Title
Cutting lemma and Zarankiewicz's problem in distal structures

Authors
Chernikov, Artem
Galvin, David
Starchenko, Sergei

Peer reviewed

CUTTING LEMMA AND ZARANKIEWICZ'S PROBLEM IN DISTAL STRUCTURES

ARTEM CHERNIKOV, DAVID GALVIN, AND SERGEI STARCHENKO

ABSTRACT. We establish a cutting lemma for definable families of sets in distal structures, as well as the optimality of the distal cell decomposition for definable families of sets on the plane in o-minimal expansions of fields. Using it, we generalize the results in [10] on the semialgebraic planar Zarankiewicz problem to arbitrary o-minimal structures, in particular obtaining an o-minimal generalization of the Szemerédi-Trotter theorem.

1. INTRODUCTION

The so called *cutting lemma* is a very useful combinatorial partition tool with numerous applications in computational and incidence geometry and related areas (see e.g. [11, Sections 4.5, 6.5] or [5] for a survey). In it's simplest form it can be stated as follows (see e.g. [11, Lemma 4.5.3]).

Fact 1.1. *For every set L of n lines in the real plane and every $1 < r < n$ there exists a $\frac{1}{r}$-cutting for L of size $O(r^2)$. That is, there is a subdivision of the plane into generalized triangles (i.e. intersections of three half-planes) $\Delta_1, \ldots, \Delta_t$ so that the interior of each Δ_i is intersected by at most $\frac{n}{r}$ lines in L, and we have $t \leq Cr^2$ for a certain constant C independent of n and r.*

This result provides a method to analyze intersection patterns in families of lines, and it has many generalizations to higher dimensional sets and/or to families of sets of more complicated shape than lines, for example for families of algebraic or semialgebraic curves of bounded complexity [6]. The proofs of these generalizations typically combine some kind of geometric "cell decomposition" result with the so-called random sampling technique of Clarkson and Shor [9].

The aim of this article is to establish a rather general version of the cutting lemma for definable (in the sense of first-order logic) families of sets in a certain model-theoretically tame class of structures (namely, for distal structures — see Section 2 for the definition), as well as to apply it to generalize some of the results in the area from the semialgebraic context to arbitrary o-minimal structures. This work can be viewed as a continuation and refinement of the work started in [8], where the connection of model-theoretic distality with a weak form of the cutting lemma was discovered (we don't assume familiarity with that paper, but recommend its introduction for an expanded discussion of the model theoretic preliminaries). We believe that distal structures provide the most general natural setting for investigating questions in "generalized incidence combinatorics".

Let us describe the main results of the paper. Our first theorem establishes a cutting lemma for a definable family of sets in a distal structure, with the bound corresponding to the bound on the size of its distal cell decomposition. This is a generalized form of Matoušek's axiomatic treatment of Clarkson's random sampling method discussed in [11, Section 6.5]. The proof relies in particular on Lemma 3.5 on correlations in set-systems to deal with the lack of the corresponding notion of "being in a general position".

Theorem. *(Theorem 3.1, Distal cutting lemma) Let $\varphi(x; y)$ be a formula admitting a distal cell decomposition $\mathcal{T}$ (given by a finite set of formulas $\Psi(x; \bar{y})$ — see Definition 2.6) with $|\mathcal{T}(S)| = O(|S|^d)$ (i.e. for some constant $C \in \mathbb{R}$, for any **non-empty** finite $S \subseteq M^{|y|}$ we have $|\mathcal{T}(S)| \leq C|S|^d$).*

Then for any finite $H \subseteq M^{|y|}$ of size n and any real r satisfying $1 < r < n$, there are subsets $X_1, \ldots, X_t$ of $M^{|x|}$ covering $M^{|x|}$ with

$$t \leq Cr^d$$

for some constant $C = C(\varphi)$ (and independent of H, r and n), and with each X_i crossed by at most n/r of the formulas $\{\varphi(x; a) : a \in H\}$.

Moreover, each X_i is the intersection of at most two sets Ψ-definable over H (see Definition 2.2).

While every formula in a distal structure admits a distal cell decomposition (see Fact 2.8), establishing optimal bounds in dimension higher than 1 is non-trivial. In our second theorem, we demonstrate that formulas in o-minimal structures admit distal cell decompositions of optimal size "on the plane".

Theorem. *(Theorem 4.1) Let $\mathcal{M}$ be an o-minimal expansion of a real closed field. For any formula $\varphi(x; y)$ with $|x| = 2$ there is a distal cell decomposition $\mathcal{T}$ with $|\mathcal{T}(S)| = O(|S|^2)$.*

In our proof, we show that a version of the vertical cell decomposition can be generalized to arbitrary o-minimal theories. This gives an optimal bound for subsets of M^2, but determining the exact bounds for distal cell decompositions in higher dimensions remains open, even in the semialgebraic case.

Finally, in Section 5 we apply these two theorems to generalize the results in [10] on the semialgebraic Zarankiewicz problem to arbitrary o-minimal structures, in the planar case (our result is more general and applies to arbitrary definable families admitting a quadratic distal cell decomposition, see Section 5 for the precise statements).

Theorem. *(Theorem 5.13) Let $\mathcal{M}$ be an o-minimal expansion of a real closed field and let $E(x, y) \subseteq M^2 \times M^d$ be a definable relation.*
(1) For every $k \in \mathbb{N}$ there is a constant $c = c(E, k)$ such that for any finite $P \subseteq M^2, Q \subseteq M^d$, $|P| = m, |Q| = n$, if $E \cap (P \times Q)$ does not contain a copy of $K_{k,k}$ (the complete bipartite graph with two parts of size k), then we have

$$|E(P, Q)| \leq c \left(m^{\frac{d}{2d-1}} n^{\frac{2d-2}{2d-1}} + m + n \right).$$

(2) There is some $k' \in \mathbb{N}$ and formulas $\varphi(x, v), \psi(y, w)$ depending only on E such that if E contains a copy of $K_{k', k'}$, then there are some parameters $b \in M^v, c \in M^w$ such that both $\varphi(M, b)$ and $\psi(M, c)$ are infinite and $\varphi(M, b) \times \psi(M, c) \subseteq E$.

Combining the two parts, it follows that either E contains a product of two infinite definable sets, or the upper bound on the number of edges in part (1) holds for all finite sets P, Q with some fixed constant $c = c(E)$.

The special case with $d = 2$ can be naturally viewed as a generalization of the classical Szemerédi-Trotter theorem for o-minimal structures.

Corollary 1.2. *Let $\mathcal{M}$ be an o-minimal expansion of a real closed field. Then for every definable relation $E(x, y) \subseteq M^2 \times M^2$ there is a constant c and some formulas $\varphi(x, v), \psi(y, w)$ depending only on E such that exactly one of the following occurs:*
(1) For any finite $P \subseteq M^2, Q \subseteq M^2$, $|P| = m, |Q| = n$ we have

$$|E(P, Q)| \leq c \left(m^{\frac{2}{3}} n^{\frac{2}{3}} + m + n \right),$$

(2) there are some parameters $b \in M^v, c \in M^w$ such that both $\varphi(M, b)$ and $\psi(M, c)$ are infinite and $\varphi(M, b) \times \psi(M, c) \subseteq E$.

Remark 1.3. While this paper was in preparation, we have learned that Basu and Raz [3] have obtained a special case of Corollary 1.2 using different methods.

2. Preliminaries and the distal cell decomposition

Let $\mathcal{M}$ be an arbitrary first-order structure in a language $\mathcal{L}$. At this point we don't make any additional assumptions on $\mathcal{M}$, e.g. we may work in "set theory", i.e. in a structure where every subset is definable. We introduce some basic notation and terminology. Given a tuple of variables x, we let $|x|$ denote its length. For each $n \in \mathbb{N}$, M^n denotes the corresponding cartesian power of M, the underlying set of $\mathcal{M}$. For a fixed formula $\varphi(x; y) \in \mathcal{L}$ with two groups of variables x and y, given $b \in M^{|y|}$ we write $\varphi(M; b)$ to denote the set $\{a \in M^{|x|} : \mathcal{M} \models \varphi(a; b)\}$. Hence the formula $\varphi(x; y)$ can be naturally associated with the definable family of sets $\{\varphi(M; b) : b \in M^{|y|}\}$. E.g., if $\mathcal{M}$ is the field of reals, all sets in such a family for a fixed $\varphi(x; y)$ are semialgebraic of description complexity bounded by some $d = d(\varphi)$ and conversely, the family of all semialgebraic sets of description complexity bounded by some fixed d can be obtained in this way for an appropriate choice of the formula $\varphi(x; y)$. We refer to [8] for a more detailed introduction and examples of the relevant model-theoretic terminology.

Definition 2.1. For sets $A, X \subseteq M^d$ we say that A *crosses* X if both $X \cap A$ and $X \cap \neg A$ are nonempty.

We extend the above definition to a set of formulas.

Definition 2.2. Let $\Phi(x; y)$ be a set of $\mathcal{L}$-formulas of the form $\varphi(x; y) \in \mathcal{L}$ and $S \subseteq M^{|y|}$.
(1) We say that a subset $A \subseteq M^{|x|}$ is $\Phi(x; S)$-*definable* if $A = \varphi(M; s)$ for some $\varphi(x; y) \in \Phi$ and $s \in S$.
(2) For a set $X \subseteq M^{|x|}$ we say that $\Phi(x; S)$ *crosses* X if some $\Phi(x; S)$-definable set crosses X. In other words $\Phi(x; S)$ does not cross X if for any $\varphi(x; y) \in \Phi(x; y)$ and $s \in S$ the formula $\varphi(x; s)$ has a constant truth value on X.

We define a very general combinatorial notion of an abstract cell decomposition for formulas (equivalently, for definable families of sets).

Definition 2.3. Let $\Phi(x; y)$ be a finite set of formulas.

(1) Given a finite set $S \subseteq M^{|y|}$, a family $\mathcal{F}$ of subsets of $M^{|x|}$ is called *an abstract cell decomposition for* $\Phi(x; S)$ if $M^{|x|} = \cup \mathcal{F}$ and every $\Delta \in \mathcal{F}$ is not crossed by $\Phi(x; S)$.
(2) *An abstract cell decomposition for* $\Phi(x; y)$ is an assignment $\mathcal{T}$ that to each finite set $S \subseteq M^{|y|}$ assigns an abstract cell decomposition $\mathcal{T}(S)$ for $\Phi(x; S)$.

Remark 2.4. In the above definition, the term "cell decomposition" is understood in a very weak sense. Firstly, the "cells" in $\mathcal{T}(S)$ are not required to have any "geometric" properties, and secondly, we don't require the family $\mathcal{T}(S)$ to partition $M^{|x|}$, but only ask for it to be a covering.

Every $\Phi(x; y)$ admits an obvious abstract cell decomposition, with $\mathcal{T}(S)$ consisting of the atoms in the Boolean algebra generated by the $\Phi(x; S)$-definable sets. In general, defining these cells would require longer and longer formulas when S grows, and the aim of the following definitions is to avoid this possibility.

