sense to talk about definable subsets of $\mathcal{R}_d(k,n;q)$ as well as the properties of a generic element of $\mathcal{R}_d(k,n;q)$.

2.3. *Definable independent sets*

Let $t \geqslant 1$ and $E \subseteq (\mathbb{F}^n)^t$. We say that a set $I \subseteq \mathbb{F}^n$ is *E-independent* if $E \cap I^t = \varnothing$, i.e., there are no $x_1,\dots,x_t \in I$ with $(x_1,\dots,x_t) \in E$.

A common feature of the techniques for proving results in the sparse random setting is their reliance on the corresponding theorems in the *dense* case. For instance, all the known proofs of Theorem 1.2 treat Szemerédi's theorem[1] as a black box. Proving a result in the dense case can often be a challenging task in its own right; and indeed, the principal obstacles to resolving several open problems in the theory of sparse random structures lie in obtaining sufficiently strong "supersaturation" bounds in the dense setting (see, e.g., [3,8]). However, the situation simplifies dramatically in the algebraic context, thanks to the following proposition, which provides a convenient criterion for when a definable hypergraph E admits a "large" definable independent set:

Proposition 2.2. *Let $t \geqslant 1$ and let $E \subseteq (\mathbb{F}^n)^t$ be a definable set. The following statements are equivalent:*

(i) *the dimension of every definable E-independent set $I \subseteq \mathbb{F}^n$ is less than n,*

(ii) *E has an irreducible component $H \subseteq E$ such that for all $i \in [t]$, we have*

$$\dim \operatorname{proj}_i H = n.$$

Proof. First, consider the case when E is irreducible. To prove (i) $\Rightarrow$ (ii), note that if $\dim \operatorname{proj}_i E < n$ for some $i \in [t]$, then $\mathbb{F}^n \setminus \operatorname{proj}_i E$ is a definable E-independent set of dimension n. Towards the other implication, assume (ii) and let $I \subseteq \mathbb{F}^n$ be a definable set of dimension n. For each $i \in [t]$, let

$$E_i := E \cap \operatorname{proj}_i^{-1}(I).$$

Since, by assumption, $\dim \operatorname{proj}_i E = \dim I = n$ and E is irreducible, the dimension of fibers theorem (specifically, Corollary 3.10) yields $\dim E_i = \dim E$. The irreducibility of E then gives

$$\dim(E \cap I^t) = \dim(E_1 \cap \cdots \cap E_t) = \dim(E) \geqslant 0,$$

and hence I is not E-independent.

Now let E be arbitrary. Since we have already verified the proposition for irreducible sets, it remains to show that if every definable *E-independent* set is of dimension less

[1] Or, more precisely, its robust version, originally due to Varnavides [25] in the case of 3-term progressions.

than n, then E has a component H such that every definable *H-independent* set is of dimension less than n as well. To that end, suppose that the components of E are $H_1, \dots, H_k$ and, for each $j \in [k]$, let $I_j \subseteq \mathbb{F}^n$ be a definable H_j-independent set of dimension n. Since $E = H_1 \cup \dots \cup H_k$, we conclude that $I_1 \cap \dots \cap I_k$ is a definable E-independent set of dimension n, which finishes the proof. ∎

An immediate consequence of Proposition 2.2 is, for example, the fact that for any $t \geqslant 3$, there is no n-dimensional definable set $I \subseteq \mathbb{F}^n$ without a t-term arithmetic progression, since the set

$$E \coloneqq \{(x_1, \dots, x_t) \in (\mathbb{F}^n)^t : x_2 - x_1 = \dots = x_t - x_{t-1} \text{ and } x_1 \neq x_2\}$$

is irreducible and satisfies $\operatorname{proj}_i E = \mathbb{F}^n$ for all $i \in [t]$. It might perhaps seem strange that this algebraic version of Szemerédi's theorem is almost trivial, while Szemerédi's theorem itself is an extremely deep result. The explanation is simple. Any definable set $I \subseteq \mathbb{F}^n$ must fulfill one of the two alternatives: either $\dim I < n$, or else, $\dim(\mathbb{F}^n \setminus I) < n$. This is analogous, in the finite setting, to I either having *density* 0 or *density* 1. On the other hand, the statement of Szemerédi's theorem is only difficult for sets whose density is *small but positive* – and they do not exist in the algebraic setting.

What makes Proposition 2.2 particularly useful is that, while property (i) is of primary interest to us, it is not apparently first-order, due to the quantification over all definable sets $I \subseteq \mathbb{F}^n$; on the other hand, property (ii) is more "concrete," and indeed, it is definable in families (see Corollary 3.6). Notice however that, to be able to verify (ii), we must have good control over the individual irreducible components of E, as it is not enough to simply know that $\dim \operatorname{proj}_i E = n$ for all $i \in [t]$. For instance, it may happen that E is expressible as a union $E = E_1 \cup E_2$ of two definable sets $E_1, E_2 \subseteq (\mathbb{F}^n)^t$ with $\dim \operatorname{proj}_1 E_1 = \dim \operatorname{proj}_2 E_2 = n$ but $\dim \operatorname{proj}_2 E_1, \dim \operatorname{proj}_1 E_2 < n$, in which case property (ii), and hence also (i), fails. (The simplest example of this situation is the set $E = (\mathbb{F}^n \times \{0\}) \cup (\{0\} \times \mathbb{F}^n) \subseteq (\mathbb{F}^n)^2$.) This issue will be a source of some important technical complications in our arguments.

2.4. *The main result*

Let $E \subseteq (\mathbb{F}^n)^t$ be a definable set. Given a definable map $f \colon \mathbb{F}^k \dashrightarrow \mathbb{F}^n$, we wish to consider the subhypergraph of E induced by the subset $f(\mathbb{F}^k) \subseteq \mathbb{F}^n$ parameterized by f. Since it will be more convenient to work directly in the parameter space $\mathbb{F}^k$, we define the *subhypergraph of E induced by f* to be the set $E[f] \subseteq (\mathbb{F}^k)^t$ given by

$$(y_1, \dots, y_t) \in E[f] \iff (f(y_1), \dots, f(y_t)) \in E.$$

A peculiar feature of this definition is that it makes sense even when $k > n$; in other words, the dimension of the vertex set of $E[f]$ can exceed that of the vertex set of E.

We say that a definable set $E \subseteq (\mathbb{F}^n)^t$ is *injective* if for all $(x_1, \dots, x_t) \in E$, the elements $x_1, \dots, x_t$ are pairwise distinct. When E is thought of as a hypergraph on $\mathbb{F}^n$, the injectivity of E means that it is "truly" t-uniform, i.e., every edge of E contains precisely t distinct vertices.

The next definition is particularly important. We say that a definable set $E \subseteq (\mathbb{F}^n)^t$ is *r-almost dense* if for all subsets $\varnothing \neq S \subseteq [t]$, we have

$$\dim \operatorname{proj}_S E \geqslant |S|n - (|S| - 1)r. \tag{2.3}$$

Observe the following chain of implications:

$$\begin{aligned} \overline{E} = (\mathbb{F}^n)^t \iff E \text{ is 0-almost dense} &\implies E \text{ is 1-almost dense} \\ &\implies E \text{ is 2-almost dense} \\ &\implies \dots . \end{aligned}$$

The notion of r-almost density for subsets $E \subseteq (\mathbb{F}^n)^t$ is mostly interesting for $r \leqslant n$, as if $r \geqslant n$, then

$$E \subseteq (\mathbb{F}^n)^t \text{ is } r\text{-almost dense} \iff \dim \operatorname{proj}_i E = n \text{ for all } i \in [t]. \tag{2.4}$$

We emphasize that in (2.3), n is the dimension of the vertex set of E; for instance, when we say that a definable set $E \subseteq (\mathbb{F}^k)^t$ is r-almost dense, then n must be replaced by k.

Observation 2.5. *Let $t \geqslant 1$ and let $E \subseteq (\mathbb{F}^n)^t$ be an irreducible definable set. If E is n-almost dense, then the dimension of every definable E-independent set $I \subseteq \mathbb{F}^n$ is less than n.*

Proof. Immediate from (2.4) and Proposition 2.2. ∎

The significance of the notion of almost density is demonstrated by the following proposition, whose proof is deferred to Section 5.1:

Proposition 2.6. *Let $t \geqslant 1$ and let $E \subseteq (\mathbb{F}^n)^t$ be an injective definable set. If there exist $d \geqslant t - 1$ and a nonzero polynomial $q \in \mathbb{F}[x_1, \dots, x_k]$ such that, for generic element $f \in \mathcal{R}_d(k, n; q)$, every definable $E[f]$-independent set $I \subseteq \mathbb{F}^k$ has dimension less than k, then E has a k-almost dense irreducible component.*

Now we are ready to state the main result of this paper, answering Question 1.3:

Theorem 2.7. *Let $t \geqslant 1$ and let $E \subseteq (\mathbb{F}^n)^t$ be an r-almost dense irreducible definable set. Fix $d \geqslant t - 1$ and a nonzero polynomial $q \in \mathbb{F}[x_1, \dots, x_k]$. If $k \geqslant r + 1$, then, for generic element $f \in \mathcal{R}_d(k, n; q)$, the following holds: Every definable $E[f]$-independent set $I \subseteq \mathbb{F}^k$ has dimension less than k. Furthermore, if E is injective, then every irreducible component of $E[f]$ is r-almost dense.*

We finish this section with a few remarks about the statement of Theorem 2.7.

Lower bound on d. Theorem 2.7 requires d to grow with t, the uniformity of the hypergraph E. Informally, one could say that the maps given by polynomials of low degree are not "random enough" for the conclusion of Theorem 2.7 to hold. In fact, the lower bound $d \geqslant t - 1$ is best possible, as the following construction shows. Take any $t \geqslant 2$, $d \leqslant t - 2$, and $n > t$. Define

$$H := \{(x_1, \dots, x_t) \in (\mathbb{F}^n)^t : \exists f \in \mathcal{R}_d(1, n; 1)\, (x_1, \dots, x_t \in f(\mathbb{F}))\}.$$

Set $E := (\mathbb{F}^n)^t \setminus H$. Using that $d \leqslant t-2$ and $n > t$, we obtain

$$\begin{aligned}\dim H &\leqslant \dim(\mathcal{R}_d(1,n;1) \times \mathbb{F}^t) = n(d+1) + t \leqslant n(t-1) + t \\ &< nt = \dim(\mathbb{F}^n)^t,\end{aligned}$$

and hence $\overline{E} = (\mathbb{F}^n)^t$. But, on the other hand, $E[f] = \varnothing$ for all $f \in \mathcal{R}_d(1,n;1)$ by definition.

Lower bound on k. If, in the setting of Theorem 2.7, we let r be the smallest integer such that E is r-almost dense, then it follows from Proposition 2.6 that k must be at least r. Actually, the conclusion of Theorem 2.7 can fail even when $k = r$ (and hence the lower bound $k \geqslant r+1$ is best possible). To see this, take any $k < n$ and define a set $E \subseteq (\mathbb{F}^n)^2$ as follows. Pick an arbitrary $(n-k)$-dimensional linear subspace $V \subseteq \mathbb{F}^n$ and put

$$(x,y) \in E \iff x \neq y \text{ and } x - y \in V.$$

Since $n - k > 0$, we have $\operatorname{proj}_1 E = \operatorname{proj}_2 E = \mathbb{F}^n$ and $\dim E = 2n - k$, so E is k-almost dense. However, for a generic affine map $\ell \in \mathcal{L}(k,n)$, the set $\ell(\mathbb{F}^k)$ is a k-dimensional affine subspace of $\mathbb{F}^n$ that intersects every translate of V in precisely one point – and hence $E[\ell] = \varnothing$.

The codegree conditions. We now briefly comment on the relationship between the statement of Theorem 2.7 and the sparse random results in the finite setting.

In order to establish a transference principle for a finite t-uniform hypergraph E, it is typical to assume that E has "many" edges, and that the edges of E are somehow "well-distributed" over the vertex set. For example, in the containers method, these assumptions take the form of the *codegree conditions*: The degree of each vertex of E is required to be close to the average, and similar restrictions are put on the codegrees of all sets of fewer than t vertices.

In Theorem 2.7, the part that forces the edges of E to be "well-distributed" is the assumption of *irreducibility*. For instance, it follows from the dimension of fibers theorem (see Section 3.2) that if a definable set $E \subseteq (\mathbb{F}^n)^t$ is irreducible and satisfies

$$\dim \operatorname{proj}_i E = n \quad \text{for all } i \in [t],$$

then the "degree" of a generic vertex $x \in \mathbb{F}^n$ – i.e., the dimension of the set of all tuples $(x_1, \dots, x_t) \in E$ such that $x \in \{x_1, \dots, x_t\}$ – is equal to the "average" value $\dim E - n$. Since irreducibility takes care of the "uniform distribution" of the edges, there is no need to explicitly bound the codegrees in the statement of Theorem 2.7, and the only numerical assumption left is that E has "many" edges – specifically, it must be r-almost dense.

Nevertheless, there are still some close parallels between the statement of Theorem 2.7 and, say, that of [2, Proposition 3.1]. Indeed, let E be a finite t-uniform hypergraph on a set X and imagine that we wish to apply [2, Proposition 3.1] to E; in particular, let p be a value between 0 and 1. The conclusion of [2, Proposition 3.1] is only interesting for independent sets of size at least $\Omega(p|X|)$. Set $n := \log|X|$ and $r := \log(p|X|)$. For each $s \in [t]$, let m_s be the logarithm of the number of s-element subsets of X that are contained

in at least one edge of E. The codegree conditions of [2, Proposition 3.1] then imply

$$m_s \geqslant sn - (s-1)r + O(1),$$

which should be compared to (2.3).

3. Preliminaries

3.1. Definability in algebraically closed fields

The following fundamental fact and its immediate consequences will be used without mention:

Theorem 3.1 (Quantifier elimination [13, Theorem 3.2.2]). *Every definable set $X \subseteq \mathbb{F}^n$ is constructible, i.e., it is a finite Boolean combination of closed sets.*

For $A \subseteq X \times Y$ and $x \in X$, we let A_x denote the *fiber of A over x*, i.e., the set

$$A_x := \{y \in Y : (x, y) \in A\}.$$

The next theorem follows from [14, Section I.8, Corollary 3] and is a special case of the fact that Morley rank is definable in strongly minimal theories [13, Lemma 6.2.20]:

Lemma 3.2 (Definability of dimension [14, Section I.8, Corollary 3]). *Let $A \subseteq \mathbb{F}^n \times \mathbb{F}^m$ be a definable set and let $k \in \mathbb{N}$. Then the set $\{x \in \mathbb{F}^n : \dim A_x = k\}$ is definable.*

We shall also require certain more advanced facts concerning definability in algebraically closed fields, all of which are well-known and can be easily derived from general results obtained by van den Dries and Schmidt in their foundational paper [24]. For additional information and alternative presentation, we also refer the reader to Chapter IV of van den Dries's thesis [23] and to Chapter 10 of Johnson's thesis [11].