Definition 2.5. Let $\Phi(x; y)$ be a finite set of formulas and $\mathcal{T}$ an abstract cell decomposition for $\Phi(x; y)$. We say that $\mathcal{T}$ is *weakly definable* if there is a finite set of formulas $\Psi(x; \bar{y}) = \Psi(x; y_1, \ldots, y_k)$ with $|y_1| = \cdots = |y_k| = |y|$ such that for any finite $S \subseteq M^{|y|}$, every $\Delta \in \mathcal{T}(S)$ is $\Psi(x; \bar{y})$-definable over S (i.e., $\Delta = \psi(M; s_1, \ldots, s_k)$ for some $s_1, \ldots, s_k \in S$ and $\psi \in \Psi$).

Notice that if an abstract cell decomposition $\mathcal{T}$ for $\Phi(x; y)$ is weakly defined by $\Psi(x; \bar{y})$ then $\Psi(x; \bar{y})$ does not determine $\mathcal{T}$ uniquely; however there is a maximal abstract cell decomposition $\mathcal{T}^{\max}$ weakly defined by $\Psi(x; \bar{y})$, where $\mathcal{T}^{\max}(S)$ consists of *all* $\Psi(x; \bar{y})$-definable over S sets Δ such that $\Phi(x; S)$ does not cross Δ.

For combinatorial applications discussed in this paper, we would like to have a cell decomposition with as few sets as possible, and we want to have control over the sets appearing in $\mathcal{T}(S)$ in a definable way.

Definition 2.6. Let $\Phi(x; y)$ be a finite set of formulas. An abstract cell decomposition $\mathcal{T}$ for Φ is *definable* if for every finite $S \subseteq M^{|y|}$ there is a family $\Psi(S)$ of subsets of $M^{|x|}$, uniformly definable in S, and for each $\Delta \in \Psi(S)$ a subset $\mathcal{I}(\Delta) \subseteq M^{|y|}$, uniformly definable in Δ, such that

$$(2.1) \qquad \mathcal{T}(S) = \{\Delta \in \Psi(S) \colon \mathcal{I}(\Delta) \cap S = \emptyset\}.$$

By the uniform definability of $\Psi(S)$ we mean the existence of a finite set of formulas $\Psi(x; \bar{y})$ as above, so that for any finite $S \subseteq M^{|y|}$ the set

$\Psi(S)$ consists of all $\Psi(x; \bar{y})$-definable over S sets; and uniform definability of $\mathcal{I}(\Delta)$ means that for every $\psi(x; \bar{y}) \in \Psi(x; \bar{y})$ there is a formula $\theta_\psi(y; \bar{y})$ such that for any $s_1, \ldots, s_k \in M^{|y|}$ if $\Delta = \psi(M; s_1, \ldots, s_k)$ then $\mathcal{I}(\Delta) = \theta_\psi(M; s_1, \ldots, s_k)$.

For example, $\mathcal{T}^{\max}$ defined above is definable with $\mathcal{I}(\Delta) = \{s \in M^{|y|} \colon \Phi(x; s) \text{ crosses } \Delta\}$.

Remark 2.7. It follows from Definition 2.6 that for every $\Psi(x; M)$-definable set $\Delta \subseteq M^{|x|}$, the set of all $s \in M^{|y|}$ such that $\Phi(x; s)$ crosses Δ is contained in $\mathcal{I}(\Delta)$ (strict containment is possible, however).

Indeed, assume that $s \in M^{|y|}$ and $\varphi(x; y) \in \Phi$ are such that $\varphi(x; s)$ crosses Δ. By Definition 2.3(1), necessarily $\Delta \notin \mathcal{T}(\{s\})$. But then $\mathcal{I}(\Delta) \cap \{s\} \neq \emptyset$ by (2.1), hence $s \in \mathcal{I}(\Delta)$.

As was realized in [8], such combinatorial definable cell decompositions have a close connection to the model-theoretic notion of distality. Distal structures were introduced in [13] for purely model theoretic purposes (we don't give the original definition here). The following was pointed out in [8] (and can be used as the definition of a distal structure in this paper).

Fact 2.8. *The following are equivalent for a first-order structure $\mathcal{M}$.*
(1) $\mathcal{M}$ is distal,
(2) for every formula $\varphi(x; y)$ there is a weakly definable cell decomposition for $\{\varphi(x; y)\}$,
(3) for every formula $\varphi(x; y)$ there is a definable cell decomposition for $\{\varphi(x; y)\}$.

Equivalence of the original definition of distality and existence of weakly definable cell decompositions is given by [7, Theorem 21]; and if $\mathcal{T}$ is a weakly definable cell decomposition for $\varphi(x; y)$, then $\mathcal{T}^{\max}$ as defined above is definable.

Examples of distal structures include (again, we refer to the introduction of [8] for a more detailed discussion):
(1) o-minimal structures,
(2) Presburger arithmetic $(\mathbb{Z}, +, 0, <)$,
(3) the field of p-adics $\mathbb{Q}_p$.

There are several contexts in model theory relevant for the topics of this paper where certain notions of cell decomposition play a prominent role (e.g. o-minimal cell decomposition, p-adic cell decomposition, etc.). These cell decompositions tend to carry more geometric information, while the one distinguished here captures combinatorial complexity. To distinguish from those cases, and in view of Fact 2.8, we

will from now on refer to a definable cell decomposition $\mathcal{T}$ for a finite set of formulas $\Phi(x; y)$ as in Definition 2.6 as a *distal cell decomposition* for $\Phi(x; y)$. Hence, **a structure $\mathcal{M}$ is distal if and only if every formula admits a distal cell decomposition**.

Distality of the examples listed above had been established by different (sometimes infinitary) methods and the question of obtaining the exact bounds on the size of the corresponding distal cell decompositions hasn't been considered. While it is easy to verify in the examples listed above that all formulas $\varphi(x, y)$ with $|x| = 1$ admit a cell decomposition $\mathcal{T}$ with the best possible bound $|\mathcal{T}(S)| = O(|S|)$, already the case of formulas with $|x| = 2$ becomes more challenging (and grows in complexity with $|x|$). In Section 4 we establish that in an o-minimal expansion of a field, all formulas with $|x| = 2$ admit a distal cell decomposition $\mathcal{T}$ with the optimal bound $|\mathcal{T}(S)| = O(|S|^2)$ (the case $|x| \geq 3$ remains open, even in the semialgebraic case).

3. Distal cutting lemma

In this section we show how a bound on the size of a distal cell decomposition for a given definable family can be used to obtain a definable cutting lemma with the corresponding bound on its size.

Theorem 3.1. *(Distal cutting lemma) Let $\varphi(x; y)$ be a formula admitting a distal cell decomposition $\mathcal{T}$ (given by a finite set of formulas $\Psi(x; y_1, \ldots, y_s)$ — see Definition 2.6) with $|\mathcal{T}(S)| = O(|S|^d)$.*

Then for any finite $H \subseteq M^{|y|}$ of size n and any real r satisfying $1 < r < n$, there are subsets $X_1, \ldots, X_t$ of $M^{|x|}$ covering $M^{|x|}$ with

$$t \leq Cr^d$$

for some constant $C = C(\varphi)$ (and independent of H, r and n), and with each X_i crossed by at most n/r of the formulas $\{\varphi(x; a) : a \in H\}$.

Moreover, each of the X_i's is an intersection of at most two $\Psi(x; H)$-definable sets.

Our proof generalizes (and closely follows) the axiomatic treatment of the Clarkson-Shor random sampling technique in [11, Section 6.5].

Let $\mathcal{T}, \Psi$ and H as in the assumption be fixed. Then (recalling Definition 2.6) for each finite $S \subseteq M^{|y|}$, we have a finite collection $\mathcal{T}(S)$ of subsets of $M^{|x|}$ that covers $M^{|x|}$ and satisfies the following conditions.

(C1) Let
$$\text{Reg} := \{\Delta : \Delta \in \mathcal{T}(S) \text{ for some } S \subseteq H\}.$$
Then every set in Reg is definable by an instance of a formula from Ψ with parameters in H.

(C2) For every $S \subseteq H$ we have
$$|\mathcal{T}(S)| \leq C' \left(|S|^d + 1\right)$$
for some constant C' depending only on φ. (The hypothesis of the theorem ensures that for *non*-empty S we have $|\mathcal{T}(S)| \leq C|S|^d$ for some constant $C = C(\varphi)$. We add "+1" here to take into account the case $S = \emptyset$.)

(C3) To each $\Delta \in \mathrm{Reg}$ we associate a collection $\mathcal{D}(\Delta)$ of subsets of H, called the *defining sets* of Δ, via
$$\mathcal{D}(\Delta) := \{S \subseteq H : |S| \leq s, \ \Delta \in \mathcal{T}(S)\}.$$
(Here s is a fixed constant corresponding to the number of parameters in $\Psi(x; y_1, \ldots, y_s)$ given by the distal cell decomposition and depending only on φ). Given $\mathcal{I}$ as in Definition 2.6, we define $\mathcal{I}_H(\Delta) := \mathcal{I}(\Delta) \cap H$. Notice that $\mathcal{I}_H(\Delta)$ contains all of the $a \in H$ such that $\varphi(x; a)$ crosses Δ (by Remark 2.7). We have:
$$\Delta \in \mathcal{T}(S) \iff \mathcal{I}_H(\Delta) \cap S = \emptyset \text{ and there is } S_0 \in \mathcal{D}(\Delta) \text{ with } S_0 \subseteq S.$$

Remark 3.2. It follows from the proof that the distal cutting lemma (Theorem 3.1) holds for any abstract cell decomposition satisfying the conditions (C1)–(C3) with an appropriately chosen relation $\mathcal{I}(\Delta)$.

Before proceeding to the proof of the distal cutting lemma (Theorem 3.1) we isolate two key tools. The first is a tail bound on the probability that a cell $\Delta \in \mathcal{T}(S)$ is crossed by many formulas, where S is a randomly chosen subset of H.

For $S \subseteq H$, let $\mathcal{T}(S)_{\geq t}$ denote the set of $\Delta \in \mathcal{T}(S)$ with $|\mathcal{I}_H(\Delta)| \geq tn/r$. Recall that for $0 \leq p \leq 1$ we say that $S \subseteq H$ is selected by *independent Bernoulli trials with success probability* p if S is selected according to the distribution μ (supported on the power set of H) given by
$$\mu(S') = p^{|S'|}(1-p)^{|H|-|S'|}$$
for each $S' \subseteq H$; observe that this is essentially the process of flipping a biased coin (biased to show heads with probability p) $|H|$ times independently, and for $1 \leq i \leq |H|$ putting the ith element of H in S if and only if the ith flip comes up heads.

Lemma 3.3. *(Tail Bound Lemma) Let $\varphi(x; y)$ be a formula as in Theorem 3.1. Let $H \subseteq M^{|y|}$ be a finite set of size n. Fix $\varepsilon > 0$ and let r be a parameter satisfying $1 \leq r \leq (1 - \varepsilon)n$. Let $S \subseteq H$ be selected by independent Bernoulli trials with success probability r/n, and let $t \geq 0$ be given. Then there is a constant $C = C(\varepsilon)$ such that*
$$\mathbf{E}_\mu \left(|\mathcal{T}(S)_{\geq t}|\right) \leq C2^{-t}r^d.$$

We use this to derive the second main tool, a cutting lemma that is weaker than Theorem 3.1.

Lemma 3.4. *(Suboptimal Cutting Lemma) Let $\varphi(x; y)$ be a formula as in Theorem 3.1. Let $H \subseteq M^{|y|}$ be a finite set of size n. Let r be a parameter satisfying $1 < r < n$. There is $S \subseteq H$ with*

$$|\mathcal{T}(S)| \leq Kr^d \log^d(r + 1)$$

for some constant K independent of H, r and n and with each $X \in \mathcal{T}(S)$ crossed by at most n/r of the formulas $\{\varphi(x; a) : a \in H\}$.

Proof. Let A be such that $3 \times 2^{2d} C A^d = 2^A$, where C is the constant appearing in Lemma 3.3. We treat separately the cases $2Ar \log(r+1) \leq n$ and $2Ar \log(r+1) \geq n$. If $2Ar \log(r+1) \geq n$ then we may take $\mathcal{T}(H)$ as our r-cutting, since it has size $C'(n^d + 1) \leq C'((2A)^d r^d \log^d(r+1) + 1) \leq Kr^d \log^d(r+1)$ for suitably large K (note that by **(C3)** no instance of $\varphi(x; y)$ over H can cross any of the sets in $\mathcal{T}(H)$).

Suppose now that $2Ar \log(r + 1) \leq n$. Set $r' = Ar \log(r + 1)$. Applying Lemma 3.3 with r' taking the role of r (valid since $r' < n/2$) and with $t = 0$ we obtain that if $S \subseteq H$ is selected by independent Bernoulli trials with success probability r'/n (with associated distribution μ') then

$$\mathbf{E}_{\mu'}(|\mathcal{T}(S)|) \leq CA^d r^d \log^d(r + 1).$$

Applying Lemma 3.3 again with $t = A \log(r + 1)$ we get

$$\mathbf{E}_{\mu'}\left(\left|\mathcal{T}(S)_{\geq A\log(r+1)}\right|\right) \leq \frac{CA^d r^d \log^d(r + 1)}{(r + 1)^A} \leq \frac{CA^d}{(r + 1)^{A-2d}} \leq 1/3,$$

the second inequality using $r \log(r + 1) \leq (r + 1)^2$ and the third using our choice of A and the fact that $r \geq 1$. By linearity of expectation

$$\mathbf{E}_{\mu'}\left(\frac{|\mathcal{T}(S)|}{3CA^d r^d \log^d(r + 1)} + \left|\mathcal{T}(S)_{\geq A\log(r+1)}\right|\right) \leq 2/3,$$

so there exists an $S \subseteq H$ such that

$$|\mathcal{T}(S)| \leq 3CA^d r^d \log^d(r + 1)$$

and $\mathcal{T}(S)_{\geq A\log(r+1)} = \emptyset$. This last condition implies that each $\Delta \in \mathcal{T}(S)$ is crossed by at most $(A\log(r + 1)n)/r' = n/r$ formulas. $\square$

We use Lemmas 3.4 and 3.3 to derive Theorem 3.1, before turning to the proof of Lemma 3.3.

Proof of Theorem 3.1. Just as in the proof of Lemma 3.4 we begin by observing that $\mathcal{T}(H)$ furnishes an r-cutting for all r, with size at most $C'(n^d + 1)$. This allows us to assume, say, $r \leq n/2$, which allows us to use Lemma 3.3.

Let $S \subseteq H$ be selected by independent Bernoulli trials with success probability r/n, and let $\mathcal{T}(S)$ be as in the assumption.

For $\Delta \in \mathcal{T}(S)$ define t_Δ by $|\mathcal{I}_H(\Delta)| = t_\Delta n/r$. Note that if $t_\Delta \leq 1$ then the number of a in H such that $\varphi(x, a)$ crosses Δ is no more than n/r.

For $\Delta \in \mathcal{T}(S)$ with $t_\Delta > 1$, consider the set $\mathcal{I}_H(\Delta)$, it contains all of $a \in H$ for which $\varphi(x, a)$ crosses Δ. By Lemma 3.4 there is $S' \subseteq \mathcal{I}_H(\Delta)$ with $\mathcal{T}(S')$ having size at most $O(t_\Delta^d \log(t_\Delta^d + 1))$ with the property that for every $\Delta' \in \mathcal{T}(S')$, the number of $a \in \mathcal{I}_H(\Delta)$ such that $\varphi(x, a)$ crosses Δ' is at most

$$\frac{|\mathcal{I}(\Delta)|}{t_\Delta} = \frac{n}{r}.$$

In particular that means that for every $\Delta' \in \mathcal{T}(S')$ the number of $a \in H$ such that $\varphi(x, a)$ crosses $\Delta' \cap \Delta$ is at most n/r.

It follows that the family of subsets of $M^{|x|}$ consisting of those $\Delta \in \mathcal{T}(S)$ for which $t_\Delta \leq 1$, together with all sets of the form $\Delta' \cap \Delta$ where $\Delta \in \mathcal{T}(S)$ has $t_\Delta > 1$ and $\Delta' \in \mathcal{T}(S')$ (with S' constructed from S via Lemma 3.4, as described above), forms a cover of $M^{|x|}$ with size at most

$$(3.1) \qquad \sum_{\Delta \in \mathcal{T}(S)} \left(\mathbf{1}_{\{t_\Delta \leq 1\}} + C t_\Delta^d \log^d(t_\Delta + 1) \mathbf{1}_{\{t_\Delta > 1\}} \right).$$

We now upper bound the expectation (with respect to μ) of this quantity. By linearity the expectation is at most

$$(3.2) \qquad \mathbf{E}_\mu \left(|\mathcal{T}(S)| \right) + C \sum_{i \geq 0} \mathbf{E}_\mu \left(\sum_{\Delta \in \mathcal{T}(S):\, 2^i \leq t_\Delta < 2^{i+1}} t_\Delta^{2d} \right)$$

(using $\log(t_\Delta + 1) \leq t_\Delta$ for $t_\Delta \geq 1$).

We bound the first term in (3.2) by an application of Lemma 3.3 with $t = 0$. This gives

$$\mathbf{E}_\mu \left(|\mathcal{T}(S)| \right) \leq O(r^d).$$

For the second term in (3.2) we have

$$\sum_{i\geq 0}\mathbf{E}_\mu\left(\sum_{\Delta\in\mathcal{T}(S)\colon 2^i\leq t_\Delta<2^{i+1}}t_\Delta^{2d}\right)\leq\sum_{i\geq 0}2^{2d(i+1)}\mathbf{E}_\mu\left(|\mathcal{T}(S)_{\geq 2^i}|\right)$$

$$\leq C\sum_{i\geq 0}2^{2d(i+1)}2^{-2^i}r^d=O(r^d),$$

with the last inequality being an application of Lemma 3.3.

We conclude that the expectation of the quantity in (3.1) is $O(r^d)$, so there is at least one choice of $S\subseteq H$ for which (3.1) is at most $O(r^d)$, proving Theorem 3.1 (the definability clause follows by **(C1)** as every set in the constructed covering is an intersection of at most two sets from Reg). $\qquad\square$

Before proving Lemma 3.3 we isolate a useful set-systems lemma.

Lemma 3.5. *Let Ω be a set of size m, and let $\{D_1,\ldots,D_q\}$ be a collection of subsets of Ω with $|D_i|\leq u$ for all i, $1\leq i\leq q$, for some u. Let*

$$\mathcal{F}=\{X\subseteq\Omega:D_i\subseteq X\text{ for some }i,\ 1\leq i\leq q\}$$

be the up-set (or filter) generated by the D_i's. Let $\tilde{p}$ and p satisfy $0<\tilde{p}\leq p\leq 1$. We have

$$(3.3)\qquad\frac{\sum_{X\in\mathcal{F}}\tilde{p}^{|X|}(1-\tilde{p})^{m-|X|}}{\sum_{X\in\mathcal{F}}p^{|X|}(1-p)^{m-|X|}}\geq\left(\frac{\tilde{p}}{p}\right)^u.$$

Proof. With each $X\in\mathcal{F}$ associate (arbitrarily) a set D_X satisfying $D_X\subseteq X$ and $D_X\in\{D_1,\ldots,D_q\}$ (such a set exists by the definition of $\mathcal{F}$).

Let $A\subseteq\Omega$ be selected by independent Bernoulli trials with success probability p, and, independently, let $B\subseteq\Omega$ be selected by independent Bernoulli trials with success probability $\tilde{p}/p$. Observe that

$$(3.4)\qquad\Pr(A\in\mathcal{F})=\sum_{X\in\mathcal{F}}p^{|X|}(1-p)^{m-|X|}$$

and

$$(3.5)\qquad\Pr(A\cap B\in\mathcal{F})=\sum_{X\in\mathcal{F}}\tilde{p}^{|X|}(1-\tilde{p})^{m-|X|},$$

with (3.5) holding by independence and because for each $\omega\in\Omega$, $\Pr(\omega\in A\cap B)=\Pr(\omega\in A)\Pr(\omega\in B)$.

Now consider the two events

$$E_1=\{A\in\mathcal{F}\text{ and }D_A\subseteq B\}$$

and
$$E_2 = \{A \cap B \in \mathcal{F}\}.$$
If $A \in \mathcal{F}$ and $D_A \subseteq B$ then $D_A \subseteq A \cap B$, so that $A \cap B \in \mathcal{F}$. It follows that $E_1 \subseteq E_2$ and

(3.6) $$\Pr(E_1) \leq \Pr(E_2).$$

Using independence we have

$$\Pr(E_1) = \sum_{X \in \mathcal{F}} \Pr(A = X \text{ and } D_X \subseteq B) = \sum_{X \in \mathcal{F}} \Pr(A = X) \left(\frac{\tilde{p}}{p}\right)^{|D(X)|}$$
$$\geq \Pr(A \in \mathcal{F}) \left(\frac{\tilde{p}}{p}\right)^u.$$

Combining with (3.4), (3.5) and (3.6) we get (3.3). $\square$

We are now ready to prove Lemma 3.3. We follow Matoušek's approach in [11, Section 6.5], but add an additional argument.

Proof of Lemma 3.3. We start by establishing

(3.7) $$\mathbf{E}_\mu (|\mathcal{T}(S)|) = O(r^d),$$

which gives Lemma 3.3 for $t \leq 1$. To see (3.7) note that **(C2)** yields $\mathbf{E}_\mu (|\mathcal{T}(S)|) \leq C\mathbf{E}_\mu (|S|^d) + 1$. Now $|S| = X_1 + \ldots + X_n$ where the $X_i's$ are independent Bernoulli random variables each with parameter $p = r/n$. We claim that for all $d \geq 1$ we have

(3.8) $$\mathbf{E}_\mu(|S|^d) \leq (r + d)^d.$$

(from which (3.7) immediately follows; note that we can drop the $+1$ since $r \geq 1$).