Lemma 3.3 (Fiberwise closure [11, Theorem 10.2.1(1)]). *Let $A \subseteq \mathbb{F}^n \times \mathbb{F}^m$ be a definable set. Then the set $\{(x, y) \in \mathbb{F}^n \times \mathbb{F}^m : y \in \overline{A_x}\}$ is definable.*

Lemma 3.4 (Irreducible fibers [11, Theorem 10.2.1(2)]). *Let $A \subseteq \mathbb{F}^n \times \mathbb{F}^m$ be a definable set. Then the set $\{x \in \mathbb{F}^n : A_x \text{ is irreducible}\}$ is definable.*

The next definability result will play a particularly important role in the sequel. For a definable set $X \subseteq \mathbb{F}^n$, let $\mathfrak{I}(X) \subseteq X \times X$ denote the relation given by

$$(x_1, x_2) \in \mathfrak{I}(X) \iff \text{there is an irreducible component } I \text{ of } X \text{ such that } x_1, x_2 \in I.$$

Note that $\mathfrak{I}(X)$ is "almost" an equivalence relation: it is reflexive and symmetric, but not necessarily transitive, as distinct irreducible components of X need not be disjoint.

Lemma 3.5 (Definability of components). *Let $A \subseteq \mathbb{F}^n \times \mathbb{F}^m$ be a definable set. Then the set*

$$\mathfrak{X} := \{(x, y_1, y_2) \in \mathbb{F}^n \times \mathbb{F}^m \times \mathbb{F}^m : (y_1, y_2) \in \mathfrak{I}(A_x)\}$$

is definable as well.

Proof. Lemma 3.5 is a straightforward consequence of [24, Theorem 2.10] and other results of that paper. For completeness, we include a direct derivation from Lemmas 3.3 and 3.4 here. This is the only place in the paper where we invoke nontrivial model-theoretic machinery. Replacing $\mathbb{F}$ by an elementary extension if necessary, we may arrange that $\mathbb{F}$ is $\aleph_1$-saturated.[2,3] For a definable set $Y \subseteq \mathbb{F}^m$, let $\Im_{k,d}(Y)$ denote the set of all pairs $(y_1, y_2) \in Y \times Y$ such that y_1 and y_2 belong to an irreducible component of $\overline{Y}$ cut out from $\mathbb{F}^m$ by at most k polynomials of degree at most d. Clearly, $\Im(Y) = \bigcup_{k,d} \Im_{k,d}(Y)$, and, since a closed set $Z \subseteq \mathbb{F}^m$ is an irreducible component of $\overline{Y}$ if and only if Z is irreducible, $Z \subseteq \overline{Y}$, and $Z \nsubseteq \overline{Y \setminus Z}$, by Lemmas 3.3 and 3.4, the set

$$\mathfrak{X}_{k,d} := \{(x, y_1, y_2) \in \mathbb{F}^n \times \mathbb{F}^m \times \mathbb{F}^m : (y_1, y_2) \in \Im_{k,d}(A_x)\}$$

is definable. Hence, $\mathfrak{X} = \bigcup_{k,d} \mathfrak{X}_{k,d}$ is a countable *union* of definable sets. On the other hand, let $\Im^{k,d}(Y)$ be the set of all pairs $(y_1, y_2) \in Y \times Y$ such that every closed subset $Z \subseteq \mathbb{F}^m$ cut out by at most k polynomials of degree at most d satisfies

$$y_1 \notin \overline{Y \setminus Z} \implies y_2 \in Z.$$

Then $\Im(Y) = \bigcap_{k,d} \Im^{k,d}(Y)$, and Lemma 3.3 yields that the set

$$\mathfrak{X}^{k,d} := \{(x, y_1, y_2) \in \mathbb{F}^n \times \mathbb{F}^m \times \mathbb{F}^m : (y_1, y_2) \in \Im^{k,d}(A_x)\}$$

is definable, so $\mathfrak{X} = \bigcap_{k,d} \mathfrak{X}^{k,d}$ is a countable *intersection* of definable sets. Since $\mathbb{F}$ is $\aleph_1$-saturated, a set that is both a countable union and a countable intersection of definable sets must itself be definable. ∎

Lemma 3.5 allows using quantification over irreducible components when forming a definable set. Below we present, as an illustration, a typical example of how Lemma 3.5 can be applied:

Corollary 3.6. *Let $E \subseteq \mathbb{F}^n \times (\mathbb{F}^m)^t$ be a definable set. Then the set*

$$\{x \in \mathbb{F}^n : \textit{the dimension of every definable } E_x\textit{-independent set } I \subseteq \mathbb{F}^m \textit{ is less than } m\}$$

is definable.

Proof. By Proposition 2.2, the set in question coincides with the set of all $x \in \mathbb{F}^n$ such that E_x has an irreducible component H with $\dim \mathrm{proj}_i H = m$ for all $i \in [t]$. The existential quantification over the components of E_x can be turned into a quantification over

[2] Actually, *every* uncountable algebraically closed field is saturated, since the theory of algebraically closed fields of any fixed characteristic is uncountably categorical, hence ω-stable; see [13, Theorem 6.5.4].

[3] Here we use the fact that if $\mathbb{K} \supseteq \mathbb{F}$ is an algebraically closed field extending $\mathbb{F}$ and $X \subseteq \mathbb{K}^n$ is a set definable in $\mathbb{K}$, then $X \cap \mathbb{F}^n$ is definable in $\mathbb{F}$. This can be verified using quantifier elimination (i.e., Theorem 3.1) and the fact that if $X \subseteq \mathbb{K}^n$ is closed in $\mathbb{K}^n$, then $X \cap \mathbb{F}^n$ is closed in $\mathbb{F}^n$. An analogous result holds more generally in arbitrary stable theories and is a consequence of definability of types; see [15, Exercise 1.29].

the points of E_x as follows: There is an element $e \in E_x$ such that:

(a) for all $h_1, h_2 \in E_x$, if $(e, h_1) \in \mathfrak{F}(E_x)$ and $(e, h_2) \in \mathfrak{F}(E_x)$, then $(h_1, h_2) \in \mathfrak{F}(E_x)$,

(b) for all $i \in [t]$, $\dim \mathrm{proj}_i\{h \in E_x : (e, h) \in \mathfrak{F}(E_x)\} = m$.

It remains to apply Lemmas 3.5 and 3.2. ∎

Throughout the rest of this paper, arguments that are similar to and just as straightforward as the proof of Corollary 3.6 are omitted.

3.2. The dimension of fibers theorem and its consequences

A central role in the sequel is played by the so-called *dimension of fibers theorem*. It is a fundamental result, with many versions and generalizations that fall far beyond the scope of this article. We only give here the statements that will be explicitly used later on; for a more thorough discussion, see, e.g., [14, Section I.8] and [22, Section 11.4].

Theorem 3.7 (Dimension of fibers, ess. [14, Section I.8, Corollary 1]). *Let $A \subseteq \mathbb{F}^n \times \mathbb{F}^m$ be an irreducible definable set. For generic $x \in \mathrm{proj}_1 A$, the dimension of every component of the set A_x is $\dim A - \dim \mathrm{proj}_1 A$.*

Since the above theorem is usually stated for varieties rather than definable sets, we include here a derivation of the general case of Theorem 3.7 from the case when A is a closed set:

Derivation of Theorem 3.7 from the closed set case. Let $B := \overline{A} \setminus A$ and note that

$$\dim B < \dim A = \dim \overline{A}.$$

Since we assume that Theorem 3.7 holds for $\overline{A}$, it suffices to argue that

$$\forall^* x \in \mathrm{proj}_1 \overline{A} \qquad \overline{A_x} = (\overline{A})_x. \tag{3.8}$$

Suppose that (3.8) fails. Since the set $\mathrm{proj}_1 \overline{A}$ is irreducible, we then have

$$\forall^* x \in \mathrm{proj}_1 \overline{A} \qquad \overline{A_x} \subsetneq (\overline{A})_x.$$

Hence, for generic $x \in \mathrm{proj}_1 \overline{A}$, at least one component of the fiber $(\overline{A})_x$ is entirely contained in $\overline{B_x} \subseteq (\overline{B})_x$. In particular,

$$\dim \mathrm{proj}_1 \overline{A} = \dim \mathrm{proj}_1 \overline{B} =: k.$$

However, Theorem 3.7, applied to the closed sets $\overline{A}$ and $\overline{B}$, yields that, for generic element $x \in \mathrm{proj}_1 \overline{A}$, the dimension of every component of $(\overline{A})_x$ is $\dim A - k$, while

$$\dim(\overline{B})_x = \dim B - k < \dim A - k.$$

This contradiction completes the proof of (3.8). ∎

Claim (3.8) is useful enough to be stated as a separate corollary:

Corollary 3.9. *Let $A \subseteq \mathbb{F}^n \times \mathbb{F}^m$ be a definable set. Then, for generic $x \in \mathrm{proj}_1 \overline{A}$, we have $\overline{A_x} = (\overline{A})_x$.*

The next result is a well-known and easy consequence of Theorem 3.7:

Corollary 3.10 (Fubini for dimension). *Let $A \subseteq \mathbb{F}^n \times \mathbb{F}^m$ be a nonempty definable set. Then*

$$\dim A = \max_{k \in \mathbb{N}} \big(k + \dim\{x \in \operatorname{proj}_1 A : \dim A_x = k\}\big).$$

The following immediate consequence of Corollary 3.9 will be used repeatedly:

Corollary 3.11. *Let $A \subseteq \mathbb{F}^n \times \mathbb{F}^m$ be a definable set. Then $\overline{\operatorname{proj}_1 A} = \overline{\operatorname{proj}_1 \overline{A}}$.*

The following statement will play a crucial role in the later stages of the proof of Theorem 2.7:

Corollary 3.12 (Generic indecomposability of fibers). *Let $A \subseteq \mathbb{F}^n \times \mathbb{F}^m$ be a nonempty irreducible definable set and let $B \subseteq A$ be a definable subset such that for generic element $x \in \operatorname{proj}_1 A$, the set B_x is a union of irreducible components of A_x. Then*

$$\textit{either} \quad \forall^* x \in \operatorname{proj}_1 A \ (B_x = \varnothing), \quad \textit{or} \quad \forall^* x \in \operatorname{proj}_1 A \ (B_x = A_x).$$

Proof. Suppose that, for generic $x \in \operatorname{proj}_1 A$, the set B_x is nonempty. Then, by Theorem 3.7, we must have $\forall^* x \in \operatorname{proj}_1 A \ (\dim B_x = \dim A - \dim \operatorname{proj}_1 A)$, and therefore,

$$\dim B \geqslant \dim \operatorname{proj}_1 A + (\dim A - \dim \operatorname{proj}_1 A) = \dim A.$$

Since A is irreducible, this yields $\overline{B} = \overline{A}$, and hence $\overline{B_x} = \overline{A_x}$ for generic $x \in \operatorname{proj}_1 A$. But if B_x is a union of components of A_x, then $\overline{B_x} = \overline{A_x}$ is equivalent to $B_x = A_x$, and we are done. ∎

3.3. An irreducibility criterion

To verify that certain sets appearing in the proof of Theorem 2.7 are irreducible, we will need the following fact (it is the only statement in this paper that requires leaving the realm of affine spaces):

Lemma 3.13 (ess. [22, Exercise 11.4.C]). *Let $A \subseteq \mathbb{F}^n \times \mathbb{F}^m$ be a nonempty definable set and let A^* be the closure of A in $\mathbb{F}^n \times \mathbb{P}^m$. Suppose that $\operatorname{proj}_1 A$ is irreducible and all the fibers $A^*_x \subseteq \mathbb{P}^m$ with $x \in \operatorname{proj}_1 A$ are irreducible and of the same dimension. Then A is irreducible.*

Proof. Let $X := \operatorname{proj}_1 A$ and let d denote the common dimension of the fibers of A^* over the points in X. Then we must have

$$d = \dim A - \dim X.$$

Let B be any irreducible component of A^* such that $\dim \operatorname{proj}_1 B = \dim X$. Since $\mathbb{P}^m$ is a complete variety [14, Section I.9, Theorem 1], $\operatorname{proj}_1 B$ is closed in $\mathbb{F}^n$, and, since X is irreducible, $X \subseteq \operatorname{proj}_1 B$. By [14, Section I.8, Theorem 2], we obtain that for all $x \in X$,

$$\dim B_x \geqslant \dim B - \dim \operatorname{proj}_1 B = \dim A - \dim X = d,$$

and thus $\dim B_x = \dim A^*_x$. Since A^*_x is irreducible, this implies $B_x = A^*_x$ and hence $B = A^*$. ∎

Corollary 3.14. *Suppose that $p_1, \dots, p_k \in \mathbb{F}[x_1, \dots, x_n, y_1, \dots, y_m]$ are polynomials that are linear in the set of variables $y_1, \dots, y_m$ and let $Z \subseteq \mathbb{F}^n \times \mathbb{F}^m$ be the set of common zeros of $p_1, \dots, p_k$. If $X \subseteq \mathbb{F}^n$ is an irreducible definable set such that* $\dim Z_x$ *is the same for all $x \in X$, then $Z \cap (X \times \mathbb{F}^m)$ is irreducible.*

Proof. If $Z_x = \varnothing$ for all $x \in X$, then $Z \cap (X \times \mathbb{F}^m) = \varnothing$ and there is nothing to prove. Otherwise, let Z^* be the closure of Z in $\mathbb{F}^n \times \mathbb{P}^m$. For each $x \in \mathbb{F}^n$, the set Z^*_x is cut out from $\mathbb{P}^m$ by a system of homogeneous linear equations. Hence, the fiber Z^*_x is irreducible and if $Z_x \neq \varnothing$, then Z_x is dense in Z^*_x. The desired conclusion now follows by Lemma 3.13. ∎

3.4. Fubini-like properties of the generic quantifier

Now we state and prove some results that allow changing the order of multiple $\forall^*$ quantifiers, in the spirit of the theorems of Fubini in measure theory and Kuratowski–Ulam in general topology.

Theorem 3.15 (Fubini for $\forall^*$). *Let $A \subseteq \mathbb{F}^n \times \mathbb{F}^m$ be an irreducible definable set and let $X := \operatorname{proj}_1 A$. Then, for any definable set $B \subseteq A$, the following statements are equivalent:*

(i) $\forall^* x \in X\ \forall^* y \in A_x\ ((x, y) \in B)$,

(ii) $\forall^* x \in X\ (\dim B_x = \dim A_x)$,

(iii) $\forall^* (x, y) \in A\ ((x, y) \in B)$, *i.e.,* $\dim B = \dim A$.

Proof. Implication (i) ⇒ (ii) is clear. For (ii) ⇒ (iii), we use the dimension of fibers theorem to observe that $\forall^* x \in X\ (\dim A_x = \dim A - \dim X)$, so if (ii) holds, then

$$\dim B \geqslant \dim X + (\dim A - \dim X) = \dim A,$$

as desired. Finally, to prove the implication (iii) ⇒ (i), consider the set $C \subseteq A$ such that for each $x \in X$, the fiber C_x is the union of all the irreducible components of A_x in which B_x is dense. It follows from the results of Section 3.1 that C is definable, and, by Corollary 3.12, we either have $\forall^* x \in X\ (C_x = A_x)$ or $\forall^* x \in X\ (C_x = \varnothing)$. In the former case, statement (i) holds (and we are done), so assume that $\forall^* x \in X\ (C_x = \varnothing)$, i.e., $\forall^* x \in X\ \forall^* y \in A_x\ ((x, y) \notin B)$. Applying (i) ⇒ (iii) with $A \setminus B$ in place of B yields $\dim(A \setminus B) = \dim A$, and hence $\dim B < \dim A$, as desired. ∎

Corollary 3.16. *Let $X \subseteq \mathbb{F}^n$, $Y \subseteq \mathbb{F}^m$, and $A \subseteq X \times Y$ be definable sets. Then the following statements are equivalent:*

(i) $\forall^* x \in X\ \forall^* y \in Y\ ((x, y) \in A)$,

(ii) $\forall^* y \in Y\ \forall^* x \in X\ ((x, y) \in A)$,

(iii) $\forall^* (x, y) \in X \times Y\ ((x, y) \in A)$.