To see (3.8), note first that by linearity we have

$$\mathbf{E}_\mu(|S|^d) = \sum_{(i_1, i_2, \ldots, i_d) \in \{1, \ldots, n\}^d} \mathbf{E}(X_{i_1} X_{i_2} \cdots X_{i_d}).$$

Let a_k be the number of tuples $(i_1, i_2, \ldots, i_d) \in \{1, \ldots, n\}^d$ such that $|\{i_1, i_2, \ldots, i_d\}| = d - k$. By independence of the X_i, and the fact that X_i^ℓ has the same distribution as X_i for any integer $\ell \geq 1$ we have

(3.9) $$\mathbf{E}_\mu(|S|^d) = \sum_{k=0}^{d} a_k p^{d-k}.$$

We claim that

(3.10) $$a_k \leq \binom{d}{k} d^k n^{d-k}.$$

Inserting into (3.9) and using the binomial theorem together with $np = r$, this gives (3.8).

To see (3.10) note that we overcount a_k by first specifying $d - k$ indices from $\{1, \ldots, d\}$ on which the i_j's are all different from each other $(\binom{d}{d-k} = \binom{d}{k}$ choices), then choosing values for these i_j's $(n(n - 1) \cdots (n - (d - k) + 1) \leq n^{d-k}$ choices), and finally choosing values for the remaining indices $((d - k)^k \leq d^k$ choices, since these indices are all constrained to lie among the $d - k$ distinct indices chosen initially). It follows that $a_k \leq \binom{d}{k} n^{d-k} d^k$, as claimed.

(We note that in the case $d = 2$ things are considerably easier: we have

$$|S|^2 = \sum_{i=1}^{n} X_i^2 + 2 \sum_{1 \leq i < j \leq n} X_i X_j$$

so

$$\mathbf{E}_\mu(|S|^2) = \sum_{i=1}^{n} \mathbf{E}(X_i^2) + 2 \sum_{1 \leq i < j \leq n} \mathbf{E}(X_i X_j)$$
$$= np + n(n - 1)p^2$$
$$\leq np + n^2 p^2 = r^2 + r.)$$

We assume from now on that $t \geq 1$. For $\Delta \in \mathrm{Reg}$ denote by $p(\Delta)$ the probability that Δ appears in $\mathcal{T}(S)$, i.e.

$$p(\Delta) = \mu(\{S \subseteq H \colon \Delta \in \mathcal{T}(S)\}) = \sum_{\Delta \in \mathcal{T}(S)} \mu(S).$$

Let $\mathrm{Reg}_{\geq t} = \{\Delta \in \mathrm{Reg} \colon |\mathcal{I}_H(\Delta)| \geq tn/r\}$. By linearity of expectation we have

(3.11) $$\mathbf{E}_\mu\left(|\mathcal{T}(S)_{\geq t}|\right) = \sum_{\Delta \in \mathrm{Reg}_{\geq t}} p(\Delta).$$

Now set $\tilde{p} = p/t$ and let $\tilde{\mu}$ be the distribution associated with selection from H by independent Bernoulli trials with success probability $\tilde{p}$. By (3.7) we have

(3.12) $$\mathbf{E}_{\tilde{\mu}}\left(|\mathcal{T}(S)|\right) = O(r^d/t^d).$$

Also, as in (3.11) we have

$$(3.13) \quad \mathbf{E}_{\tilde{\mu}} \left(|\mathcal{T}(S)| \right) = \sum_{\Delta \in \mathrm{Reg}} \tilde{p}(\Delta) \geq \sum_{\Delta \in \mathrm{Reg}_{\geq t}} \tilde{p}(\Delta)$$

$$= \sum_{\Delta \in \mathrm{Reg}_{\geq t}} p(\Delta) \frac{\tilde{p}(\Delta)}{p(\Delta)} \geq \min \left\{ \frac{\tilde{p}(\Delta)}{p(\Delta)} : \Delta \in \mathrm{Reg}_{\geq t} \right\} \sum_{\Delta \in \mathrm{Reg}_{\geq t}} p(\Delta)$$

$$= \min \left\{ \frac{\tilde{p}(\Delta)}{p(\Delta)} : \Delta \in \mathrm{Reg}_{\geq t} \right\} \mathbf{E}_{\mu} \left(|\mathcal{T}(S)_{\geq t}| \right).$$

We now estimate from below the quantity $\tilde{p}(\Delta)/p(\Delta)$ for $\Delta \in \mathrm{Reg}_{\geq t}$. Fix such a Δ and let $\mathcal{F}(\Delta)$ be the up-set on ground set $H \setminus \mathcal{I}_H(\Delta)$ generated by $\mathcal{D}(\Delta)$. Using **(C3)** we see that

$$p(\Delta) = (1-p)^{|\mathcal{I}_H(\Delta)|} \sum_{X \subseteq \mathcal{F}(\Delta)} p^{|X|}(1-p)^{|H \setminus \mathcal{I}_H(\Delta)| - |X|}$$

with an analogous expression for $\tilde{p}(\Delta)$. Recalling $\tilde{p}/p = 1/t$ and that defining sets have size at most s, an application of Lemma 3.5 immediately yields

$$\frac{\tilde{p}(\Delta)}{p(\Delta)} \geq \frac{(1-\tilde{p})^{|\mathcal{I}_H(\Delta)|}}{(1-p)^{|\mathcal{I}_H(\Delta)|}} \left(\frac{1}{t} \right)^s$$

$$\geq \left(\frac{1-\tilde{p}}{1-p} \right)^{tn/r} \left(\frac{1}{t} \right)^s$$

$$\geq \left(\frac{e^{-c\tilde{p}}}{e^{-p}} \right)^{tn/r} \left(\frac{1}{t} \right)^s$$

$$(3.14) \qquad\qquad = e^{t-c} t^{-s},$$

with the second inequality using $(1-\tilde{p})/(1-p) \geq 1$ and $|\mathcal{I}_H(\Delta)| \geq tn/r$, and the third inequality using the standard bound $1-p \leq e^{-p}$ (valid for all real p). In the third inequality we also use that for $0 \leq \tilde{p} \leq (1-\varepsilon)n$ (which certainly holds, since $\tilde{p} \leq p \leq (1-\varepsilon)n$) we have $1 - \tilde{p} \geq e^{-c\tilde{p}}$ for some sufficiently large $c = c(\varepsilon)$ ($c = \log(1/\varepsilon)/(1-\varepsilon)$ will do).

Inserting (3.14) into (3.13) and combining with (3.12) we finally get

$$\mathbf{E}_{\mu} \left(|\mathcal{T}(S)_{\geq t}| \right) \leq t^s e^{c-t} O(r^d/t^d) \leq C 2^{-t} r^d$$

for sufficiently large C. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

4. Optimal distal cell decomposition on the plane in o-minimal expansions of fields

Our goal in this section is to prove the following theorem.

Theorem 4.1. *Let $\mathcal{M}$ be an o-minimal expansion of a real closed field. Then any formula $\varphi(x; y)$ with $|x| = 2$ admits a distal cell decomposition $\mathcal{T}$ with $|\mathcal{T}(S)| = O(|S|^2)$.*

Towards this purpose, we fix a formula $\varphi(x; y)$ with $|x| = 2$ (and often we will write x as (x_1, x_2)).

We first construct a finite set of formulas $\Phi(x; y)$ such that for any $s \in M^{|y|}$ the set $\varphi(M; s)$ is a Boolean combination of $\Phi(x; s)$-definable sets, and formulas in $\Phi(x; y)$ have a very simple form, and then we construct a definable cell decomposition $\mathcal{T}$ for $\Phi(x; y)$ (hence also for φ) with $|\mathcal{T}(S)| = O(|S|^2)$.

Let $\varphi_1(x_1; x_2, y)$ be the formula $\varphi(x_1, x_2; y)$.

Using o-minimality and definable choice we can find definable functions $h_1, \ldots, h_k \colon M \times M^{|y|} \to M$ such that

$$h_1(a, s) \le h_2(a, s) \le \cdots \le h_k(a, s) \text{ for all } a \in M, s \in M^{|y|}$$

and for all $a \in M, s \in M^{|y|}$ and $i = 0, \ldots, k$ we have

$$h_i(a, s) < x_1, x_1' < h_{i+1}(a, s) \to [\varphi(x_1; a, s) \leftrightarrow \varphi(x_1'; a, s)],$$

where for convenience we let $h_0(a, s) = -\infty$ and $h_{k+1}(a, s) = +\infty$.

At this point we have that for a fixed $i = 0, \ldots, k$ for all $a \in M, s \in M^{|y|}$ the truth value of $\varphi(x_1, a; s)$ is constant on the interval $h_i(a, s) < x_1 < h_{i+1}(a, s)$ but may vary if we perturb a. We need to partition M into pieces where this truth value does not depend on a.

For $a, a' \in M$ and $s \in M^{|y|}$ we define the relation $a \sim_s a'$ as

$a \sim_s a'$ iff for all $i = 0, \ldots, k$
$\qquad$ and any $h_i(a, s) < x_1 < h_{i+1}(a, s), h_i(a', s) < x_1' < h_{i+1}(a', s)$
$\qquad\qquad$ we have $\varphi_1(x_1; a, s) \leftrightarrow \varphi(x_1'; a', s).$

Clearly $\sim_s$ is an equivalence relation on M with at most 2^{k+1}-classes uniformly definable in terms of s. Using o-minimality and definable choice, we can find definable functions $u_i \colon M^{|y|} \to M$, $i = 1, \ldots, l$ with $u_1(y) \le u_2(y) \le \cdots \le u_l(y)$ such that for all $s \in M^{|y|}$ and $i = 0, \ldots, l$ we have

$$u_i(s) < x_2, x_2' < u_{i+1}(s) \to x_2 \sim_s x_2',$$

where again for convenience we use $u_0(y) = -\infty$ and $u_{l+1}(y) = +\infty$.

We would prefer that for $s \in M^{|y|}$ each of the functions $x_2 \mapsto h_i(x_2, s)$ is continuous.

For $k \in \mathbb{N}$, we will write $[k]$ to denote the set $\{1, 2, \ldots, k\}$. Since every definable function is piecewise continuous, we can further partition M and in addition require that for any $i = 0, \ldots, l$, $j \in [k]$ and

every $s \in M^{|y|}$ the function $x_2 \mapsto h_j(x_2, s)$ is continuous on the interval $u_i(s) < x_2 < u_{i+1}(s)$.

We take $\Phi(x; y)$ to be the following set of formulas (recall that $x = (x_1, x_2)$):

$$\{x_2 = u_i(y) \colon i \in [l]\} \cup \{x_2 < u_i(y) \colon i \in [l]\}$$
$$\cup \{x_2 > u_i(y) \colon i \in [l]\} \cup \{x_1 = h_i(x_2, y) \colon i \in [k]\}$$
$$\cup \{x_1 < h_i(x_2, y) \colon i \in [k]\} \cup \{x_1 > h_i(x_2, y) \colon i \in [k]\}.$$

It is not hard to see that for any $s \in M^{|y|}$ the set $\varphi(M; s)$ is a Boolean combination of $\Phi(x; s)$-definable sets.