Proof. Since the components of $X \times Y$ are precisely the products of the components of X and Y, we may assume that X and Y are irreducible. An application of Theorem 3.15 completes the proof. ∎

In the next corollary, it is instructive to think of $A \subseteq X \times X$ as being an equivalence relation on X (and this is the context in which this corollary will be used later on).

Corollary 3.17. *Let $X \subseteq \mathbb{F}^n$ be an irreducible definable set and let $A \subseteq X \times X$ be an irreducible definable set such that $\operatorname{proj}_1 A = \operatorname{proj}_2 A = X$. Then, for any definable set $Y \subseteq X$, the following statements are equivalent:*

(i) $\forall^* x \in X\ \forall^* x' \in A_x\ (x' \in Y)$,

(ii) $\forall^* x \in X\ (x \in Y)$, *i.e.,* $\dim Y = \dim X$.

Proof. By Theorem 3.15, condition (i) is equivalent to

(i') $\dim(A \cap (X \times Y)) = \dim A$,

and it follows from the dimension of fibers theorem that condition (i') is equivalent to $\dim Y = \dim X$. ∎

4. Expansion in algebraic bipartite graphs

4.1. Overview

The central results of this section are Theorem 4.15 and its Corollary 4.23. They play a key role in the proof of Theorem 2.7 and are also interesting in their own right as natural extensions of the dimension of fibers theorem.

Consider an irreducible definable set $A \subseteq \mathbb{F}^n \times \mathbb{F}^m$ such that $\dim \operatorname{proj}_1 A = n$. Given an irreducible definable subset $X \subseteq \mathbb{F}^n$, what is the dimension of the fiber of A over X, i.e., of the set $A \cap (X \times \mathbb{F}^m)$? For simplicity, assume that every component B of the set $A \cap (X \times \mathbb{F}^m)$ is *dominant*, i.e., $\dim \operatorname{proj}_1 B = \dim X$ (the dimension of nondominant components is harder to control). The "expected" answer is

$$\dim(A \cap (X \times \mathbb{F}^m)) = \dim A - n + \dim X, \tag{4.1}$$

and we hope that (4.1) holds for "typical" X. This hope is justified by the dimension of fibers theorem. Indeed, let C be the closure of the set of all elements $x \in \mathbb{F}^n$ such that $\dim A_x \neq \dim A - n$. By Theorem 3.7, $\dim C < n$, and if (4.1) fails, then $X \subseteq C$. In other words, the dimension of fibers theorem gives us a "small" definable set C that contains every counterexample to (4.1).

We wish to obtain a version of this result for the dimension of the *second projection* $\operatorname{proj}_2(A \cap (X \times \mathbb{F}^m))$. Notice that $\operatorname{proj}_2(A \cap (X \times \mathbb{F}^m))$ has a natural combinatorial interpretation: If we think of A as the edge set of a bipartite graph with bipartition $(\mathbb{F}^n, \mathbb{F}^m)$, then $\operatorname{proj}_2(A \cap (X \times \mathbb{F}^m))$ is the *neighborhood* of X in this graph (see Figure 1). There are two obvious upper bounds on $\dim \operatorname{proj}_2(A \cap (X \times \mathbb{F}^m))$: First,

$$\dim \operatorname{proj}_2(A \cap (X \times \mathbb{F}^m)) \leqslant \dim(A \cap (X \times \mathbb{F}^m)) \overset{(4.1)}{=} \dim A - n + \dim X;$$

and second, since $A \cap (X \times \mathbb{F}^m) \subseteq A$,

$$\dim \operatorname{proj}_2(A \cap (X \times \mathbb{F}^m)) \leqslant \dim \operatorname{proj}_2 A.$$

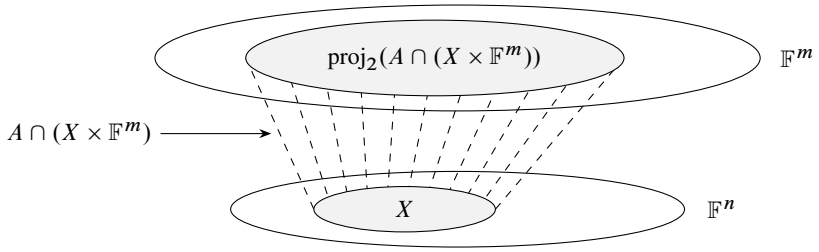

Fig. 1. A combinatorial interpretation of $\mathrm{proj}_2(A \cap (X \times \mathbb{F}^m))$.

The goal of this section is to show that, for "typical" X, at least one of these bounds must be tight:

$$\dim \mathrm{proj}_2(A \cap (X \times \mathbb{F}^m)) = \min\{\dim A - n + \dim X, \dim \mathrm{proj}_2 A\}. \tag{4.2}$$

In other words, algebraic graphs are "maximally expanding": the dimension of the neighborhood of a "typical" set in such a graph is as large as it can possibly be.

What is meant by a "typical" set here is a somewhat subtle issue. Based on the preceding discussion, it is tempting to conjecture that there should be a "small" definable set C that contains every counterexample X to (4.2). This, however, need not be the case, as the following construction illustrates. Suppose that $n = 2$, $m = 1$, and $A \subset \mathbb{F}^2 \times \mathbb{F}$ is given by

$$A := \{(x, y, z) \in \mathbb{F}^2 \times \mathbb{F} : y = zx \text{ and } x \neq 0\}. \tag{4.3}$$

Then A is an irreducible set of dimension 2 and $\mathrm{proj}_2 A = \mathbb{F}$. Thus, for this A and for any 1-dimensional set $X \subset \mathbb{F}^2$, (4.2) turns into $\dim \mathrm{proj}_2(A \cap (X \times \mathbb{F})) = 1$. But if $X \subset \mathbb{F}^2$ is a straight line passing through the origin $(0, 0)$, then $\dim \mathrm{proj}_2(A \cap (X \times \mathbb{F})) \leqslant 0$, and the union of all such lines is all of $\mathbb{F}^2$.

Our approach is to allow the "container" C to *vary* with X, but in a very *limited way*. (This idea is inspired by the hypergraph containers method, where one builds a container for a given independent set I based on a small fingerprint $F \subseteq I$.) To be more precise, let $\mathcal{F}$ be a family of k-dimensional counterexamples to (4.2). (As in the statement of Theorem 2.7, we shall only work with *parameterized* sets, so $\mathcal{F}$ is really a family of *definable functions* rather than sets, but for the purposes of the current informal discussion, this technicality may be ignored.) Imagine that Alice and Bob are playing the following game: Alice secretly chooses a set $X \in \mathcal{F}$. Then she picks an r-dimensional subset $Y \subset X$, where $r < k$, and shows it to Bob. Bob's goal is to find, based on Y alone, a "container" C such that $\dim C < n$ and $X \subseteq C$. Corollary 4.23 asserts, roughly speaking, that there is a definable construction that allows Bob to win for all $X \in \mathcal{F}$ and for generic $Y \subset X$.

As an illustration, consider again the set $A \subset \mathbb{F}^2 \times \mathbb{F}$ given by (4.3) and let $\mathcal{F}$ be the family of all straight lines in $\mathbb{F}^2$ passing through the origin. Whatever $X \in \mathcal{F}$ Alice chooses, when she shows Bob a generic point $y \in X$, he can win simply by making C be the unique straight line passing through $(0, 0)$ and y, in agreement with Corollary 4.23.

Before we proceed to formal statements and proofs, there is one more detail that is worth pointing out, namely what we mean by a *generic* r-dimensional subset $Y \subset X$. As mentioned before, we only work with parametrized sets, so let $f\colon \mathbb{F}^k \dashrightarrow \mathbb{F}^n$ be the parameterization of X. We can then pick a generic affine map $\ell \in \mathcal{L}(r,k)$ and take Y to be the subset of X parameterized by the *composition* $f \circ \ell\colon \mathbb{F}^r \dashrightarrow \mathbb{F}^n$. The reader may be wondering why we restrict our attention to *affine* maps ℓ instead of allowing rational maps of arbitrary degree. The reason, roughly speaking, is that we wish the subset $Y \subset X$ to contain "less information" than X itself. When ℓ is affine, this intuition is justified as $f \circ \ell$ is a map of the same degree as f but in fewer variables. Note that this construction can be iterated; in other words, given $\lambda \in \mathcal{L}(s,r)$, we can look at the triple composition $f \circ \ell \circ \lambda\colon \mathbb{F}^s \dashrightarrow \mathbb{F}^n$, which can be interpreted as picking a generic s-dimensional subset of a generic r-dimensional subset of X. Unsurprisingly, this operation is essentially equivalent to directly picking a generic s-dimensional subset of X (see Lemma 4.10; this again relies on the fact that the maps ℓ and λ are affine, so their composition $\ell \circ \lambda$ is affine as well), and this simple fact will be crucial for our arguments (in particular, in the proof of Proposition 4.13).

4.2. Definable families of functions

By a *definable family of functions* from $\mathbb{F}^k$ to $\mathbb{F}^n$ we mean a pair $(\mathcal{F}, \mathrm{e})$, where $\mathcal{F}$ is a definable set (in some power of $\mathbb{F}$) and $\mathrm{e}\colon \mathcal{F} \times \mathbb{F}^k \dashrightarrow \mathbb{F}^n$ is a definable partial function, called the *evaluation map*, such that $\forall^* f \in \mathcal{F}\ \forall^* y \in \mathbb{F}^k\ (\mathrm{e}(f,y)$ is defined). The evaluation map is usually clear from the context, so we omit it and simply write $\mathcal{F}$ instead of $(\mathcal{F}, \mathrm{e})$. We write $\mathcal{F}\colon \mathbb{F}^k \rightrightarrows \mathbb{F}^n$ to indicate that $\mathcal{F}$ is a definable family of functions from $\mathbb{F}^k$ to $\mathbb{F}^n$. Each $f \in \mathcal{F}$ gives rise to the partial map

$$\mathbb{F}^k \dashrightarrow \mathbb{F}^n\colon y \mapsto \mathrm{e}(f,y),$$

which we also denote by f (so expressions like $f(y)$ or $f \circ g$ must be interpreted accordingly). The basic example of a definable family of functions from $\mathbb{F}^k$ to $\mathbb{F}^n$ is $\mathcal{R}_d(k,n;q)$ for a nonzero polynomial $q \in \mathbb{F}[x_1,\dots,x_k]$ (with the natural evaluation map).

Definition 4.4. A definable family of functions $\mathcal{F}\colon \mathbb{F}^k \rightrightarrows \mathbb{F}^n$ is *comprehensive* if the following statements are equivalent for every definable set $C \subseteq \mathbb{F}^n$:

(C1) $\forall^* f \in \mathcal{F}\ \forall^* y \in \mathbb{F}^k\ (f(y) \in C)$,

(C2) $\forall^* x \in \mathbb{F}^n\ (x \in C)$, i.e., $\dim C = n$.

For instance, the family $\mathcal{R}_d(k,n;q)$ is comprehensive, since for all $y \in \mathbb{F}^k$ with $q(y) \neq 0$ and for all $x, x' \in \mathbb{F}^n$,

$$\dim\{f \in \mathcal{R}_d(k,n;q) : f(y) = x\} = \dim\{f \in \mathcal{R}_d(k,n;q) : f(y) = x'\};$$

indeed, the map

$$\begin{aligned}\{f \in \mathcal{R}_d(k,n;q) : f(y) = x\} &\to \{f \in \mathcal{R}_d(k,n;q) : f(y) = x'\},\\ f &\mapsto f + q(y)(x' - x)/q\end{aligned}$$

is a definable bijection.

Proposition 4.5. *Let $\mathcal{F}\colon \mathbb{F}^k \Rightarrow \mathbb{F}^n$ be a definable family of functions. If $\mathcal{F}$ is irreducible and satisfies the implication (C1) $\Rightarrow$ (C2), then $\mathcal{F}$ is comprehensive.*

Proof. We have to prove (C2) $\Rightarrow$ (C1). To that end, let $C \subseteq \mathbb{F}^n$ be a definable set for which (C1) fails. Since $\mathcal{F}$ is irreducible, the negation of condition (C1) is equivalent to $\forall^* f \in \mathcal{F}\ \forall^* y \in \mathbb{F}^k\ (f(y) \notin C)$, so applying (C1) $\Rightarrow$ (C2) with $\mathbb{F}^n \setminus C$ in place of C yields $\dim(\mathbb{F}^n \setminus C) = n$, i.e., $\dim C < n$, as desired. ∎

The following construction will be useful. Let $\mathcal{F}\colon \mathbb{F}^k \Rightarrow \mathbb{F}^n$ and $\mathcal{G}\colon \mathbb{F}^r \Rightarrow \mathbb{F}^k$ and assume that $\mathcal{G}$ is comprehensive. Define $\mathcal{F} \otimes \mathcal{G}\colon \mathbb{F}^r \Rightarrow \mathbb{F}^n$, the *composition* of $\mathcal{F}$ and $\mathcal{G}$, as follows: As a set, $\mathcal{F} \otimes \mathcal{G}$ is equal to $\mathcal{F} \times \mathcal{G}$, and the evaluation map on $\mathcal{F} \otimes \mathcal{G}$ is given by

$$(f, g)(z) := (f \circ g)(z) \quad \text{for all } f \in \mathcal{F},\ g \in \mathcal{G}, \text{ and } z \in \mathbb{F}^r.$$

Proposition 4.6. *The above definition is correct; that is, for all families $\mathcal{F}\colon \mathbb{F}^k \Rightarrow \mathbb{F}^n$ and $\mathcal{G}\colon \mathbb{F}^r \Rightarrow \mathbb{F}^k$, if $\mathcal{G}$ is comprehensive, then $\mathcal{F} \otimes \mathcal{G}$ is a definable family of functions from $\mathbb{F}^r$ to $\mathbb{F}^n$.*

Proof. Due to Corollary 3.16, we have to verify that

$$\forall^* f \in \mathcal{F} \quad \forall^* g \in \mathcal{G} \quad \forall^* z \in \mathbb{F}^r \qquad (f \circ g)(z) \text{ is defined.}$$

Since the family $\mathcal{G}$ is comprehensive, this is equivalent to

$$\forall^* f \in \mathcal{F} \quad \forall^* y \in \mathbb{F}^k \qquad f(y) \text{ is defined,}$$

which holds by definition, since $\mathcal{F}$ is a definable family of functions. ∎

Proposition 4.7. *Let $\mathcal{F}\colon \mathbb{F}^k \Rightarrow \mathbb{F}^n$ and $\mathcal{G}\colon \mathbb{F}^r \Rightarrow \mathbb{F}^k$. If $\mathcal{F}$ and $\mathcal{G}$ are comprehensive, then so is $\mathcal{F} \otimes \mathcal{G}$.*

Proof. For any definable set $C \subseteq \mathbb{F}^n$, we have

$$\begin{aligned}
&\forall^* f \in \mathcal{F} \quad \forall^* g \in \mathcal{G} \quad \forall^* z \in \mathbb{F}^r \quad (f \circ g)(z) \in C \\
&\iff \forall^* f \in \mathcal{F} \quad \forall^* y \in \mathbb{F}^k \quad f(y) \in C && [\mathcal{G} \text{ is comprehensive}] \\
&\iff \forall^* x \in \mathbb{F}^n \quad x \in C && [\mathcal{F} \text{ is comprehensive}].
\end{aligned}$$ ∎