We now proceed with a construction of a definable cell decomposition for $\Phi(x; y)$.

Geometrically we view M^2 as (x_1, x_2)-plane, with x_1 being on the vertical axis and x_2 on horizontal. Then $\Phi(x; S)$-definable sets partition the plain by vertical lines $x_2 = u_i(s)$ and "horizontal" "curves" $x_1 = h_j(x_2, s)$.

Unfortunately we cannot use complete Φ-types over S as $\mathcal{T}(S)$. Since S is finite every complete Φ-type is equivalent to a formula however in general we cannot get uniform definability.

Consider a simple example of a partition of a plane by straight lines, i.e. the case when we don't have functions u_i and have only one $h(x_2, a, b)$ defining the straight lines $x_1 = ax_2 + b$. In the example below all points in the gray area have the same Φ-type, but we need at least 5 lines to describe the region; and in general, this number may be as big as one wants.
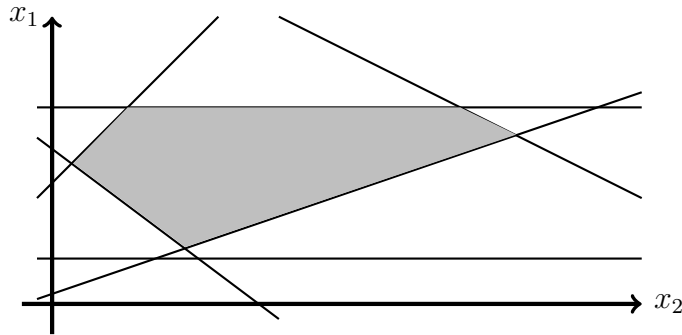
We could solve this problem by using also vertical lines through all points of intersections, as shown below, but then the size of the partition would be $O(|S|^3)$.
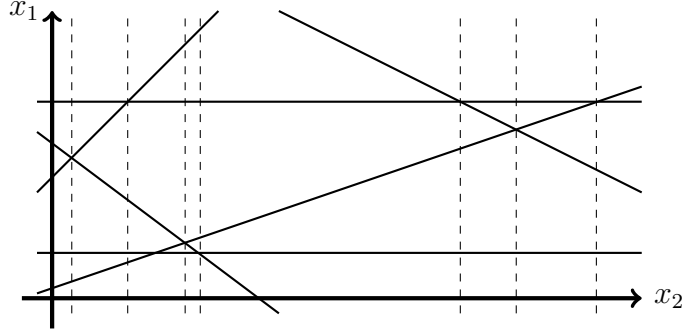
Using the idea of "vertical decomposition" from [9] we add only vertical line segments where they are needed, i.e. from an intersection point to the first line above (or plus infinity) and the first line below (or minus infinity), as in the following picture.
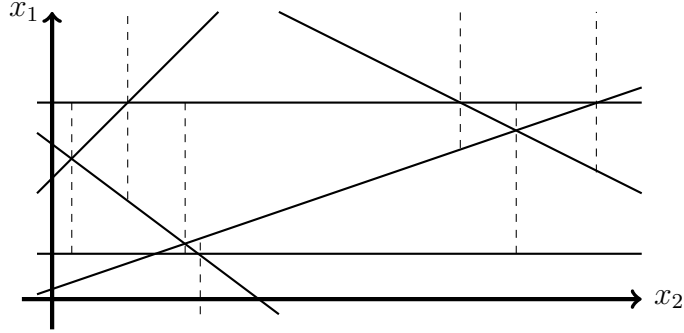
Our general case is slightly more complicated since the functions $x_2 \mapsto h_i(x_2, s)$ are not linear and even not continuous, just piecewise continuous, so their graphs may intersect without one crossing another.

For $i \in [l]$ and $s \in M^{|y|}$ we will denote by $\hat{u}_i(s)$ the corresponding vertical line

$$\hat{u}_i(s) := \{(x_1, x_2) \in M^2 \colon x_2 = u_i(s)\},$$

and also for $i \in [k]$ and $s \in M^{|y|}$ we will denote by $\hat{h}_i(s)$ the "curve"

$$\hat{h}_i(s) := \{(x_1, x_2) \in M^2 \colon x_1 = h_i(x_2, s)\}.$$

For $i, j \in [k], s_1, s_2 \in M^{|y|}$ and $(a, b) \in \mathbb{R}^2$ we say that $\hat{h}_i(s_1)$ and $\hat{h}_j(s_2)$ *properly intersect* at (a, b) if $(a, b) \in \hat{h}_i(s_1) \cap \hat{h}_j(s_2)$ and $\hat{h}_i(s_1), \hat{h}_j(s_2)$ have different germs at (a, b). Formally it means that $a = h_i(b, s_1) = h_j(b, s_2)$ and for any $\varepsilon > 0$ there is $b' \in (b - \varepsilon, b + \varepsilon)$ with $h_i(b', s_1) \neq h_j(b', s_2)$. We will denote by $\hat{h}_i(s_1) \sqcap \hat{h}_j(s_2)$ the set of all points $(a, b) \in M^2$ where $\hat{h}_i(s_1)$ and $\hat{h}_j(s_2)$ intersect properly. It is easy to see using o-minimality that the set $\hat{h}_i(s_1) \sqcap \hat{h}_j(s_2)$ is finite and there is $N_l \in \mathbb{N}$ such that $|\hat{h}_i(s_1) \sqcap \hat{h}_j(s_2)| \leq N_l$ for all $i, j \in [k]$

and $s_1, s_2 \in M^{|y|}$. Also all points in $\hat{h}_i(s_1) \sqcap \hat{h}_j(s_2)$ are definable over s_1, s_2, i.e. there are definable functions $f_{i,j}^m(y_1, y_2)$ with $m \in [N_l]$ such that for all s_1, s_2 the set $\hat{h}_i(s_1) \sqcap \hat{h}_j(s_2)$ is either empty or it is exactly $\{f_{ij}^m(s_1, s_2) \colon m \in [N_l]\}$.

We will construct a definable cell decomposition $\mathcal{T}(S)$ for $\Phi(x; y)$ as a union of 5 families of cells:
- $\mathcal{T}_0(S)$ – 0-dimensional cells, i.e. points;
- $\mathcal{T}_1^u(S)$ – 1-dimensional "vertical" cells;
- $\mathcal{T}_1^e(S)$ – extra 1-dimensional vertical cells;
- $\mathcal{T}_1^h(S)$ – 1-dimensional "horizontal" cells;
- $\mathcal{T}_2(S)$ – 2-dimensional cells.

For each family $\mathcal{T}_\star^\star(S)$ we will have $|\mathcal{T}_\star^\star(S)| = O(|S|^2)$, and also we will have appropriate uniformly definable $\Psi_\star^\star(S)$ and $\mathcal{I}_\star^\star(\Delta)$ so that $\mathcal{T}_\star^\star(S) = \{\Delta \in \Psi_\star^\star(S) \colon \mathcal{I}_\star^\star(\Delta) \cap S = \emptyset\}$.

The family $\mathcal{T}_0(\mathbf{S})$. We take $\mathcal{T}_0(S)$ to be the set of all points of intersections of vertical lines $\hat{u}_i(s)$ and curves $\hat{h}_j(s')$ together with all points where curves $\hat{h}_i(s)$ and $\hat{h}_j(s')$ intersect properly. I.e.,

$$\mathcal{T}_0(S) = \cup\{(\hat{u}_i(s) \cap \hat{h}_j(s') \colon i \in [l]; j \in [k]; s, s' \in S\}$$
$$\cup \{\hat{h}_i(s) \sqcap \hat{h}_j(s') \colon i, j \in [k]; s, s' \in S\}.$$

We take $\Psi_0(S) := \mathcal{T}(S)$ and $\mathcal{I}_0(\Delta) := \emptyset$.

It is easy to see that $\Psi_0(S)$ is uniformly definable.

We also have $|\mathcal{T}_0(S)| \leq kl|S|^2 + N_l k^2 |S|^2 = O(|S|^2)$.

The set $\mathcal{T}_1^{\mathbf{u}}(\mathbf{S})$. For fixed $i \in [l]$ and $s \in S$ let I_i^s be the set of all definably connected components of $\hat{u}_i(s) \setminus \mathcal{T}_0(S)$.

Since

$$\hat{u}_i(s) \cap \mathcal{T}_0(S) = \{\hat{u}_i(s) \cap \hat{h}_j(s') \colon j \in [k], s' \in S\},$$

we have $|I_i^s| \leq (k+1)|S|$, and every $\Delta \in I_i^s$ has form

$$\Delta = \{(x_1, x_2) \in M^2 \colon x_2 = u_i(s); h_j(s_1) < x_1 < h_{j'}(s_2)\},$$

for some $j, j' \in \{0, \ldots, k+1\}$, and $s_1, s_2 \in S$.

We take $\mathcal{T}_1^u(S)$ to be the union of all I_i^s for $i \in [l]$ and $s \in S$. Clearly $|\mathcal{T}_1^u(S)| \leq l(k+1)|S|^2 = O(|S|^2)$.

We take $\Psi_1^u(S)$ to be the set of all vertical lines segments of the form $\{(x_1, x_2) \in M^2 \colon x_2 = u_i(s); h_j(s_1) < x_1 < h_{j'}(s_2)\}$, for $i \in [l]$, $j, j' \in \{0, \ldots, k+1\}$, $s, s_1, s_2 \in S$. For $\Delta \in \Psi_1^u(S)$ we take $\mathcal{I}_1^u(\Delta) := \{s \in M^{|y|} \colon \Phi(x; s) \text{ crosses } \Delta\}$.

It is not hard to see that Ψ_1^u and $\mathcal{I}_1^u$ are uniformly definable and $\mathcal{T}_1^u(S) = \{\Delta \in \Psi_1^u(S) \colon \mathcal{I}_1^u(\Delta) \cap S = \emptyset\}$.

The set $\mathcal{T}_1^{\mathbf{e}}(\mathbf{S})$. For each point where two horizontal curves intersect properly we add two vertical line segments: one from the point to the curve above (or to plus infinity if there is no curve above) and one to the curve below (or to minus infinity if there is no curve below).

Let $i, j \in [k]$, $s, s_1 \in S$ and $p = (p_1, p_2) \in \hat{h}_i(s) \sqcap \hat{h}_j(s_1)$.