4.3. Containers

A crucial concept for the results of this section is that of an *r-container*. Informally, an r-container $\mathfrak{C}$ for a definable family $\mathcal{F}\colon \mathbb{F}^k \Rightarrow \mathbb{F}^n$ is a definable rule that, given $f \in \mathcal{F}$ and $\ell \in \mathcal{L}(r, k)$, outputs a subset $\mathfrak{C}_{f\circ\ell} \subseteq \mathbb{F}^n$ that only depends on the composition $f \circ \ell$ and not on f and ℓ themselves. Here is the precise definition:

Definition 4.8. Let $\mathcal{F}\colon \mathbb{F}^k \Rightarrow \mathbb{F}^n$ and let $r < k$. An *r-container* for $\mathcal{F}$ is a definable set

$$\mathfrak{C} \subseteq \mathcal{F} \times \mathcal{L}(r, k) \times \mathbb{F}^n$$

such that for all $f, g \in \mathcal{F}$ and $\ell, \lambda \in \mathcal{L}(r, k)$ with $f \circ \ell = g \circ \lambda$, and for all $x \in \mathbb{F}^n$, we have

$$(f, \ell, x) \in \mathfrak{C} \iff (g, \lambda, x) \in \mathfrak{C}.$$

If $\mathfrak{C}$ is an r-container for $\mathcal{F}$, then, for any definable map $\varphi\colon \mathbb{F}^r \rightharpoonup \mathbb{F}^n$, we write

$$\mathfrak{C}_\varphi := \{x \in \mathbb{F}^n : (f, \ell, x) \in \mathfrak{C} \text{ for some } f \in \mathcal{F} \text{ and } \ell \in \mathcal{L}(r,k) \text{ such that } f \circ \ell = \varphi\}.$$

Definition 4.9. Let $\mathcal{F}\colon \mathbb{F}^k \Rrightarrow \mathbb{F}^n$ and $r \in \{-1, 0, 1, \dots, k-1\}$. The family $\mathcal{F}$ is called *r-uncontainable* if either $r = -1$ and $\mathcal{F}$ is comprehensive, or else, $r \in \mathbb{N}$ and the following statements are equivalent for every r-container $\mathfrak{C}$ for $\mathcal{F}$:

(U1) $\forall^* f \in \mathcal{F}\ \forall^* \ell \in \mathcal{L}(r,k)\ \forall^* y \in \mathbb{F}^k\ (f(y) \in \mathfrak{C}_{f\circ\ell})$,

(U2) $\forall^* f \in \mathcal{F}\ \forall^* \ell \in \mathcal{L}(r,k)\ \forall^* x \in \mathbb{F}^n\ (x \in \mathfrak{C}_{f\circ\ell})$.

The implication (U1) $\Rightarrow$ (U2) in Definition 4.9 can be informally summarized as, "It is impossible to put most elements of $\mathcal{F}$ into small containers," hence the term "uncontainable." The opposite implication (U2) $\Rightarrow$ (U1) says that, conversely, "large containers must capture most elements of $\mathcal{F}$." It is convenient to include the second implication as part of Definition 4.9, even though in most cases of interest to us it will follow automatically, see Proposition 4.12.

When working with r-uncontainable families, we often rely on the following basic fact:

Lemma 4.10. *Suppose that $s \leqslant k$, $s \leqslant r$ and let $\lambda \in \mathcal{L}(s,r)$ be an injective affine map. Then, for every definable set $A \subseteq \mathcal{L}(s,k)$, we have*

$$\forall^* \ell \in \mathcal{L}(r,k) \quad \ell \circ \lambda \in A \qquad \Longleftrightarrow \qquad \forall^* \varphi \in \mathcal{L}(s,k) \quad \varphi \in A.$$

Proof. Since $s \leqslant k$, a generic affine map $\varphi \in \mathcal{L}(s,k)$ is injective, and, similarly, $\ell \circ \lambda$ is injective for generic $\ell \in \mathcal{L}(r,k)$. It remains to observe that for all injective $\varphi \in \mathcal{L}(s,k)$, the dimension of the set $\{\ell \in \mathcal{L}(r,k) : \ell \circ \lambda = \varphi\}$ is the same. ∎

Proposition 4.11. *Let $\mathcal{F}\colon \mathbb{F}^k \Rrightarrow \mathbb{F}^n$ and $s, r \in \{-1, 0, 1, \dots, k-1\}$. If $s \leqslant r$, then*

$$\mathcal{F} \text{ is } r\text{-uncontainable} \implies \mathcal{F} \text{ is } s\text{-uncontainable}.$$

Proof. Suppose that $0 \leqslant s \leqslant r < k$ and $\mathcal{F}$ is r-uncontainable (the proof in the case $s = -1$ is similar, and we omit it). Let $\mathfrak{C}$ be an arbitrary s-container for $\mathcal{F}$. Fix any injective map $\lambda \in \mathcal{L}(s,r)$ and define an r-container $\tilde{\mathfrak{C}}$ by $\tilde{\mathfrak{C}}_{f\circ\ell} := \mathfrak{C}_{f\circ\ell\circ\lambda}$ for all $f \in \mathcal{F}$ and $\ell \in \mathcal{L}(r,k)$. We claim that (U1) and (U2) for $\mathfrak{C}$ are equivalent to the corresponding statements for $\tilde{\mathfrak{C}}$, which implies that $\mathcal{F}$ is s-uncontainable. Indeed, (U1) for $\tilde{\mathfrak{C}}$ takes the form

$$\forall^* f \in \mathcal{F} \quad \forall^* \ell \in \mathcal{L}(r,k) \quad \forall^* y \in \mathbb{F}^k \qquad f(y) \in \mathfrak{C}_{f\circ\ell\circ\lambda}.$$

By Lemma 4.10, this is equivalent to

$$\forall^* f \in \mathcal{F} \quad \forall^* \varphi \in \mathcal{L}(s,k) \quad \forall^* y \in \mathbb{F}^k \qquad f(y) \in \mathfrak{C}_{f\circ\varphi},$$

which is precisely (U1) for $\mathfrak{C}$. The argument for (U2) is similar. ∎

Proposition 4.12. *Let $\mathcal{F}\colon \mathbb{F}^k \Rrightarrow \mathbb{F}^n$ and $r \in \{-1, 0, 1, \dots, k-1\}$. If $\mathcal{F}$ is irreducible and satisfies the implication (U1) $\Rightarrow$ (U2), then $\mathcal{F}$ is r-uncontainable.*

Proof. The proof is the same, *mutatis mutandis*, as the proof of Proposition 4.5. ∎

The family $\mathcal{L}(k,n)$ is easily seen to be $(k-1)$-uncontainable (this is a special case of the more general Proposition 5.8 proved in Section 5). This fact yields the following extension of Proposition 4.7:

Proposition 4.13. *Let $\mathcal{F}\colon \mathbb{F}^k \rightrightarrows \mathbb{F}^n$ and $r \in \{-1, 0, 1, \dots, k-1\}$. Suppose that $s \leqslant k$ and let $t := \min\{r, s-1\}$. If $\mathcal{F}$ is r-uncontainable, then $\mathcal{F} \otimes \mathcal{L}(s,k)$ is t-uncontainable.*

Proof. The case $t = -1$ is handled by Proposition 4.7, so, from now on, assume that $t \in \mathbb{N}$.

Let $\mathfrak{C}$ be a t-container for $\mathcal{F} \otimes \mathcal{L}(s,k)$. Then (U1) takes the form

$$\forall^* f \in \mathcal{F} \quad \forall^* \ell \in \mathcal{L}(s,k) \quad \forall^* \lambda \in \mathcal{L}(t,s) \quad \forall^* z \in \mathbb{F}^s \qquad (f \circ \ell)(z) \in \mathfrak{C}_{f\circ\ell\circ\lambda}.$$

Since $\mathcal{L}(s,k)$ is $(s-1)$-uncontainable and $t \leqslant s-1$, this yields

$$\forall^* f \in \mathcal{F} \quad \forall^* \ell \in \mathcal{L}(s,k) \quad \forall^* \lambda \in \mathcal{L}(t,s) \quad \forall^* y \in \mathbb{F}^k \qquad f(y) \in \mathfrak{C}_{f\circ\ell\circ\lambda},$$

which, by Lemma 4.10, is equivalent to

$$\forall^* f \in \mathcal{F} \quad \forall^* \varphi \in \mathcal{L}(t,k) \quad \forall^* y \in \mathbb{F}^k \qquad f(y) \in \mathfrak{C}_{f\circ\varphi}.$$

Since $\mathcal{F}$ is r-uncontainable and $t \leqslant r$, we conclude that

$$\forall^* f \in \mathcal{F} \quad \forall^* \varphi \in \mathcal{L}(t,k) \quad \forall^* x \in \mathbb{F}^n \qquad x \in \mathfrak{C}_{f\circ\varphi},$$

which, by Lemma 4.10 again, can be rewritten as

$$\forall^* f \in \mathcal{F} \quad \forall^* \ell \in \mathcal{L}(s,k) \quad \forall^* \lambda \in \mathcal{L}(t,s) \quad \forall^* x \in \mathbb{F}^n \qquad x \in \mathfrak{C}_{f\circ\ell\circ\lambda}.$$

But the last expression is exactly the relevant instance of (U2). ∎

4.4. Typical dimension of neighborhoods

Let $A \subseteq \mathbb{F}^n \times \mathbb{F}^m$ be a definable set. We say that A is *dominant* if $\dim \operatorname{proj}_1 A = n$. For a definable map $\varphi\colon \mathbb{F}^k \dashrightarrow \mathbb{F}^n$, the *fiber of A over φ* is the set $A_\varphi \subseteq \mathbb{F}^k \times \mathbb{F}^m$ given by

$$(y_1, x_2) \in A_\varphi \ :\Longleftrightarrow\ (\varphi(y_1), x_2) \in A.$$

The following lemma is a precise statement of the corollary of the dimension of fibers theorem discussed in the beginning of Section 4.1:

Lemma 4.14. *Let $A \subseteq \mathbb{F}^n \times \mathbb{F}^m$ be a dominant definable set and let $\mathcal{F}\colon \mathbb{F}^k \rightrightarrows \mathbb{F}^n$. If $\mathcal{F}$ is comprehensive, then, for generic $f \in \mathcal{F}$, the set $A_f \subseteq \mathbb{F}^k \times \mathbb{F}^m$ is dominant and the dimension of every dominant irreducible component of A_f is $\dim A - n + k$.*

Note that, by Lemma 3.5, the set of all elements $f \in \mathcal{F}$ that satisfy the conclusion of Lemma 4.14 is definable.

Proof. Let $C \subseteq \mathbb{F}^n$ be the set of all $x \in \mathbb{F}^n$ such that $A_x \neq \varnothing$ and the dimension of every irreducible component of A_x is $\dim A - n$. Since A is dominant, by Theorem 3.7, $\dim C = n$. Applying the implication (C2) $\Rightarrow$ (C1), we obtain that for generic $f \in \mathcal{F}$ and

for generic $y \in \mathbb{F}^k$, $f(y) \in C$; in other words, $(A_f)_y \neq \varnothing$ and the dimension of every irreducible component of $(A_f)_y$ is $\dim A - n$. Since $(A_f)_y \neq \varnothing$ for generic $y \in \mathbb{F}^k$, we conclude that A_f is dominant. Let $B \subseteq A_f$ be any dominant component of A_f. For generic $y \in \mathbb{F}^k$, we have $\dim B_y = \dim B - k$. Since $B_y \subseteq (A_f)_y$, this implies that $\dim B \leqslant \dim A - n + k$. Suppose that $\dim B < \dim A - n + k$. Then, for generic $y \in \mathbb{F}^k$, $\dim B_y$ is less than the dimension of every irreducible component of $(A_f)_y$, and hence the set B_y is contained in the closure of $(A_f)_y \setminus B_y$. This means that $B \subseteq \overline{A_f \setminus B}$, contradicting the fact that B is a component of A_f. ∎

We are now ready to state and prove the central result of this section:

Theorem 4.15. *Let $A \subseteq \mathbb{F}^n \times \mathbb{F}^m$ be a dominant irreducible definable set and suppose that $\mathcal{F}: \mathbb{F}^k \rightrightarrows \mathbb{F}^n$ is an irreducible definable family of functions. Let*

$$r := \min\{k, \dim \mathrm{proj}_2 A + n - \dim A\} - 1.$$

If $\mathcal{F}$ is r-uncontainable, then, for generic $f \in \mathcal{F}$, every dominant component B of A_f satisfies

$$\dim \mathrm{proj}_2 B = \min\{\dim A - n + k, \dim \mathrm{proj}_2 A\}.$$

Proof. The proof is by induction on k. The base case $k = 0$ is a restatement of Theorem 3.7, so suppose $k \geqslant 1$. Let $\mathcal{F}: \mathbb{F}^k \rightrightarrows \mathbb{F}^n$ be an r-uncontainable irreducible definable family. It follows from Lemma 4.14 that, for generic $f \in \mathcal{F}$, every dominant component B of A_f satisfies

$$\begin{aligned} \dim \mathrm{proj}_2 B &\leqslant \min\{\dim B, \dim \mathrm{proj}_2 A\} \\ &= \min\{\dim A - n + k, \dim \mathrm{proj}_2 A\}. \end{aligned} \tag{4.16}$$

Suppose that the conclusion of Theorem 4.15 fails for $\mathcal{F}$. Due to (4.16) and since $\mathcal{F}$ is irreducible, this means that for generic $f \in \mathcal{F}$, the set A_f has a dominant component B with

$$\dim \mathrm{proj}_2 B < \min\{\dim A - n + k, \dim \mathrm{proj}_2 A\}. \tag{4.17}$$

Consider the family $\mathcal{F} \otimes \mathcal{L}(k-1, k): \mathbb{F}^{k-1} \rightrightarrows \mathbb{F}^n$. Since the families $\mathcal{F}$ and $\mathcal{L}(k-1, k)$ are comprehensive, so is $\mathcal{F} \otimes \mathcal{L}(k-1, k)$ by Proposition 4.7. By Lemma 4.14, this implies that for generic $f \in \mathcal{F}$ and $\ell \in \mathcal{L}(k-1, k)$, every dominant component C of $A_{f \circ \ell}$ satisfies

$$\dim C = \dim A - n + k - 1. \tag{4.18}$$

We can use the inductive assumption and apply Theorem 4.15 to $\mathcal{F} \otimes \mathcal{L}(k-1, k)$. Set

$$s := \min\{k - 1, \dim \mathrm{proj}_2 A + n - \dim A\} - 1.$$

Then $s = \min\{r, k-2\}$, and, since $\mathcal{F}$ is r-uncontainable, Proposition 4.13 yields that $\mathcal{F} \otimes \mathcal{L}(k-1, k)$ is s-uncontainable. Hence, the conclusion of Theorem 4.15 holds for $\mathcal{F} \otimes \mathcal{L}(k-1, k)$; in other words, for generic $f \in \mathcal{F}$ and $\ell \in \mathcal{L}(k-1, k)$, every dominant component C of $A_{f \circ \ell}$ satisfies

$$\dim \mathrm{proj}_2 C = \min\{\dim A - n + k - 1, \dim \mathrm{proj}_2 A\}. \tag{4.19}$$

Claim 4.15.1. *The following statements are valid:*

(a) $\dim \operatorname{proj}_2 A \geqslant \dim A - n + k$, *and hence* $r = k - 1$,

(b) *for generic* $f \in \mathcal{F}$ *and* $\ell \in \mathcal{L}(k-1,k)$, *there exist a dominant component* B *of* A_f *and a dominant component* C *of* $A_{f\circ\ell}$ *such that*

$$\overline{\operatorname{proj}_2 B} = \overline{\operatorname{proj}_2 C}$$

and

$$\dim \operatorname{proj}_2 B = \dim \operatorname{proj}_2 C = \dim A - n + k - 1.$$

Proof. Consider generic $f \in \mathcal{F}$ and $\ell \in \mathcal{L}(k-1,k)$. Let B be an arbitrary dominant component of A_f satisfying (4.17). Recall that, by Lemma 4.14, $\dim B = \dim A - n + k$. Since the family $\mathcal{L}(k-1,k)$ is comprehensive, we may apply Lemma 4.14 to B in order to conclude that the set B_ℓ is dominant and the dimension of every dominant component of B_ℓ is $\dim B - k + (k-1) = \dim A - n + k - 1$. Let B' be any dominant component of B_ℓ. Since $B_\ell \subseteq A_{f\circ\ell}$, there is a (necessarily also dominant) component C of $A_{f\circ\ell}$ such that $B' \subseteq C$. From (4.18), we see that $\dim B' = \dim C$, and hence the closures of B' and C coincide. By Corollary 3.11, projections commute with closures, and therefore

$$\overline{\operatorname{proj}_2 B'} = \overline{\operatorname{proj}_2 \overline{B'}} = \overline{\operatorname{proj}_2 \overline{C}} = \overline{\operatorname{proj}_2 C},$$

so

$$\dim \operatorname{proj}_2 B' = \dim \operatorname{proj}_2 C.$$

Combining this with (4.17) and (4.19), we obtain the following chain of (in)equalities:

$$\begin{aligned} \min\{\dim A - n + k, \dim \operatorname{proj}_2 A\} &> \dim \operatorname{proj}_2 B \geqslant \dim \operatorname{proj}_2 B_\ell \geqslant \dim \operatorname{proj}_2 B' \\ &= \dim \operatorname{proj}_2 C \\ &= \min\{\dim A - n + k - 1, \dim \operatorname{proj}_2 A\}. \end{aligned} \tag{4.20}$$

By comparing the first and the last terms in (4.20), we see that

$$\dim \operatorname{proj}_2 A \geqslant \dim A - n + k$$

and

$$\dim \operatorname{proj}_2 B = \dim \operatorname{proj}_2 C = \dim A - n + k - 1.$$

Since $\operatorname{proj}_2 B$ is irreducible and $\overline{\operatorname{proj}_2 B} \supseteq \overline{\operatorname{proj}_2 C}$, this yields $\overline{\operatorname{proj}_2 B} = \overline{\operatorname{proj}_2 C}$, and we are done. ∎

Now we define a $(k-1)$-container $\mathfrak{C}$ for the family $\mathcal{F}$ as follows: For each $f \in \mathcal{F}$ and $\ell \in \mathcal{L}(k-1,k)$, let $\mathfrak{C}_{f\circ\ell}$ be the set of all $x \in \mathbb{F}^n$ such that the set $A_{f\circ\ell}$ has a dominant component C with

$$\dim \operatorname{proj}_2 C = \dim A - n + k - 1 \quad \text{and} \quad \dim(A_x \cap \operatorname{proj}_2 C) \geqslant \dim A - n.$$

Note that $\mathfrak{C}$ is definable due to Lemma 3.5 (and other results of Section 3.1).

Claim 4.15.2. *The* $(k-1)$*-container* $\mathfrak{C}$ *satisfies* (U1)*; in other words, we have*

$$\forall^* f \in \mathcal{F} \quad \forall^* \ell \in \mathcal{L}(k-1,k) \quad \forall^* y \in \mathbb{F}^k \qquad f(y) \in \mathfrak{C}_{f\circ\ell}.$$

Proof. Pick generic $f \in \mathcal{F}$ and $\ell \in \mathcal{L}(k-1,k)$ and let B and C be dominant components of A_f and $A_{f\circ\ell}$ respectively with the properties specified by Claim 4.15.1 (b). Since $\dim \mathrm{proj}_2 C = \dim A - n + k - 1$, every $y \in \mathbb{F}^k$ with $f(y) \notin \mathfrak{C}_{f\circ\ell}$ must satisfy

$$\dim(B_y \cap \mathrm{proj}_2 C) \leqslant \dim((A_f)_y \cap \mathrm{proj}_2 C) < \dim A - n. \tag{4.21}$$

We claim that (4.21) fails for generic $y \in \mathbb{F}^k$, which gives the desired conclusion. Indeed, we have $\dim B = \dim A - n + k$, so, for generic $y \in \mathbb{F}^k$, $\dim B_y = \dim A - n$, and if such y satisfies (4.21), then

$$\dim(B_y \setminus \mathrm{proj}_2 C) = \dim A - n. \tag{4.22}$$

If (4.22) holds for generic $y \in \mathbb{F}^k$, then

$$\dim(B \setminus (\mathbb{F}^k \times \mathrm{proj}_2 C)) \geqslant \dim A - n + k = \dim B.$$

Since B is irreducible, Corollary 3.11 then yields

$$\dim(\mathrm{proj}_2 B \setminus \mathrm{proj}_2 C) = \dim \mathrm{proj}_2(B \setminus (\mathbb{F}^k \times \mathrm{proj}_2 C)) = \dim \mathrm{proj}_2 B,$$

which contradicts the fact that $\overline{\mathrm{proj}_2 B} = \overline{\mathrm{proj}_2 C}$. ∎

Since, by Claim 4.15.1 (a), the family $\mathcal{F}$ is $(k-1)$-uncontainable, we deduce from Claim 4.15.2 that

$$\forall^* f \in \mathcal{F} \quad \forall^* \ell \in \mathcal{L}(k-1,k) \quad \forall^* x \in \mathbb{F}^n \qquad x \in \mathfrak{C}_{f\circ\ell}.$$

Take any $f \in \mathcal{F}$ and $\ell \in \mathcal{L}(k-1,k)$ such that $\forall^* x \in \mathbb{F}^n$ $(x \in \mathfrak{C}_{f\circ\ell})$. The set $A_{f\circ\ell}$ has only finitely many components, so we can choose a dominant component C of $A_{f\circ\ell}$ so that

$$\dim \mathrm{proj}_2 C = \dim A - n + k - 1$$

and

$$\forall^* x \in \mathbb{F}^n \, (\dim(A_x \cap \mathrm{proj}_2 C) \geqslant \dim A - n).$$

Then we have

$$\dim(A \cap (\mathbb{F}^n \times \mathrm{proj}_2 C)) \geqslant \dim A,$$

and, since A is irreducible, Corollary 3.11 yields

$$\dim \mathrm{proj}_2 C = \dim \mathrm{proj}_2(A \cap (\mathbb{F}^n \times \mathrm{proj}_2 C)) = \dim \mathrm{proj}_2 A.$$

But by Claim 4.15.1 (a), $\dim \mathrm{proj}_2 A \geqslant \dim A - n + k > \dim \mathrm{proj}_2 C$; a contradiction. ∎

To establish a connection between Theorem 4.15 and the motivating discussion in Section 4.1, consider the contrapositive of Theorem 4.15. Let $A \subseteq \mathbb{F}^n \times \mathbb{F}^m$ be a dominant irreducible definable set and suppose that $\mathcal{F} : \mathbb{F}^k \rightrightarrows \mathbb{F}^n$ is an irreducible definable family consisting of functions f that *violate* the conclusion of Theorem 4.15. Such a family $\mathcal{F}$ cannot be r-uncontainable, and hence, by Proposition 4.12, there exists an r-container $\mathfrak{C}$ for $\mathcal{F}$ such that for generic $f \in \mathcal{F}$ and $\ell \in \mathcal{L}(r,k)$, we have

$$\dim \mathfrak{C}_{f\circ\ell} < n, \quad \text{yet} \quad \dim f^{-1}(\mathfrak{C}_{f\circ\ell}) = k.$$

In other words, based on $f \circ \ell$ alone, we can definably build a "small" set $\mathfrak{C}_{f\circ\ell}$ that contains "most" of $f(\mathbb{F}^k)$. This result can be further strengthened in two ways: first, we can actually do this for *all* $f \in \mathcal{F}$ (and not just for a *generic* subset); second, the family $\mathcal{F}$ need not be irreducible:

Corollary 4.23. *Let $A \subseteq \mathbb{F}^n \times \mathbb{F}^m$ be a dominant irreducible definable set and let*

$$r := \min\{k, \dim \operatorname{proj}_2 A + n - \dim A\} - 1.$$

Let $\mathcal{F} : \mathbb{F}^k \rightrightarrows \mathbb{F}^n$ be a definable family of functions such that for all elements $f \in \mathcal{F}$, $\dim \operatorname{dom}(f) = k$. Suppose that for every $f \in \mathcal{F}$, either the set A_f is not dominant, or else, A_f has a dominant component B with

$$\dim \operatorname{proj}_2 B \neq \min\{\dim A - n + k, \dim \operatorname{proj}_2 A\}.$$

Then there is an r-container $\mathfrak{C}$ for $\mathcal{F}$ such that for all $f \in \mathcal{F}$ and for generic $\ell \in \mathcal{L}(r,k)$, we have

$$\dim \mathfrak{C}_{f\circ\ell} < n, \quad \textit{yet} \quad \dim f^{-1}(\mathfrak{C}_{f\circ\ell}) = k.$$

Proof. For fixed $A \subseteq \mathbb{F}^n \times \mathbb{F}^m$ and k, let $\mathcal{F} : \mathbb{F}^k \rightrightarrows \mathbb{F}^n$ be a counterexample to Corollary 4.23 that minimizes $\dim \mathcal{F}$ and has the fewest irreducible components among all counterexamples with dimension equal to $\dim \mathcal{F}$. Clearly, $\mathcal{F} \neq \varnothing$. Since for all $f \in \mathcal{F}$, $\dim \operatorname{dom}(f) = k$, every definable subset of $\mathcal{F}$ forms a definable family of functions with the evaluation map inherited from $\mathcal{F}$. Call a definable subfamily $\mathcal{F}' \subseteq \mathcal{F}$ *small* if there is an r-container $\mathfrak{C}$ for $\mathcal{F}'$ such that for all $f \in \mathcal{F}'$ and for generic $\ell \in \mathcal{L}(r,k)$, we have $\dim \mathfrak{C}_{f\circ\ell} < n$ but $\dim f^{-1}(\mathfrak{C}_{f\circ\ell}) = k$. By assumption, $\mathcal{F}$ itself is not small.

Claim 4.23.1. *If $\mathcal{F}_1, \mathcal{F}_2 \subseteq \mathcal{F}$ are small definable subfamilies, then so is $\mathcal{F}_1 \cup \mathcal{F}_2$.*

Proof. For each $i \in [2]$, let $\mathfrak{C}_i$ be an r-container for $\mathcal{F}_i$ such that for all $f \in \mathcal{F}_i$ and for generic $\ell \in \mathcal{L}(r,k)$, we have $\dim(\mathfrak{C}_i)_{f\circ\ell} < n$ and $\dim f^{-1}((\mathfrak{C}_i)_{f\circ\ell}) = k$. We may in fact assume that $\dim(\mathfrak{C}_i)_{f\circ\ell} < n$ for *all* $\ell \in \mathcal{L}(r,k)$, since otherwise we can replace $\mathfrak{C}_i$ with $\mathfrak{C}_i'$ given by

$$(\mathfrak{C}_i')_{f\circ\ell} := \begin{cases} (\mathfrak{C}_i)_{f\circ\ell} & \text{if } \dim(\mathfrak{C}_i)_{f\circ\ell} < n, \\ \varnothing & \text{otherwise.} \end{cases}$$

Define an r-container $\mathfrak{C}$ for $\mathcal{F}_1 \cup \mathcal{F}_2$ by setting $\mathfrak{C}_{f\circ\ell} := (\mathfrak{C}_1)_{f\circ\ell} \cup (\mathfrak{C}_2)_{f\circ\ell}$. It is easy to see that $\mathfrak{C}$ has all the desired properties. ∎

We claim that $\mathcal{F}$ is irreducible. Otherwise, we can write $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$, where each of $\mathcal{F}_1, \mathcal{F}_2$ is a definable family of functions that has fewer irreducible components than $\mathcal{F}$. By the choice of $\mathcal{F}$, the families $\mathcal{F}_1$ and $\mathcal{F}_2$ are small, and hence so is $\mathcal{F}$ by Claim 4.23.1, which is a contradiction.

If $\mathcal{F}$ were r-uncontainable, we would be able to use Lemma 4.14 and Theorem 4.15 to conclude that for generic $f \in \mathcal{F}$, the set A_f is dominant and every dominant component B of A_f satisfies

$$\dim \operatorname{proj}_2 B = \min\{\dim A - n + k, \dim \operatorname{proj}_2 A\}.$$

This conclusion contradicts the assumptions on $\mathcal{F}$, so $\mathcal{F}$ cannot be r-uncontainable. From Proposition 4.12, it follows that there exists an r-container $\mathfrak{C}$ for $\mathcal{F}$ that satisfies (U1) but fails (U2); in other words, for generic $f \in \mathcal{F}$ and $\ell \in \mathcal{L}(r,k)$, we have $\dim f^{-1}(\mathfrak{C}_{f\circ\ell}) = k$ but $\dim \mathfrak{C}_{f\circ\ell} < n$. Let

$$\mathcal{F}_1 := \{f \in \mathcal{F} : \forall^* \ell \in \mathcal{L}(r,k)\,(\dim f^{-1}(\mathfrak{C}_{f\circ\ell}) = k \text{ and } \dim \mathfrak{C}_{f\circ\ell} < n)\}$$

and

$$\mathcal{F}_2 := \mathcal{F} \setminus \mathcal{F}_1.$$

The r-container $\mathfrak{C}$ certifies that $\mathcal{F}_1$ is small. But $\dim \mathcal{F}_2 < \dim \mathcal{F}$, so $\mathcal{F}_2$ is also small by the choice of $\mathcal{F}$. Claim 4.23.1 then implies that $\mathcal{F}$ itself is small – a contradiction. ∎

5. Proof of Theorem 2.7

5.1. *Almost density is necessary*

The following simple linear-algebraic fact will be rather useful (it is partially responsible for the lower bound $d \geqslant t-1$ in the statement of Theorem 2.7):

Lemma 5.1. *Let $t \geqslant 1$. Fix $d \geqslant t-1$ and a nonzero polynomial $q \in \mathbb{F}[x_1,\dots,x_k]$. Let $y_1,\dots,y_t \in \mathbb{F}^k$ be a sequence of pairwise distinct points such that $q(y_i) \neq 0$ for all $i \in [t]$. Then, for any $x_1,\dots,x_t \in \mathbb{F}^n$,*

$$\begin{aligned}\dim\{f \in \mathcal{R}_d(k,n;q) : f(y_1) = x_1,\dots,f(y_t) = x_t\} &= \dim \mathcal{R}_d(k,n;q) - tn\\ &= \left(\binom{k+d}{d} - t\right)n.\end{aligned}$$

Proof. We can express each element $f \in \mathcal{R}_d(k,n;q)$ as $f = (p_1/q,\dots,p_n/q)$, where $p_1,\dots,p_n$ are polynomials of degree at most d. For each $i \in [t]$, write

$$x_i = (x_i(1),\dots,x_i(n)) \quad \text{and} \quad y_i = (y_i(1),\dots,y_i(k)).$$

The conditions $f(y_1) = x_1,\dots,f(y_t) = x_t$ are then equivalent to

$$p_j(y_i) = q(y_i)\cdot x_i(j) \quad \text{for all } i \in [t] \text{ and } j \in [n]. \tag{5.2}$$

This is a system of tn linear equations in the coefficients of the polynomials $p_1,\dots,p_n$. Therefore, the statement of Lemma 5.1 is equivalent to the assertion that equations (5.2) are independent. Furthermore, it is enough to establish the independence of the equations corresponding to the same $j \in [n]$, since the equations corresponding to distinct j share no common variables. Since the rows of a matrix M are linearly independent if and only if the linear operator determined by M is surjective, it remains to show that for all $b_1,\dots,b_t \in \mathbb{F}$, there exists a polynomial p of degree at most d such that

$$p(y_i) = b_i \quad \text{for all } i \in [t]. \tag{5.3}$$

For each $i \in [t] \setminus \{1\}$, choose some $j_i \in [k]$ such that $y_i(j_i) \neq y_1(j_i)$ and let

$$q_1(v_1,\dots,v_k) := \frac{(v_{j_2} - y_2(j_2))\cdots(v_{j_k} - y_k(j_k))}{(y_1(j_2) - y_2(j_2))\cdots(y_1(j_k) - y_k(j_k))} \in \mathbb{F}[v_1,\dots,v_k].$$

Then q_1 is a polynomial of degree $t-1 \leqslant d$ such that $q_1(y_1) = 1$ and $q_1(y_i) = 0$ for all $i \in [t] \setminus \{1\}$. Similarly, there exist polynomials $q_2, \ldots, q_t$ such that $q_i(y_i) = 1$ and $q_j(y_i) = 0$ for $i \neq j$. Clearly, the polynomial $p := b_1 q_1 + \cdots + b_t q_t$ is a solution to equation (5.3), as desired. ∎

Recall that we call a set $E \subseteq (\mathbb{F}^n)^t$ injective if for all $(x_1, \ldots, x_t) \in E$, the elements $x_1, \ldots, x_t$ are pairwise distinct.