Let

$$p^+ := \inf\{h_m(p_2, s'): m = 1, \ldots, k+1; s' \in S; h_m(p_2, s') > p_1\},$$

and

$$p^- := \sup\{h_m(p_2, s'): m = 0, \ldots, k; s' \in S; h_m(p_2, s') < p_1\}.$$

We define $I_p^+ := \{(x_1, x_2) \in M^2: x_2 = p_2; p_1 < x_1 < p^+\}$, $I_p^- := \{(x_1, x_2) \in M^2: x_2 = p_2; p^- < x_1 < p_1\}$; and take

$$\mathcal{T}_1^e(S) := \{I_p^+, I_p^-: p \in \hat{h}_i(s) \sqcap \hat{h}_j(s_1); i, j \in [k]; s, s_1 \in S\}.$$

Obviously $|\mathcal{T}_1^e(S)| \le 2N_l k^2 |S|^2 = O(|S|^2)$.

We take $\Psi_1^e(S)$ to be the family of all sets of the form

$$\{(x_1, x_2) \in M^2: x_2 = p_2; p_1 < x_1 < h_m(p_2, s')\}$$

for all $i, j \in [k]$, $m \in \{1, \ldots, k+1\}$, $s, s_1, s' \in S$, and $p = (p_1, p_2) \in \hat{h}_i(s) \sqcap \hat{h}_j(s_1)$; and of the form

$$\{(x_1, x_2) \in M^2: x_2 = p_2; h_m(p_2, s') < x_1 < p_1\}$$

for all $i, j \in [k]$, $m \in \{0, \ldots, k\}$, $s, s_1, s' \in S$, and $p = (p_1, p_2) \in \hat{h}_i(s) \sqcap \hat{h}_j(s_1)$. It is not hard to see that $\Psi(S)$ is uniformly definable.

For $\Delta \in \Psi_1^e(S)$ we take $\mathcal{I}_1^e(\Delta) := \{s \in M^{|y|}: \Phi(x; s) \text{ crosses } \Delta\}$. It is not hard to see $\mathcal{I}_1^e(\Delta)$ is uniformly definable and $\mathcal{T}_1^e(S) = \{\Delta \in \Psi_1^e(S): \mathcal{I}_1^e(\Delta) \cap S = \emptyset\}$.

The set $\mathcal{T}_1^{\mathbf{h}}(\mathbf{S})$. Given $i \in [k]$ and $s \in S$, let J_i^s be the set of all definably connected components of $\hat{h}_i(s) \setminus \mathcal{T}_0(S)$. It is easy to see that

$$\hat{h}_i(s) \cap \mathcal{T}_0(S) = \{\hat{h}_i(s) \cap \hat{u}_j(s'): j \in [l]; s' \in S\}$$
$$\cup \{\hat{h}_i(s) \sqcap \hat{h}_j(s'): j \in [k]; s' \in S\}.$$

In particular $|J_i^s| \le (l + N_l k + 1)|S|$.

We take $\mathcal{T}_1^h(S)$ to be the union of all J_i^s for $i \in [k]$, $s \in S$. Clearly $|\mathcal{T}_1^h(S)| \le k(l + N_l k + 1)|S|^2 = O(|S|^2)$.

Given $i \in [k]$, $s \in S$ and $s_1, s_2 \in S$ let $\mathcal{A}_{i,s}[s_1, s_2]$ be the family of all sets of the form $\{(x_1, x_2) \in \hat{h}_i(s); c_1 < x_2 < c_2\}$, with

$$c_1, c_2 \in \{u_j(s_1): j \in [l]\}$$
$$\cup \{p_2: (p_1, p_2) \in \hat{h}_i(s) \sqcap \hat{h}_j(s_2) \text{ for some } p_1\} \cup \{\pm\infty\}.$$

We take $\Psi_1^h(S)$ to be the union of all $\mathcal{A}_{i,s}[s_1, s_2]$ with $i \in [k]$ and $s, s_1, s_2 \in S$. It is not hard to see that $\Psi_1^h(S)$ is uniformly definable and $\mathcal{T}_1^h(S) = \{\Delta \in \Psi_1^h(S) \colon \mathcal{I}_1^h(\Delta) \cap S = \emptyset\}$, where $\mathcal{I}_1^h(\Delta) = \{s \in M^{|y|} \colon \Phi(x; s)$ crosses $\Delta\}$.

The set $\mathcal{T}_2(\mathbf{S})$. For the family $\mathcal{T}_2(S)$ we take all definably connected components of $M^2 \setminus (\mathcal{T}_0(S) \cup \mathcal{T}_1^u(S) \cup \mathcal{T}_1^e(S) \cup \mathcal{T}_1^h(S))$.

Given $i, j \in \{0, \ldots, k+1\}$, $s_1, s_2 \in S$ and $c_1 < c_2 \in M \cup \{\pm\infty\}$ with $h_i(x_2, s_1) < h_j(x_2, s_2)$ for all $x_2 \in (c_1, c_2)$, let $A_{i,s_1}^{j,s_2}(c_1, c_2)$ be the set

$$A_{i,s_1}^{j,s_2}(c_1, c_2) = \{(x_1, x_2) \in M^2 \colon c_1 < x_2 < c_2; \; h_i(x_2, s_1) < x_1 < h_j(x_2, s_2)\}.$$

It is not hard to see that if $\Delta \in \mathcal{T}_2(S)$ then $\Delta = A_{i,s_1}^{j,s_2}(c_1, c_2)$ for some $i, j \in \{0, \ldots, k+1\}$, $s_1, s_2 \in S$ and c_1, c_2 belonging to the following set:

$$S_{i,s_1}^{j,s_2} = \{u_{i'}(s') \colon i' \in \{0, \ldots l+1\}; s' \in S\}$$

$$\cup \{p_2 \colon (p_1, p_2) \in \hat{h}_i(s_1) \sqcap \hat{h}_{i'}(s') \text{ for some } i' \in [k], s' \in S, p_1 \in M\}$$

$$\cup \{p_2 \colon (p_1, p_2) \in \hat{h}_j(s_2) \sqcap \hat{h}_{i'}(s') \text{ for some } i' \in [k], s' \in S, p_1 \in M\}.$$

We take $\Psi_2(S)$ to be the family of all $A_{i,s_1}^{j,s_2}(c_1, c_2)$, for all $c_1, c_2 \in S_{i,s_1}^{j,s_2}$.

It is not hard to see that $\Psi_2(S)$ is uniformly definable family, and we have $\mathcal{T}_2(S) \subseteq \Psi_2(S)$.

It is also not hard to see that a set $\Delta \in \Psi_2(S)$ is in $\mathcal{T}_2(S)$ if and only if it is not crossed by $\Phi(x; S)$, and is also not crossed by any line segment in $\mathcal{T}_1^e(S)$.

Hence a set $\Delta = A_{i,s_1}^{j,s_2}(c_1, c_2) \in \Psi_2(S)$ **is not in** $\mathcal{T}_2(S)$ if and only if there is $s \in S$ satisfying at least one of the following conditions.
(C1) $\Phi(x; s)$ crosses Δ.
(C2) There are $i' \in [k]$ and $(p_1, p_2) \in \hat{h}_i(s_1) \sqcap \hat{h}_{i'}(s)$ with $c_1 < p_2 < c_2$.
(C3) There are $i' \in [k]$ and $(p_1, p_2) \in \hat{h}_j(s_2) \sqcap \hat{h}_{i'}(s)$ with $c_1 < p_2 < c_2$.

For $\Delta \in \Psi_2(S)$ we take $\mathcal{I}_2(\Delta)$ to be the set of all $s \in M^{|y|}$ satisfying any of the conditions $(C1) - (C3)$. It is not hard to see that $\mathcal{I}_2(\Delta)$ is uniformly definable and $\mathcal{T}_2(S) = \{\Delta \in \Psi_2(S) \colon \mathcal{I}_2(\Delta) \cap S = \emptyset\}$.

We are left to check that $|\mathcal{T}_2(S)| = O(|S|^2)$.

Since $\mathcal{T}_2(S)$ consists of definably connected components of $M^2 \setminus (\mathcal{T}_0(S) \cup \mathcal{T}_1^u(S) \cup \mathcal{T}_1^e(S) \cup \mathcal{T}_1^h(S))$, any two $\Delta, \Delta' \in \mathcal{T}_2(S)$ are either disjoint or coincide, hence every $\Delta \in \mathcal{T}_2(S)$ is completely determined by its "left lower corner", i.e. if $\Delta = A_{i,s_1}^{j,s_2}(c_1, c_2)$ and $\Delta' = A_{i,s_1}^{j',s_2'}(c_1, c_2')$ are in $\mathcal{T}_2$ then $\Delta = \Delta'$.

We divide $\mathcal{T}_2(S)$ into 4 disjoint families:
• The family $F_1(S)$ of all $A_{i,s_1}^{j,s_2}(c_1, c_2) \in \mathcal{T}_2(S)$ with $c_1 = -\infty$.

- The family $F_2(S)$ of all $A^{j,s_2}_{i,s_1}(c_1, c_2) \in \mathcal{T}_2(S)$ with $c_1 = u_{i'}(s')$ for some $i' \in [l]$ and $s' \in S$.
- The family $F_3(S)$ of all $A^{j,s_2}_{i,s_1}(c_1, c_2) \in \mathcal{T}_2(S)$ that are not in $F_2(S)$ and $(p_1, c_1) \in \hat{h}_i(s_1) \sqcap \hat{h}_{i'}(s')$ for some $i' \in [k]$, $s' \in S$, and $p_1 \in M$.
- The family $F_4(S)$ of all $A^{j,s_2}_{i,s_1}(c_1, c_2) \in \mathcal{T}_2(S)$ that are not in $F_1(S) \cup F_2(S) \cup F_3(S)$. In this case we have that $\{(x_1, c_1) : h_i(c_1, s_1) < x_1 < h_j(c_1, s_2)\} \in \mathcal{T}_1^e(S)$.

Every $A^{j,s_2}_{i,s_1}(c_1, c_2) \in F_1(S)$ is completely determined by i and s_1, hence $|F_1(S)| \leq (k+1)|S|$ (we get $k+1$, since we allow $i = 0$).

Every $A^{j,s_2}_{i,s_1}(c_1, c_2) \in F_2(S)$ is completely determined by i, s_1, some $i' \in [l]$ and $s' \in S$. Hence $|F_2(S)| \leq (k+1)l|S|^2$.

Since $\hat{h}_i(s_1) \sqcap \hat{h}_{i'}(s') \leq N_l$ we have $|F_3(S)| \leq k^2 N_l |S|^2$.

Finally, each $A^{j,s_2}_{i,s_1}(c_1, c_2) \in F_4(S)$ is completely determined by its "left side" $\{(x_1, c_1) : h_i(c_1, s_1) < x_1 < h_j(c_1, s_2)\}$ that is in $\mathcal{T}_1^e(S)$. Since $|\mathcal{T}_1^e(S)| = O(|S|^2)$, we also have $|F_4(S)| = O(|S|^2)$.

Therefore $|\mathcal{T}_2(S)| = O(|S|^2)$.

Taking $\mathcal{T}(S) = \mathcal{T}_0(S) \cup \mathcal{T}_1^u(S) \cup \mathcal{T}_1^e(S) \cup \mathcal{T}_1^h(S) \cup \mathcal{T}_2(S)$ we obtain a definable cell decomposition for $\Phi(x; y)$ with $|\mathcal{T}(S)| = O(|S|^2)$.