Lemma 5.4. *Let $t \geqslant 1$ and let $E \subseteq (\mathbb{F}^n)^t$ be an injective definable set. Suppose that for some $d \geqslant t-1$ and $q \in \mathbb{F}[x_1, \ldots, x_k] \setminus \{0\}$, we have $\forall^* f \in \mathcal{R}_d(k,n;q)\,(\dim E[f] \geqslant k)$. Then $\dim E \geqslant tn - (t-1)k$.*

Proof. Consider the set

$$\mathfrak{G} := \{(f; y_1, \ldots, y_t; x_1, \ldots, x_t) \in \mathcal{R}_d(k,n;q) \times (\mathbb{F}^k)^t \times E : f(y_1) = x_1, \ldots, f(y_t) = x_t\}.$$

By definition, $(y_1, \ldots, y_t) \in E[f]$ when for some $(x_1, \ldots, x_t) \in E$, we have

$$(f; y_1, \ldots, y_t; x_1, \ldots, x_t) \in \mathfrak{G}.$$

Since $\forall^* f \in \mathcal{R}_d(k,n;q)\,(\dim E[f] \geqslant k)$, this implies that

$$\dim \mathfrak{G} \geqslant \dim \mathcal{R}_d(k,n;q) + k. \tag{5.5}$$

On the other hand, consider any $(f; y_1, \ldots, y_t; x_1, \ldots, x_t) \in \mathfrak{G}$. By definition, we have $\{y_1, \ldots, y_t\} \subseteq \mathrm{dom}(f)$, i.e., $q(y_i) \neq 0$ for all $i \in [t]$. Since E is injective, the points $x_1, \ldots, x_t$, and hence also $y_1, \ldots, y_t$, are pairwise distinct. By Lemma 5.1, if we choose any $(x_1, \ldots, x_t) \in E$ and any sequence $y_1, \ldots, y_t \in \mathbb{F}^k$ of pairwise distinct points such that $q(y_i) \neq 0$ for all $i \in [t]$, then

$$\dim\{f \in \mathcal{R}_d(k,n;q) : (f; y_1, \ldots, y_t; x_1, \ldots, x_t) \in \mathfrak{G}\} = \dim \mathcal{R}_d(k,n;q) - tn.$$

Therefore,

$$\dim \mathfrak{G} = \dim E + tk + \dim \mathcal{R}_d(k,n;q) - tn. \tag{5.6}$$

Combining (5.5) and (5.6) yields the desired conclusion. ∎

With Lemma 5.4 in hand, we derive Proposition 2.6, restated below for the reader's convenience:

Proposition 5.7. *Let $t \geqslant 1$ and let $E \subseteq (\mathbb{F}^n)^t$ be an injective definable set. If there exist $d \geqslant t-1$ and a nonzero polynomial $q \in \mathbb{F}[x_1, \ldots, x_k]$ such that, for generic element $f \in \mathcal{R}_d(k,n;q)$, every definable $E[f]$-independent set $I \subseteq \mathbb{F}^k$ has dimension less than k, then E has a k-almost dense irreducible component.*

Proof. First we show that E itself is k-almost dense. Let $\varnothing \neq S \subseteq [t]$. Since every $(\mathrm{proj}_S E)[f]$-independent set is also $E[f]$-independent, we conclude that for generic element $f \in \mathcal{R}_d(k,n;q)$, every definable $(\mathrm{proj}_S E)[f]$-independent set $I \subseteq \mathbb{F}^k$ has dimension less than k. By Proposition 2.2, this implies that $\dim(\mathrm{proj}_S E)[f] \geqslant k$. Then, by Lemma 5.4, $\dim \mathrm{proj}_S E \geqslant |S|n - (|S|-1)k$, as desired.

Now suppose that the components of E are $H_1, \dots, H_s$, and consider any element $f \in \mathcal{R}_d(k,n;q)$. Notice that

$$E[f] = H_1[f] \cup \cdots \cup H_s[f],$$

and thus if there is no k-dimensional definable $E[f]$-independent set, then there is also no k-dimensional definable $H_i[f]$-independent set for some $i \in [s]$ (as the intersection of finitely many dense definable sets is dense). Since the set $\mathcal{R}_d(k,n;q)$ is irreducible, there is some $i \in [s]$ such that for generic element $f \in \mathcal{R}_d(k,n;q)$, every definable $H_i[f]$-independent set has dimension less than k. By the above argument, H_i is k-almost dense, and we are done. ∎

Lemma 5.1 will be used again in Sections 5.3 and 5.4.

5.2. *Restrictions*

Fix a nonzero polynomial $q \in \mathbb{F}[x_1, \dots, x_k]$. Given an element $g \in \mathcal{R}_d(k,n;q)$ and a subset $L \subseteq \mathcal{L}_{<k}(k)$, we use $\mathcal{R}_d(k,n;q)[g;L]$ to denote the set of all $f \in \mathcal{R}_d(k,n;q)$ such that

$$f \circ \ell = g \circ \ell \quad \text{for all } \ell \in L.$$

We view $\mathcal{R}_d(k,n;q)[g;L]$ as a definable family of functions, with the evaluation map inherited from $\mathcal{R}_d(k,n;q)$, and call it the $(g;L)$-*restriction* of $\mathcal{R}_d(k,n;q)$. Note that, by definition, $\mathcal{R}_d(k,n;q)[g;L] \neq \varnothing$ as $g \in \mathcal{R}_d(k,n;q)[g;L]$. As the set $\mathcal{R}_d(k,n;q)[g;L]$ is cut out from $\mathcal{R}_d(k,n;q)$ by linear equations, it is irreducible.

The next proposition allows us to apply the results of Section 4 to $\mathcal{R}_d(k,n;q)[g;L]$ (it is another reason for the lower bound $d \geqslant t-1$ in Theorem 2.7):

Proposition 5.8. *Fix $q \in \mathbb{F}[x_1, \dots, x_k] \setminus \{0\}$. Let $g \in \mathcal{R}_d(k,n;q)$ and $L \subseteq \mathcal{L}_{<k}(k)$.*

(a) *If $|L| \leqslant d$, then the family $\mathcal{R}_d(k,n;q)[g;L]$ is comprehensive.*

(b) *If $|L| < d$, then the family $\mathcal{R}_d(k,n;q)[g;L]$ is $(k-1)$-uncontainable.*

Proof. (a) Assume that $|L| \leqslant d$ and suppose that $\mathcal{R}_d(k,n;q)[g;L]$ is not comprehensive. Due to Proposition 4.5, there is a definable set $C \subset \mathbb{F}^n$ such that $\dim C < n$ but

$$\forall^* f \in \mathcal{R}_d(k,n;q)[g;L] \quad \forall^* y \in \mathbb{F}^k \qquad f(y) \in C. \tag{5.9}$$

We may replace C by its closure and assume that C is closed, i.e., it is the set of common zeros of a family of n-variable polynomials. Let $Z \subset \mathbb{F}^k$ be the zero locus of q. Since C is closed, the set

$$\{(f,y) \in \mathcal{R}_d(k,n;q)[g;L] \times (\mathbb{F}^k \setminus Z) : f(y) \in C\}$$

is cut out from $\mathcal{R}_d(k,n;q)[g;L] \times (\mathbb{F}^k \setminus Z)$ by a family of polynomial equations; in other words, it is relatively closed in $\mathcal{R}_d(k,n;q)[g;L] \times (\mathbb{F}^k \setminus Z)$. Since it is dense in $\mathcal{R}_d(k,n;q)[g;L] \times (\mathbb{F}^k \setminus Z)$ by (5.9), we can replace generic quantifiers in (5.9) by universal ones and conclude that

$$\forall f \in \mathcal{R}_d(k,n;q)[g;L] \quad \forall y \in \mathbb{F}^k \setminus Z \qquad f(y) \in C.$$

To arrive at a contradiction, we shall exhibit $f \in \mathcal{R}_d(k,n;q)[g;L]$ and $y \in \mathbb{F}^k \setminus Z$ such that $f(y) \notin C$. Actually, for any given $x \in \mathbb{F}^n$, we will find $f \in \mathcal{R}_d(k,n;q)[g;L]$ and $y \in \mathbb{F}^k \setminus Z$ with $f(y) = x$.

Let $L =: \{\ell_1, \dots, \ell_s\}$, where $s \leqslant d$. Since each function ℓ_i is defined on a space of dimension less than k, we can choose nonconstant affine maps $\varphi_i \colon \mathbb{F}^k \to \mathbb{F}$ such that $\operatorname{im}(\ell_i) \subseteq \ker(\varphi_i)$. Let

$$p \coloneqq \varphi_1 \cdots \varphi_s.$$

Then p is a nonzero polynomial in k variables of degree $s \leqslant d$ such that $p \circ \ell_i = 0$ for all $i \in [s]$. For any sequence of coefficients $a_1, \dots, a_n \in \mathbb{F}$, define a rational map $f_{a_1,\dots,a_n} \in \mathcal{R}(k,n)$ via

$$f_{a_1,\dots,a_n} \coloneqq g + \left(\frac{a_1 p}{q}, \dots, \frac{a_n p}{q}\right).$$

By the choice of p, we have $f_{a_1,\dots,a_n} \in \mathcal{R}_d(k,n;q)[g;L]$. Furthermore, if $y \in \mathbb{F}^k \setminus Z$ is such that $p(y) \neq 0$, then, by varying $a_1, \dots, a_n$, we can force $f_{a_1,\dots,a_n}(y)$ to take an arbitrary value $x \in \mathbb{F}^n$, as desired.

(b) Assume that $|L| < d$ and let $\mathfrak{C}$ be a $(k-1)$-container for $\mathcal{R}_d(k,n;q)[g;L]$. Then (U1) becomes

$$\forall^* f \in \mathcal{R}_d(k,n;q)[g;L] \quad \forall^* \ell \in \mathcal{L}(k-1,k) \quad \forall^* y \in \mathbb{F}^k \qquad f(y) \in \mathfrak{C}_{f\circ\ell},$$

or, equivalently,

$$\forall^* \ell \in \mathcal{L}(k-1,k) \quad \forall^* f \in \mathcal{R}_d(k,n;q)[g;L] \quad \forall^* y \in \mathbb{F}^k \qquad f(y) \in \mathfrak{C}_{f\circ\ell}. \tag{5.10}$$

For $\ell \in \mathcal{L}(k-1,k)$, define an equivalence relation R_ℓ on $\mathcal{R}_d(k,n;q)[g;L]$ by

$$(f,h) \in R_\ell \ :\Longleftrightarrow\ f \circ \ell = h \circ \ell.$$

By definition, the R_ℓ-equivalence class of f is the set $\mathcal{R}_d(k,n;q)[f;L \cup \{\ell\}]$. Since R_ℓ is cut out from $\mathcal{R}_d(k,n;q)[g;L] \times \mathcal{R}_d(k,n;q)[g;L]$ by linear equations, it is irreducible, and since R_ℓ is reflexive, we have $\operatorname{proj}_1 R_\ell = \operatorname{proj}_2 R_\ell = \mathcal{R}_d(k,n;q)[g;L]$. Thus, we may apply implication (ii) $\Rightarrow$ (i) of Corollary 3.17 to rewrite (5.10) as follows:

$$\begin{aligned} &\forall^* \ell \in \mathcal{L}(k-1,k) \quad \forall^* f \in \mathcal{R}_d(k,n;q)[g;L] \\ &\qquad \forall^* h \in \mathcal{R}_d(k,n;q)[f;L \cup \{\ell\}] \quad \forall^* y \in \mathbb{F}^k \qquad h(y) \in \mathfrak{C}_{h\circ\ell}. \end{aligned} \tag{5.11}$$

Note that if $(f,h) \in R_\ell$, then $h \circ \ell = f \circ \ell$, and hence $\mathfrak{C}_{h\circ\ell} = \mathfrak{C}_{f\circ\ell}$. Also, since $|L| < d$, by part (a), the family $\mathcal{R}_d(k,n;q)[f;L \cup \{\ell\}]$ is comprehensive. Therefore, we have

$$\begin{aligned} &\forall^* h \in \mathcal{R}_d(k,n;q)[f;L \cup \{\ell\}] \quad \forall^* y \in \mathbb{F}^k \quad h(y) \in \mathfrak{C}_{h\circ\ell} \\ &\iff \forall^* h \in \mathcal{R}_d(k,n;q)[f;L \cup \{\ell\}] \quad \forall^* y \in \mathbb{F}^k \quad h(y) \in \mathfrak{C}_{f\circ\ell} \qquad [\mathfrak{C}_{h\circ\ell} = \mathfrak{C}_{f\circ\ell}] \\ &\iff \forall^* x \in \mathbb{F}^n \quad x \in \mathfrak{C}_{f\circ\ell} \qquad [\text{comprehensiveness}]. \end{aligned}$$

Hence, (5.11) is equivalent to

$$\forall^* \ell \in \mathcal{L}(k-1,k) \quad \forall^* f \in \mathcal{R}_d(k,n;q)[g;L] \quad \forall^* x \in \mathbb{F}^n \qquad x \in \mathfrak{C}_{f\circ\ell},$$

which turns into (U2) after switching the order of the first two quantifiers. ∎

5.3. Iterative applications of Theorem 4.15

Let $t \geqslant 1$ and let $E \subseteq (\mathbb{F}^n)^t$ be a definable set. Let $s \leqslant t-1$. Given $f \in \mathcal{R}(k,n)$ and affine maps $\ell_1, \dots, \ell_s \in \mathcal{L}(r,k)$, define $E(f;\ell_1,\dots,\ell_s) \subseteq (\mathbb{F}^n)^{t-s}$ as the set of all tuples $(x_1,\dots,x_{t-s}) \in (\mathbb{F}^n)^{t-s}$ such that

$$\exists z_1,\dots,z_s \in \mathbb{F}^r \quad ((f\circ\ell_1)(z_1),\dots,(f\circ\ell_s)(z_s),\, x_1,\dots,x_{t-s}) \in E.$$

For $s=0$, we set $E(f;\varnothing) := E$. It is clear from this definition that for $s \geqslant 1$,

$$E(f;\ell_1,\dots,\ell_s) = E(f;\ell_1,\dots,\ell_{s-1})(f;\ell_s).$$

We view $E(f;\ell_1,\dots,\ell_s)$ as a subset of the product space $(\mathbb{F}^n)^{t-s}$ with the coordinates indexed by $[t-s]$.