5. Planar Zarankiewicz's problem in distal structures

5.1. **Zarankiewicz's problem.** Zarankiewicz's problem in graph theory asks to determine the largest possible number of edges in a bipartite graph on a given number of vertices that has no complete bipartite subgraphs of a given size.

In [10] the authors investigate Zarankiewicz's problem for semialgebraic graphs of bounded description complexity, a setting which in particular subsumes a lot of different incidence-type questions.

In particular, they prove the following upper bound on the number of edges (they have more general results in $\mathbb{R}^n$ for arbitrary n as well, but here we will be only concerned with the "planar" case).

Fact 5.1. [10, Theorem 1.1] *Let $E \subseteq \mathbb{R}^2 \times \mathbb{R}^2$ be a semi-algebraic relation such that E has description complexity at most t (i.e., E can be defined as a Boolean combination of at most t polynomial inequalities, with all of the polynomials involved of degree at most t). Then for any $k \in \mathbb{N}$ there is some constant $c = c(t, k)$ satisfying the following.*

If $P, Q \subseteq \mathbb{R}^2$ with $|P| = m, |Q| = n$ are such that $E \cap (P \times Q)$ doesn't contain a copy of $K_{k,k}$ (the complete bipartite graph with both parts of size k), then

$$|E(P, Q)| \leq c\left((mn)^{\frac{2}{3}} + m + n\right),$$

where $E(P, Q) = E \cap (P \times Q)$.

Remark 5.2. This result is a natural generalization of the Szemerédi-Trotter theorem over $\mathbb{R}$ [14]. Namely, if P a set of points on the plane, Q the dual of the lines (i.e. lines are semialgebraically coded by points in $\mathbb{R}^2$), and E the incidence relationship (which is also clearly semialgebraic), then $E(P, Q)$ is $K_{2,2}$-free as any two distinct lines intersect in at most one point.

We will give a common generalization of Fact 5.1 and the semialgebraic "points / planar curves" incidence bound from [12, Theorem 4] to arbitrary definable families admitting a quadratic distal cell decomposition (e.g. any definable family of subsets of M^2 in an o-minimal expansion of a field). To state the result, we first recall the notion of the VC-density of a formula (and refer to [2] for a detailed discussion).

Definition 5.3. (1) Given a set X and a family $\mathcal{F}$ of subsets of X, the *shatter function* $\pi_{\mathcal{F}} : \mathbb{N} \to \mathbb{N}$ of $\mathcal{F}$ is defined as

$$\pi_{\mathcal{F}}(n) := \max\{|\mathcal{F} \cap A| : A \subseteq X, |A| = n\},$$

where $\mathcal{F} \cap A = \{S \cap A : S \in \mathcal{F}\}$.
(2) The *VC-density* of $\mathcal{F}$, or $\mathrm{vc}(\mathcal{F})$, is defined as the infimum of all real numbers r such that $\pi_{\mathcal{F}}(n) = O(n^r)$ (and $\mathrm{vc}(\mathcal{F}) = \infty$ if there is no such r).
(3) Given a formula $\varphi(x; y)$, we define its VC density $\mathrm{vc}(\varphi)$ as the VC-density of the family $\{\varphi(M, b) : b \in M^{|y|}\}$ of subsets of $M^{|x|}$.
(4) Given a formula $\varphi(x; y)$, we consider its dual formula $\varphi^*(y; x) := \varphi(x; y)$ obtained by interchanging the roles of the variables. It is easy to see then that the family $\{\varphi^*(y; a) : a \in M^{|x|}\}$ of subsets of $M^{|y|}$ is the dual set system for the family $\{\varphi^*(x; b) : b \in M^{|y|}\}$ of subsets of $M^{|x|}$.

VC-density in various classes of NIP structures is investigated e.g. in [1, 2], and the optimal bounds are known in some cases including the o-minimal structures.

Fact 5.4. [2, Theorem 6.1] *Let $\mathcal{M}$ be an o-minimal structure, and let $\varphi(x; y)$ be any formula. Then $\mathrm{vc}(\varphi^*) \leq |x|$.*

Remark 5.5. Let $\varphi(x; y)$ be a formula admitting a distal cell decomposition $\mathcal{T}$ with $|\mathcal{T}(S)| = O(|S|^d)$. Then $\mathrm{vc}(\varphi^*) \leq d$.

Indeed, recalling Definition 2.6, given any finite $S \subseteq M^{|y|}$ and $\Delta \in \mathcal{T}(S)$, $S \cap \varphi^*(M, a) = S \cap \varphi^*(M, a')$ for any $a, a' \in \Delta$ (and the sets in $\mathcal{T}(S)$ give a covering of $M^{|x|}$), hence at most $|S|^d$ different subsets of S are cut out by the instances of $\varphi^*(y; x)$.

We will need the following weaker bound that applies to graphs of bounded VC-density.

Fact 5.6. [10, Theorem 2.1] *For every c, d, k there is some constant $c_1 = c_1(c, d, k)$ such that the following holds.*

Let $E \subseteq P \times Q$ be a bipartite graph with $|P| = m, |Q| = n$ such that the family of sets $\mathcal{F} = \{E(q) : q \in Q\}$ satisfies $\pi_{\mathcal{F}}(z) \leq cz^d$ for all $z \in \mathbb{N}$ (where $E(q) = \{p \in P : (p, q) \in E\}$). Then if E is $K_{k,k}$-free, we have

$$|E(P, Q)| \leq c_1(mn^{1-1/d} + n).$$

We are ready to prove the main theorem of this section.

Theorem 5.7. *Let $\mathcal{M}$ be a structure, and assume that $E(x, y)$ is a formula admitting a distal cell decomposition $\mathcal{T}$ with $|\mathcal{T}(S)| = O(|S|^2)$ and such that $\mathrm{vc}(E) \leq d$. Then for any $k \in \mathbb{N}$ there is a constant $c = c(E, k)$ satisfying the following.*

For any finite $P \subseteq M^{|x|}, Q \subseteq M^{|y|}$, $|P| = m, |Q| = n$, if $E(P, Q)$ is $K_{k,k}$-free, then we have:

$$|E(P, Q)| \leq c \left(m^{\frac{d}{2d-1}} n^{\frac{2d-2}{2d-1}} + m + n \right).$$

Proof. Our argument is a generalization of the proofs of [10, Theorem 3.2] and [12, Theorem 4].

If $n > m^d$, then by Fact 5.6 and assumption we have

$$|E(P, Q)| \leq c_0(mn^{1-\frac{1}{d}} + n) \leq c_0(n^{\frac{1}{d}} n^{1-\frac{1}{d}} + n) = 2c_0 n$$

for some $c_0 = c_0(E, d, k)$, and we are done. Hence we assume $n \leq m^2$.

Let $r := \dfrac{m^{\frac{d}{2d-1}}}{n^{\frac{1}{2d-1}}}$, and consider the family $\Sigma = \{E(M, q) : q \in Q\}$ of subsets of $M^{|x|}$. By assumption and Theorem 3.1, there is a family $\mathcal{C}$ of subsets of $M^{|x|}$ giving a $1/r$-cutting with respect to Σ. That is, $M^{|x|}$ is covered by the union of the sets in $\mathcal{C}$ and any of the sets $C \in \mathcal{C}$ is crossed by at most $|\Sigma|/r$ elements from Σ. Moreover, $|\mathcal{C}| \leq c_1 r^2$ for some $c_1 = c_1(E)$.

Then there is a set $C \in \mathcal{C}$ containing at least $\dfrac{m}{c_1 r^2} = \dfrac{n^{\frac{2}{2d-1}}}{c_1 m^{\frac{1}{2d-1}}}$ points from P. Let $P' \subseteq P \cap C$ be a subset of size exactly $\left\lceil \dfrac{n^{\frac{2}{2d-1}}}{c_1 m^{\frac{1}{2d-1}}} \right\rceil$.

If $|P'| < k$, we have $\dfrac{n^{\frac{2}{2d-1}}}{c_1 m^{\frac{1}{2d-1}}} \leq |P'| < k$, so $n < k^{\frac{2d-1}{2}} c_1^{\frac{2d-1}{2}} m^{\frac{1}{2}}$. Then by Fact 5.6 applied to the dual formula E^* (and using Remark 5.5) we have

$$|E(P, Q)| \leq c_2(nm^{1-\frac{1}{2}} + m) \leq c_2(k^{\frac{2d-1}{2}} c_1^{\frac{2d-1}{2}} m^{\frac{1}{2}} m^{\frac{1}{2}} + m) \leq c_3 m$$

for some $c_3 = c_3(E, k)$, so we are done.

Hence we may assume that $|P'| \geq k$. Let Q' be the set of all points $q \in Q$ such that $E(M, q)$ crosses C. We know that

$$|Q'| \leq \frac{|Q|}{r} \leq \frac{n n^{\frac{1}{2d-1}}}{m^{\frac{d}{2d-1}}} = \frac{n^{\frac{2d}{2d-1}}}{m^{\frac{d}{2d-1}}} \leq c_1^d |P'|^d.$$

Again by Fact 5.6 we get

$$|E(P', Q')| \leq c_4(|P'||Q'|^{1-\frac{1}{d}} + |Q'|) \leq c_4(|P'|c_1^d|P'|^{d-1} + c_1^d|P'|^d) \leq c_5|P'|^d$$

for some $c_5 = c_5(E, k)$. Hence there is a point $p \in P'$ such that $|E(p) \cap Q'| \leq c_5|P'|^{d-1}$.

Since $E(P, Q)$ is $K_{k,k}$-free, there are at most $k-1$ points in $Q \setminus Q'$ from $E(p)$ (otherwise, since none of those points crosses C and C contains P', which is of size $\geq k$, we would have a copy of $K_{k,k}$). Hence

$$|E(p)| \leq c_5|P'| + (k-1) \leq \frac{c_5}{c_1} \frac{n^{\frac{2(d-1)}{2d-1}}}{m^{\frac{d-1}{2d-1}}} + (k-1).$$

We remove p and repeat the argument until we have no vertices remaining in P, and see that

$$|E(P, Q)| \leq (c/2)(n+m) + \sum_{i=n^{\frac{1}{d}}}^{m} \left(\frac{c_5}{c_1} \frac{n^{\frac{2(d-1)}{2d-1}}}{i^{\frac{d-1}{2d-1}}} + (k-1) \right)$$

$$\leq (c/2)(n+m) + \frac{c_5}{c_1} n^{\frac{2(d-1)}{2d-1}} \sum_{i=n^{\frac{1}{d}}}^{m} \frac{1}{i^{\frac{d-1}{2d-1}}} + (k-1)m$$

$$\leq (c/2)(n+m) + \frac{c_5}{c_1} n^{\frac{2(d-1)}{2d-1}} m^{1-\frac{d-1}{2d-1}} + (k-1)m$$

$$\leq c(m^{\frac{d}{2d-1}} n^{\frac{2d-2}{2d-1}} + m + n)$$

for a sufficiently large $c = c(E, k)$. $\square$

5.2. Omitting $K_{k,k}$ versus omitting infinite complete bipartite graphs. We recall a theorem of Bukh and Matoušek.

Fact 5.8. [4, Theorem 1.9] *For every d, D and k there exists N such that for every semialgebraic relation $R(x_1, \ldots, x_k)$ with $|x_1| = \ldots = |x_k| = d$ of description complexity D, the following two conditions are equivalent.*
(1) There exist $A_1, \ldots, A_k \subseteq \mathbb{R}^d$ such that $|A_1| = \ldots = |A_k| = N$ and $A_1 \times \ldots \times A_k \subseteq R$.
(2) There exist infinite *sets $A_1, \ldots, A_k \subseteq \mathbb{R}^d$ such that $A_1 \times \ldots \times A_k \subseteq R$.*

We give a generalization of this result for any distal structure in which finite sets in every definable family have a uniform bound on their size. Recall:

Definition 5.9. An $\mathcal{L}$-structure $\mathcal{M}$ *eliminates* $\exists^\infty$ if for every $\varphi(x, y) \in \mathcal{L}$ there is some $n_\varphi \in \mathbb{N}$ such that for any $b \in M^{|y|}$, $\varphi(M, b)$ is infinite if and only if $|\varphi(M, b)| \geq n_\varphi$.