Lemma 5.12. *Let $t \geqslant 1$ and let $E \subseteq (\mathbb{F}^n)^t$ be an injective r-almost dense irreducible definable set. Fix $d \geqslant t-1$ and a nonzero polynomial $q \in \mathbb{F}[x_1,\dots,x_k]$. If $k \geqslant r+1$ and $s \leqslant t-1$, then, for generic elements $f \in \mathcal{R}_d(k,n;q)$ and $\ell_1,\dots,\ell_s \in \mathcal{L}(r,k)$, the subset $E(f;\ell_1,\dots,\ell_s)$ of $(\mathbb{F}^n)^{t-s}$ is nonempty and all its irreducible components are r-almost dense.*

Proof. The proof is by induction on s. The base case $s=0$ is trivial, so assume $s \geqslant 1$. We first show that the set $E(f;\ell_1,\dots,\ell_s)$ itself is r-almost dense (and, in particular, nonempty). The same argument as in the proof of Proposition 5.8 (b) shows that the sequence of quantifiers

$$\forall^* f \in \mathcal{R}_d(k,n;q) \quad \forall^* \ell_1,\dots,\ell_s \in \mathcal{L}(r,k)$$

can be replaced by

$$\forall^* g \in \mathcal{R}_d(k,n;q) \quad \forall^* \ell_1,\dots,\ell_{s-1} \in \mathcal{L}(r,k)$$
$$\forall^* f \in \mathcal{R}_d(k,n;q)[g;\{\ell_1,\dots,\ell_{s-1}\}] \quad \forall^* \ell_s \in \mathcal{L}(r,k).$$

Consider generic $g \in \mathcal{R}_d(k,n;q)$ and $\ell_1,\dots,\ell_{s-1} \in \mathcal{L}(r,k)$ and let

$$\mathcal{F} := \mathcal{R}_d(k,n;q)[g;\{\ell_1,\dots,\ell_{s-1}\}] \quad \text{and} \quad E' := E(g;\ell_1,\dots,\ell_{s-1}).$$

Notice that if $f \in \mathcal{F}$, then $E(f;\ell_1,\dots,\ell_{s-1}) = E'$, and thus for any $\ell_s \in \mathcal{L}(r,k)$,

$$E(f;\ell_1,\dots,\ell_s) = E'(f;\ell_s).$$

By the inductive assumption, $E' \neq \varnothing$ and every irreducible component of E' is r-almost dense. Let H be any component of E' (so H is r-almost dense) and let $\varnothing \neq S \subseteq [t-s]$. Set $S' := \{i+1 : i \in S\} \subset [t-s+1]$. Then for each $\ell_s \in \mathcal{L}(r,k)$, we have

$$\mathrm{proj}_S(E(f;\ell_1,\dots,\ell_s)) = \mathrm{proj}_S(E'(f;\ell_s)) \supseteq \mathrm{proj}_S(H(f;\ell_s)) = (\mathrm{proj}_{\{1\}\cup S'}H)(f;\ell_s).$$

Define $A := \mathrm{proj}_{\{1\}\cup S'}H$. We can view A as a subset of the product space $\mathbb{F}^n \times (\mathbb{F}^n)^{S'}$. Then

$$A(f;\ell_s) = \mathrm{proj}_2(A_{f\circ\ell_s}),$$

where $A_{f\circ\ell_s}$ denotes the fiber of A over $f\circ\ell_s$ (as defined in Section 4.4). Since H is r-almost dense, we have

$$\dim \operatorname{proj}_1 A = \dim \operatorname{proj}_1 H = n,$$

i.e., the set A is dominant. As $r<k$ and $s-1\leqslant t-2<d$, Proposition 5.8 (b) implies that the family $\mathcal{F}$ is $(k-1)$-uncontainable, and, by Proposition 4.13, the family $\mathcal{F}\otimes\mathcal{L}(r,k)$ is $(r-1)$-uncontainable. Therefore, since A is irreducible, we can apply Lemma 4.14 and Theorem 4.15 to conclude that, for generic $f\in\mathcal{F}$ and $\ell_s\in\mathcal{L}(r,k)$,

$$\begin{aligned}\dim \operatorname{proj}_S(E(f;\ell_1,\dots,\ell_s)) &\geqslant \dim \operatorname{proj}_2(A_{f\circ\ell_s})\\ &\geqslant \min\{\dim A-n+r, \dim \operatorname{proj}_2 A\}.\end{aligned}\tag{5.13}$$

Observe that, since H is r-almost dense,

$$\dim A = \dim \operatorname{proj}_{\{1\}\cup S'} H \geqslant |\{1\}\cup S'|n-(|\{1\}\cup S'|-1)r = (|S|+1)n-|S|r,$$

so $\dim A-n+r\geqslant |S|n-(|S|-1)r$. Similarly,

$$\dim \operatorname{proj}_2 A = \dim \operatorname{proj}_{S'} H \geqslant |S'|n-(|S'|-1)r = |S|n-(|S|-1)r.$$

Thus, (5.13) yields

$$\dim \operatorname{proj}_S(E(f;\ell_1,\dots,\ell_s)) \geqslant |S|n-(|S|-1)r,$$

and hence $E(f;\ell_1,\dots,\ell_s)$ is r-almost dense, as claimed.

To prove that every component of $E(f;\ell_1,\dots,\ell_s)$ is r-almost dense, let $\mathfrak{X}$ be the set of all tuples

$$(f;\ell_1,\dots,\ell_s;x_1,\dots,x_{t-s})\in\mathcal{R}_d(k,n;q)\times(\mathcal{L}(r,k))^s\times(\mathbb{F}^n)^{t-s}$$

such that $(x_1,\dots,x_{t-s})\in E(f;\ell_1,\dots,\ell_s)$.

Claim 5.12.1. *The set $\mathfrak{X}$ is irreducible.*

Proof. Writing $f=(p_1/q,\dots,p_n/q)$, let $\mathfrak{Z}$ be the set of all tuples

$$\begin{aligned}&(f;\ell_1,\dots,\ell_s;z_1,\dots,z_s;y_1,\dots,y_s;x_1,\dots,x_t)\\ &\quad\in\mathcal{R}_d(k,n;q)\times(\mathcal{L}(r,k))^s\times(\mathbb{F}^r)^s\times(\mathbb{F}^k)^s\times(\mathbb{F}^n)^t\end{aligned}$$

such that $\ell_i(z_i)=y_i$ and $(p_1(y_i),\dots,p_n(y_i))=q(y_i)\cdot x_i$ for all $i\in[s]$. Let D denote the set of all tuples $(y_1,\dots,y_s)\in(\mathbb{F}^k)^s$ of pairwise distinct points such that $q(y_i)\neq 0$ for all $i\in[s]$, and let

$$\mathfrak{Y}:=\mathfrak{Z}\cap(\mathcal{R}_d(k,n;q)\times(\mathcal{L}(r,k))^s\times(\mathbb{F}^r)^s\times D\times E).$$

Note that D is irreducible, as $\overline{D}=(\mathbb{F}^k)^s$. Since E is injective, $\mathfrak{X}$ is the image of $\mathfrak{Y}$ under the projection

$$(f;\ell_1,\dots,\ell_s;z_1,\dots,z_s;y_1,\dots,y_s;x_1,\dots,x_t)\mapsto(f;\ell_1,\dots,\ell_s;x_{s+1},\dots,x_t),$$

so it suffices to prove that $\mathfrak{Y}$ is irreducible. The defining equations for $\mathfrak{Z}$ are linear in the coefficients of $p_1, \dots, p_n$ and $\ell_1, \dots, \ell_s$, so we may use Corollary 3.14. The set $(\mathbb{F}^r)^s \times D \times E$ is irreducible, so we just need to show that for all

$$(z_1, \dots, z_s; y_1, \dots, y_s; x_1, \dots, x_t) \in (\mathbb{F}^r)^s \times D \times E,$$

the dimension of the set

$$\{(f; \ell_1, \dots, \ell_s) \in \mathcal{R}_d(k, n; q) \times (\mathcal{L}(r, k))^s : \\ (f; \ell_1, \dots, \ell_s; z_1, \dots, z_s; y_1, \dots, y_s; x_1, \dots, x_t) \in \mathfrak{Z}\}$$

is the same. But since $(y_1, \dots, y_s) \in D$, this is indeed the case by Lemma 5.1, according to which the dimension of this set equals $\dim \mathcal{R}_d(k, n; q) - sn + s(\dim \mathcal{L}(r, k) - k)$. ∎

We can now finish the proof of Lemma 5.12. To this end, take any $\varnothing \neq S \subseteq [t - s]$ and let $\mathfrak{X}_S \subseteq \mathfrak{X}$ be the set such that for all $f \in \mathcal{R}_d(k, n; q)$ and $\ell_1, \dots, \ell_s \in \mathcal{L}(r, k)$, the fiber of $\mathfrak{X}_S$ over $(f; \ell_1, \dots, \ell_s)$ is the union of all the irreducible components H of $E(f; \ell_1, \dots, \ell_s)$ with

$$\dim \operatorname{proj}_S H \geqslant |S|n - (|S| - 1)r. \tag{5.14}$$

It follows from the results of Section 3.1 that $\mathfrak{X}_S$ is definable. We already know that for generic $f, \ell_1, \dots, \ell_s$, the set $E(f; \ell_1, \dots, \ell_s)$ is r-almost dense, and hence

$$\dim \operatorname{proj}_S(E(f; \ell_1, \dots, \ell_s)) \geqslant |S|n - (|S| - 1)r,$$

which means that $E(f; \ell_1, \dots, \ell_s)$ has a component H satisfying (5.14), i.e., the fiber of $\mathfrak{X}_S$ over $(f; \ell_1, \dots, \ell_s)$ is nonempty. Since $\mathfrak{X}$ is irreducible, Corollary 3.12 implies that for generic $f, \ell_1, \dots, \ell_s$, the fiber of $\mathfrak{X}_S$ over $(f; \ell_1, \dots, \ell_s)$ must be equal to $E(f; \ell_1, \dots, \ell_s)$. In other words, every irreducible component H of $E(f; \ell_1, \dots, \ell_s)$ satisfies (5.14), as desired. ∎

Applying Lemma 5.12 with $s = t - 1$ yields the following:

Corollary 5.15. *Let $t \geqslant 1$ and let $E \subseteq (\mathbb{F}^n)^t$ be an injective r-almost dense irreducible definable set. Fix $d \geqslant t - 1$ and a nonzero polynomial $q \in \mathbb{F}[x_1, \dots, x_k]$. If $k \geqslant r + 1$, then*

$$\forall^* f \in \mathcal{R}_d(k, n; q) \quad \forall^* \ell_1, \dots, \ell_{t-1} \in \mathcal{L}(r, k) \quad \forall^* y \in \mathbb{F}^k \\ \exists z_1, \dots, z_{t-1} \in \mathbb{F}^r \qquad (\ell_1(z_1), \dots, \ell_{t-1}(z_{t-1}), y) \in E[f].$$

Proof. We again observe that the sequence of quantifiers

$$\forall^* f \in \mathcal{R}_d(k, n; q) \quad \forall^* \ell_1, \dots, \ell_{t-1} \in \mathcal{L}(r, k)$$

can be replaced by

$$\forall^* g \in \mathcal{R}_d(k, n; q) \quad \forall^* \ell_1, \dots, \ell_{t-1} \in \mathcal{L}(r, k) \quad \forall^* f \in \mathcal{R}_d(k, n; q)[g; \{\ell_1, \dots, \ell_{t-1}\}].$$

Consider generic $g \in \mathcal{R}_d(k, n; q)$ and $\ell_1, \dots, \ell_{t-1} \in \mathcal{L}(r, k)$ and let

$$\mathcal{F} \coloneqq \mathcal{R}_d(k, n; q)[g; \{\ell_1, \dots, \ell_{t-1}\}] \quad \text{and} \quad C \coloneqq E(g; \ell_1, \dots, \ell_{t-1}) \subseteq \mathbb{F}^n.$$

Note that for all $f \in \mathcal{F}$, $E(f;\ell_1,\dots,\ell_{t-1}) = C$. Since $d \geqslant t-1$, Proposition 5.8 (a) shows that $\mathcal{F}$ is comprehensive, and hence we obtain

$$\begin{aligned}
&\forall^* f \in \mathcal{F} \quad \forall^* y \in \mathbb{F}^k \quad \exists z_1,\dots,z_{t-1} \in \mathbb{F}^r \quad (\ell_1(z_1),\dots,\ell_{t-1}(z_{t-1}),y) \in E[f] \\
&\iff \forall^* f \in \mathcal{F} \quad \forall^* y \in \mathbb{F}^k \quad f(y) \in E(f;\ell_1,\dots,\ell_{t-1}) \\
&\iff \forall^* f \in \mathcal{F} \quad \forall^* y \in \mathbb{F}^k \quad f(y) \in C \\
&\iff \forall^* x \in \mathbb{F}^n \quad x \in C \\
&\iff \dim C = n.
\end{aligned}$$

But $\dim C = n$ is precisely the conclusion of Lemma 5.12 for $s = t-1$, so we have proved the corollary. ∎

5.4. *Finishing the proof*

We are finally ready to prove Theorem 2.7. For the reader's convenience, we state it again below:

Theorem 5.16. *Let $t \geqslant 1$ and let $E \subseteq (\mathbb{F}^n)^t$ be an r-almost dense irreducible definable set. Fix $d \geqslant t-1$ and a nonzero polynomial $q \in \mathbb{F}[x_1,\dots,x_k]$. If $k \geqslant r+1$, then, for generic $f \in \mathcal{R}_d(k,n;q)$, the following holds: Every definable $E[f]$-independent set $I \subseteq \mathbb{F}^k$ has dimension less than k. Furthermore, if E is injective, then every irreducible component of $E[f]$ is r-almost dense.*

Proof. Let us first assume that E is injective. As in the proof of Lemma 5.12, we start by showing that for generic $f \in \mathcal{R}_d(k,n;q)$, the set $E[f]$ itself is r-almost dense (hence nonempty). Consider any $\varnothing \neq S \subseteq [t]$. For concreteness, we may assume $S = [s] \cup \{t\}$ for some $s \leqslant t-1$. It follows from Corollary 5.15 that a generic map $f \in \mathcal{R}_d(k,n;q)$ satisfies

$$\begin{aligned}
\forall^* \ell_1,\dots,\ell_s \in \mathcal{L}(r,k) \quad \forall^* y \in \mathbb{F}^k \quad \exists z_1,\dots,z_s \in \mathbb{F}^r \quad \exists y_{s+1},\dots,y_{t-1} \in \mathbb{F}^k \\
(\ell_1(z_1),\dots,\ell_s(z_s),y_{s+1},\dots,y_{t-1},y) \in E[f],
\end{aligned}$$

which can be rewritten as

$$\begin{aligned}
\forall^* \ell_1,\dots,\ell_s \in \mathcal{L}(r,k) \quad \forall^* y \in \mathbb{F}^k \quad \exists z_1,\dots,z_s \in \mathbb{F}^r \\
(\ell_1(z_1),\dots,\ell_s(z_s),y) \in \operatorname{proj}_S(E[f]).
\end{aligned} \tag{5.17}$$

Let $\mathfrak{G}$ be the set of all tuples

$$(\ell_1,\dots,\ell_s;y_1,\dots,y_{s+1};z_1,\dots,z_s) \in (\mathcal{L}(r,k))^s \times \operatorname{proj}_S(E[f]) \times (\mathbb{F}^r)^s$$

such that $\ell_i(z_i) = y_i$ for all $i \in [s]$. From (5.17), we conclude that

$$\dim \mathfrak{G} \geqslant s \dim \mathcal{L}(r,k) + k. \tag{5.18}$$

On the other hand, for all $y \in \mathbb{F}^k$ and $z \in \mathbb{F}^r$, we have

$$\dim\{\ell \in \mathcal{L}(r,k) : \ell(z) = y\} = \dim \mathcal{L}(r,k) - k$$

by Lemma 5.1. Hence, for all $(y_1, \dots, y_s, y_{s+1}) \in \mathrm{proj}_S(E[f])$ and $z_1, \dots, z_s \in \mathbb{F}^r$, the set of all tuples $(\ell_1, \dots, \ell_s) \in (\mathcal{L}(r,k))^s$ with $\ell_1(z_1) = y_1, \dots, \ell_s(z_s) = y_s$ has dimension $s(\dim \mathcal{L}(r,k) - k)$, so

$$\dim \mathscr{G} = \dim \mathrm{proj}_S(E[f]) + sr + s(\dim \mathcal{L}(r,k) - k). \tag{5.19}$$

Comparing (5.18) and (5.19), we obtain

$$\dim \mathrm{proj}_S(E[f]) \geqslant (s+1)k - sr = |S|k - (|S|-1)r,$$

as desired.

To deduce that every irreducible component of $E[f]$ is r-almost dense, we use the same trick as in the proof of Lemma 5.12. Define

$$\mathfrak{X} := \{(f; y_1, \dots, y_t) \in \mathcal{R}_d(k,n;q) \times (\mathbb{F}^k)^t : (y_1, \dots, y_t) \in E[f]\}.$$

Claim 2.7.1. *The set $\mathfrak{X}$ is irreducible.*

Proof. The argument is analogous to the proof of Claim 5.12.1 above. Indeed, writing $f = (p_1/q, \dots, p_n/q)$, let $\mathfrak{Z}$ be the set of all tuples

$$(f; y_1, \dots, y_t; x_1, \dots, x_t) \in \mathcal{R}_d(k,n;q) \times (\mathbb{F}^k)^t \times (\mathbb{F}^n)^t$$

such that $(p_1(y_i), \dots, p_n(y_i)) = q(y_i) \cdot x_i$ for all $i \in [t]$. Let D denote the set of all tuples $(y_1, \dots, y_t) \in (\mathbb{F}^k)^t$ of pairwise distinct points such that $q(y_i) \neq 0$ for all $i \in [t]$, and let

$$\mathfrak{Y} := \mathfrak{Z} \cap (\mathcal{R}_d(k,n;q) \times D \times E).$$

Since E is injective, $\mathfrak{X}$ is the image of $\mathfrak{Y}$ under the projection

$$(f; y_1, \dots, y_t; x_1, \dots, x_t) \mapsto (f; y_1, \dots, y_t),$$

so it suffices to prove that $\mathfrak{Y}$ is irreducible. The defining equations for $\mathfrak{Z}$ are linear in the coefficients of $p_1, \dots, p_n$, so we may use Corollary 3.14. The set $D \times E$ is irreducible, so we just need to show that for all $(y_1, \dots, y_t; x_1, \dots, x_t) \in D \times E$, the dimension of the set

$$\{f \in \mathcal{R}_d(k,n;q) : (f; y_1, \dots, y_t; x_1, \dots, x_t) \in \mathfrak{Z}\}$$

is the same. But since $(y_1, \dots, y_t) \in D$ and $d \geqslant t-1$, this is the case by Lemma 5.1, according to which the dimension of this set is equal to $\dim \mathcal{R}_d(k,n;q) - tn$. ∎

Treating each $f \in \mathcal{R}_d(k,n;q)$ as a point in the space $\mathcal{R}_d(k,n;q)$, we may consider the fiber $\mathfrak{X}_f$ of $\mathfrak{X}$ over f. By definition,

$$\mathfrak{X}_f = \{(y_1, \dots, y_t) \in (\mathbb{F}^k)^t : (f; y_1, \dots, y_t) \in \mathfrak{X}\} = E[f].$$

Now take any $\varnothing \neq S \subseteq [t]$ and let $\mathfrak{X}_S \subseteq \mathfrak{X}$ be the set such that for all $f \in \mathcal{R}_d(k,n;q)$, the fiber

$$(\mathfrak{X}_S)_f = \{(y_1, \dots, y_t) \in (\mathbb{F}^k)^t : (f; y_1, \dots, y_t) \in \mathfrak{X}_S\} \subseteq E[f]$$

is the union of all the irreducible components H of $E[f]$ with

$$\dim \mathrm{proj}_S H \geqslant |S|k - (|S| - 1)r.$$

Then $\mathcal{X}_S$ is definable, and we have shown that for generic $f \in \mathcal{R}_d(k, n; q)$, $(\mathcal{X}_S)_f \neq \varnothing$. Since $\mathcal{X}$ is irreducible, it follows from Corollary 3.12 that for generic $f \in \mathcal{R}_d(k, n; q)$, $(\mathcal{X}_S)_f = E[f]$, i.e., every component H of $E[f]$ satisfies

$$\dim \mathrm{proj}_S H \geqslant |S|k - (|S| - 1)r,$$

as desired. Recall that since, for generic $f \in \mathcal{R}_d(k, n; q)$, the set $E[f]$ is nonempty and all its components are r-almost dense, Observation 2.5 yields that there is no k-dimensional definable $E[f]$-independent set.

Now suppose that E is not necessarily injective. We need to show that for generic element $f \in \mathcal{R}_d(k, n; q)$, every definable $E[f]$-independent set has dimension less than k. Towards a contradiction, let E be a counterexample with the smallest value of t. Let

$$E_0 \coloneqq \{(x_1, \ldots, x_t) \in E : x_1, \ldots, x_t \text{ are pairwise distinct}\},$$

and for $1 \leqslant i < j \leqslant t$, define

$$E_{ij} \coloneqq \{(x_1, \ldots, x_t) \in E : x_i = x_j\}.$$

Since E is irreducible, at least one of the sets E_0, E_{ij}, $1 \leqslant i < j \leqslant t$, is dense in E. If $\overline{E_0} = \overline{E}$, then the set E_0 is irreducible and, by Corollary 3.11, r-almost dense. Since E_0 is injective by definition, we conclude that for generic $f \in \mathcal{R}_d(k, n; q)$, the set $E_0[f]$ does not admit a k-dimensional definable independent set, and hence the same is true for $E[f] \supseteq E_0[f]$. If, on the other hand, $\overline{E_{ij}} = \overline{E}$ for some i and j with $1 \leqslant i < j \leqslant t$, then the set E_{ij} is irreducible and r-almost dense. Let $S \coloneqq [t] \setminus \{j\}$ and $E' \coloneqq \mathrm{proj}_S E_{ij}$. Then E' is also irreducible and r-almost dense, so, by the minimality of t, for generic $f \in \mathcal{R}_d(k, n; q)$, there is no k-dimensional definable $E'[f]$-independent set. But a set $I \subseteq \mathbb{F}^k$ is $E'[f]$-independent if and only if it is $E_{ij}[f]$-independent. Hence, the set $E_{ij}[f]$, and thus also $E[f] \supseteq E_{ij}[f]$, does not admit a k-dimensional definable independent set, and the proof is complete. ∎

6. Further directions

In this paper we worked with hypergraphs definable in an algebraically closed field $\mathbb{F}$. A natural next step would be to study hypergraphs definable in strongly minimal structures, with Morley rank assuming the role of dimension. It seems especially promising to look at strongly minimal structures with the so-called *definable multiplicity property*, which was isolated by Hrushovski in [10], as they satisfy natural analogs of many basic facts stated in Section 3. While it appears likely that most our arguments could be extended to this more general setting without too much difficulty, at certain places we invoke properties that are very special to algebraically closed fields. In particular, the proofs of Claims 5.12.1 and 2.7.1 use Corollary 3.14 and Lemma 3.13, which ultimately

rely on $\mathbb{P}^n$ being a complete variety – a fact that has no obvious analog in arbitrary strongly minimal structures. Nevertheless, we suspect that at least the results of Section 4 should have their counterparts in the strongly minimal setting.

Another structure of interest is the real field $\mathbb{R}$. There our arguments cannot be easily adapted, because, in contrast to algebraically closed fields, $\mathbb{R}$ lacks a well-behaved notion of irreducibility. In particular, the "size" of an $\mathbb{R}$-definable set is characterized not only by its dimension, but also by its *measure*. It makes sense, therefore, to define the *independence ratio* $\alpha(E)$ of an $\mathbb{R}$-definable hypergraph $E \subseteq ([0,1]^n)^t$ as $\alpha(E) := \sup_I \lambda_n(I)$, where λ_n is the n-dimensional Lebesgue measure and the supremum is taken over all $\mathbb{R}$-definable E-independent sets $I \subseteq [0,1]^n$. (We are considering hypergraphs on $[0,1]^n$ to make the total measure of the vertex set 1.) The goal is to isolate the properties of E which guarantee that, for a generic $\mathbb{R}$-definable map $f\colon [0,1]^k \to [0,1]^n$, the induced subhypergraph $E[f] \subseteq ([0,1]^k)^t$ satisfies $\alpha(E[f]) \leqslant \alpha(E)$. A similar problem can also be studied over finite fields $\mathbb{F}_q$, using the (normalized) counting measure instead of the Lebesgue measure.

Acknowledgments. We are very grateful to the anonymous referee for carefully reading the manuscript and providing many useful comments and suggestions.

Funding. Research of the first author is supported in part by NSF grant DMS-2045412. Research of the second author is supported in part by NSERC Discovery Grant No. 2019-04269 and AMS-Simons Travel Grant. Research of the third author is supported in part by NSF Grants DMS-1501036 and DMS-1855648.

References

[1] Alon, N., Pach, J., Pinchasi, R., Radoičić, R., Sharir, M.: Crossing patterns of semi-algebraic sets. J. Combin. Theory Ser. A **111**, 310–326 (2005) Zbl 1099.14048 MR 2156215

[2] Balogh, J., Morris, R., Samotij, W.: Independent sets in hypergraphs. J. Amer. Math. Soc. **28**, 669–709 (2015) Zbl 1310.05154 MR 3327533

[3] Balogh, J., Morris, R., Samotij, W.: The method of hypergraph containers. In: Proceedings of the International Congress of Mathematicians-Rio de Janeiro 2018. Vol. IV. Invited Lectures, World Scientific, Hackensack, 3059–3092 (2018) Zbl 1448.05147 MR 3966523

[4] Bernshteyn, A., Delcourt, M., Towsner, H., Tserunyan, A.: A short nonalgorithmic proof of the containers theorem for hypergraphs. Proc. Amer. Math. Soc. **147**, 1739–1749 (2019) Zbl 07020584 MR 3910438

[5] Chernikov, A., Starchenko, S.: Regularity lemma for distal structures. J. Eur. Math. Soc. (JEMS) **20**, 2437–2466 (2018) Zbl 06966829 MR 3852184

[6] Conlon, D.: Combinatorial theorems relative to a random set. In: Proceedings of the International Congress of Mathematicians – Seoul 2014. Vol. IV, Kyung Moon Sa, Seoul, 303–327 (2014) Zbl 1373.05175 MR 3727614

[7] Conlon, D., Gowers, W. T.: Combinatorial theorems in sparse random sets. Ann. of Math. (2) **184**, 367–454 (2016) Zbl 1351.05204 MR 3548529

[8] Ferber, A., McKinley, G., Samotij, W.: Supersaturated sparse graphs and hypergraphs. Int. Math. Res. Not. IMRN **2020**, no. 2, 378–402 (2020) Zbl 1433.05152 MR 4073190

[9] Fox, J., Gromov, M., Lafforgue, V., Naor, A., Pach, J.: Overlap properties of geometric expanders. J. Reine Angew. Math. **671**, 49–83 (2012) Zbl 1306.05171 MR 2983197

[10] Hrushovski, E.: Strongly minimal expansions of algebraically closed fields. Israel J. Math. **79**, 129–151 (1992) Zbl 0773.12005 MR 1248909

[11] Johnson, W.: Fun with fields. Ph.D. thesis, University of California, Berkeley (2016) MR 3564042

[12] Kleitman, D. J., Winston, K. J.: On the number of graphs without 4-cycles. Discrete Math. **41**, 167–172 (1982) Zbl 0491.05035 MR 676877

[13] Marker, D.: Model Theory. Grad. Texts in Math. 217, Springer, New York (2002) Zbl 1003.03034 MR 1924282

[14] Mumford, D.: The Red Book of Varieties and Schemes. Lecture Notes in Math. 1358, Springer, Berlin (1988) Zbl 0658.14001 MR 971985

[15] Pillay, A.: An Introduction to Stability Theory. Oxford Logic Guides 8, Oxford University, New York (1983) Zbl 0526.03014 MR 719195

[16] Sapozhenko, A.: Systems of containers and enumeration problems. In: Stochastic Algorithms: Foundations and Applications, Lecture Notes in Comput. Sci. 3777, Springer, Berlin, 1–13 (2005) Zbl 1159.68656

[17] Saxton, D., Thomason, A.: Hypergraph containers. Invent. Math. **201**, 925–992 (2015) Zbl 1320.05085 MR 3385638

[18] Schacht, M.: Extremal results for random discrete structures. Ann. of Math. (2) **184**, 333–365 (2016) Zbl 1351.05207 MR 3548528

[19] Suk, A.: Semi-algebraic Ramsey numbers. J. Combin. Theory Ser. B **116**, 465–483 (2016) Zbl 1327.05331 MR 3425253

[20] Szemerédi, E.: On sets of integers containing no k elements in arithmetic progression. Acta Arith. **27**, 199–245 (1975) Zbl 0303.10056 MR 0369312

[21] Tao, T.: Expanding polynomials over finite fields of large characteristic, and a regularity lemma for definable sets. Contrib. Discrete Math. **10**, 22–98 (2015) Zbl 1370.11135 MR 3386249

[22] Vakil, R.: The rising sea: foundations of algebraic geometry. https://math.stanford.edu/~vakil/216blog/FOAGnov1817public.pdf (2017)

[23] Van den Dries, L. P. D.: Model theory of fields: Decidability, and bounds for polynomial ideals. Ph.D. thesis, Rijksuniversiteit Utrecht, Utrecht (1978)

[24] Van den Dries, L., Schmidt, K.: Bounds in the theory of polynomial rings over fields. A non-standard approach. Invent. Math. **76**, 77–91 (1984) Zbl 0539.13011 MR 739626

[25] Varnavides, P.: On certain sets of positive density. J. Lond. Math. Soc. **34**, 358–360 (1959) Zbl 0088.25702 MR 106865