We will use the definable strong Erdős-Hajnal property for hypergraphs in distal structures from [8] (and we will use some terminology from that paper in our argument).

Fact 5.10. [8, Corollary 4.6] *Let $\mathcal{M}$ be a distal $\mathcal{L}$-structure. Then for every formula $\varphi(x_1, \ldots, x_k; z) \in \mathcal{L}$ there are some $\alpha > 0$ and formulas $\psi_i(x_i, y_i) \in \mathcal{L}$ for $1 \leq i \leq k$ such that the following holds.*
For any generically stable Keisler measures μ_i on $M^{|x_i|}$ and any $c \in M^{|z|}$, there are some $b_i \in M^{|y_i|}$ such that $\mu_i(\psi_i(M^{|x_i|}, b_i)) \geq \alpha$ and either $\prod_{1 \leq i \leq k} \psi_i(M^{|x_i|}, b_i) \subseteq \varphi(M^{|x_1|}, \ldots, M^{|x_k|}; c)$ or $\prod_{1 \leq i \leq k} \psi_i(M^{|x_i|}, b_i) \subseteq \neg\varphi(M^{|x_1|}, \ldots, M^{|x_k|}; c)$.

Theorem 5.11. *Let $\mathcal{M}$ be a distal $\mathcal{L}$-structure eliminating $\exists^\infty$. Then for any formula $\varphi(x_1, \ldots, x_k; z) \in \mathcal{L}$ there is some $N \in \mathbb{N}$ and $\psi_i(x_i, y_i) \in \mathcal{L}$, for $1 \leq i \leq k$, such that the following are equivalent for any $c \in M^{|z|}$, letting $R \subseteq M^{|x_1|} \times \ldots \times M^{|x_k|}$ be given by $R := \varphi(M^{|x_1|}, \ldots, M^{|x_k|}, c)$.*
(1) There exist $A_i \subseteq M^{|x_i|}$ for $1 \leq i \leq k$ such that $|A_1| = \ldots = |A_k| = N$ and $A_1 \times \ldots \times A_k \subseteq R$.
(2) There are some $b_i \in M^{|y_i|}$ such that $\psi_i(M^{|x_i|}, b_i)$ is infinite for all $1 \leq i \leq k$ and $\psi_1(M^{|x_1|}, b_1) \times \ldots \times \psi_k(M^{|x_k|}, b_k) \subseteq R$.

Proof. Let $\alpha > 0$ and $\psi_i(x_i, y_i) \in \mathcal{L}$, for $1 \leq i \leq k$, be as given by Fact 5.10 for $\varphi(x_1, \ldots, x_k; z)$. Let $n_i \in \mathbb{N}$ be as given by Definition 5.9 for $\psi_i(x_i, y_i)$, and let $n := \max\{n_i : 1 \leq i \leq k\}$. We take $N := \lceil \frac{n}{\alpha} \rceil$, then $N = N(\varphi)$.

Let $c \in M^{|z|}$ be arbitrary, and let $R := \varphi(M^{|x_1|}, \ldots, M^{|x_k|}, c)$. Assume that (1) holds. That is, there are some $A_i \subseteq M^{|x_i|}$ such that $|A_1| = \ldots = |A_k| = N$ and $A_1 \times \ldots \times A_k \subseteq R$. Let μ_i be a Keisler measure on $M^{|x_i|}$ defined by $\mu_i(X) := \frac{|A_i \cap X|}{|A_i|}$ for all definable $X \subseteq M^{|x_i|}$, then μ_i is generically stable for all $1 \leq i \leq k$. Applying Fact 5.10, we find some $b_i \in M^{|y_i|}$ such that $\mu_i(\psi_i(M^{|x_i|}, b_i)) \geq \alpha$ and $\prod_{1 \leq i \leq k} \psi_i(M^{|x_i|}, b_i) \subseteq R$ (note that $\prod_{1 \leq i \leq k} \psi_i(M^{|x_i|}, b_i) \subseteq \neg R$ is impossible as $\prod_{1 \leq i \leq k} A_i \subseteq R$). Now for any $1 \leq i \leq k$, $\mu_i(\psi_i(M^{|x_i|}, b_i)) \geq \alpha$ implies $|\psi_i(A_i, b_i)| \geq \alpha N \geq n_i$, hence $\psi_i(M^{|x_i|}, b_i)$ is infinite by the choice of n_i, as wanted. $\square$

Remark 5.12. Examples of structures satisfying the assumption of Theorem 5.11 are given by arbitrary o-minimal structures and p-minimal structures (e.g. the field $\mathbb{Q}_p$). Hence Fact 5.8 follows by applying it to the field of reals.

5.3. **The o-minimal case.** Theorem 5.11 implies that in Theorem 5.7, assuming $\mathcal{M}$ eliminates $\exists^\infty$, we can relax the assumption to just assuming that E doesn't contain a copy of an infinite complete bipartite graph.

We conclude by observing that all of these results apply to o-minimal expansions of fields.

Theorem 5.13. *Let $\mathcal{M}$ be an o-minimal expansion of a field and let $E(x; y) \subseteq M^2 \times M^d$ be a definable relation.*
(1) For every $k \in \mathbb{N}$ there is a constant $c = c(E, k)$ such that for any finite $P \subseteq M^2, Q \subseteq M^d$, $|P| = m, |Q| = n$, if $E(P, Q)$ does not contain a copy of $K_{k,k}$ (the complete bipartite graph with two parts of size k), then we have

$$|E(P, Q)| \le c \left(m^{\frac{d}{2d-1}} n^{\frac{2d-2}{2d-1}} + m + n \right).$$

(2) There is some $k' \in \mathbb{N}$ and formulas $\varphi(x, v), \psi(y, w)$ depending only on E such that if E contains a copy of $K_{k',k'}$, then there are some parameters $b \in M^v, c \in M^w$ such that both $\varphi(M, b)$ and $\psi(M, c)$ are infinite and $\varphi(M, b) \times \psi(M, c) \subseteq E$.

Proof. (1) Follows by applying Theorem 5.7. Its assumptions are satisfied for E by Theorem 4.1 and by Fact 5.4 (applied to the dual formula $E^*(y; x)$).

(2) Follows by Theorem 5.11 as o-minimal theories eliminate the $\exists^\infty$ quantifier.

$\square$

Corollary 5.14. *In the setting of Theorem 5.13, there is a constant c and formulas $\varphi(x, v), \psi(y, w)$ depending only on E such that either*

$$|E(P, Q)| \le c \left(m^{\frac{d}{2d-1}} n^{\frac{2d-2}{2d-1}} + m + n \right)$$

for all finite $P \subseteq M^2, Q \subseteq M^d$ with $|P| = m, |Q| = n$, or there are some $b \in M^v, c \in M^w$ such that both $\varphi(M, b)$ and $\psi(M, c)$ are infinite and $\varphi(M, b) \times \psi(M, c) \subseteq E$.

Proof. Immediate combining (1) and (2) in Theorem 5.13 (let k', φ, ψ be as given by (2) for $E(x, y)$, and let c be as given by (1) for this k').
$\square$

Remark 5.15. The special case with $d = 2$ and E satisfying an additional assumption of 1-dimensionality of its fibers was obtained independently by Basu and Raz [3] using different methods.

REFERENCES

[1] Matthias Aschenbrenner, Alf Dolich, Deirdre Haskell, Dugald Macpherson, and Sergei Starchenko, *Vapnik–Chervonenkis density in some theories without the independence property, II*, Notre Dame Journal of Formal Logic **54** (2013), no. 3-4, 311–363.

[2] ———, *Vapnik–Chervonenkis density in some theories without the independence property, I*, Transactions of the American Mathematical Society **368** (2016), no. 8, 5889–5949.

[3] Saugata Basu and Orit E Raz, *An o-minimal Szemerédi-Trotter theorem*, Preprint, arXiv:1611.07362 (2016).

[4] Boris Bukh and Jiří Matoušek, *Erdős–Szekeres-type statements: Ramsey function and decidability in dimension* 1, Duke Mathematical Journal **163** (2014), no. 12, 2243–2270.

[5] Bernard Chazelle, *Cuttings*, Handbook of data structures and applications, 2005, pp. 25.1–25.10.

[6] Bernard Chazelle, Herbert Edelsbrunner, Leonidas J Guibas, and Micha Sharir, *A singly exponential stratification scheme for real semi-algebraic varieties and its applications*, Theoretical Computer Science **84** (1991), no. 1, 77–105.

[7] Artem Chernikov and Pierre Simon, *Externally definable sets and dependent pairs II*, Transactions of the American Mathematical Society **367** (2015), no. 7, 5217–5235.

[8] Artem Chernikov and Sergei Starchenko, *Regularity lemma for distal structures*, Journal of the European Mathematical Society, accepted (arXiv:1507.01482) (2015).

[9] Kenneth L Clarkson and Peter W Shor, *Applications of random sampling in computational geometry. II*, Discrete & Computational Geometry. An International Journal of Mathematics and Computer Science **4** (1989), no. 5, 387–421.

[10] Jacob Fox, János Pach, Adam Sheffer, Andrew Suk, and Joshua Zahl, *A semi-algebraic version of Zarankiewicz's problem*, Journal of the European Mathematical Society, accepted (arXiv:1407.5705) (2014).

[11] Jiří Matoušek, *Lectures on discrete geometry*, Graduate Texts in Mathematics, vol. 212, Springer-Verlag, New York, 2002. MR1899299

[12] János Pach and Micha Sharir, *Repeated angles in the plane and related problems*, Journal of Combinatorial Theory, series A **59** (1992), no. 1, 12–22.

[13] Pierre Simon, *Distal and non-distal NIP theories*, Annals of Pure and Applied Logic **164** (2013), no. 3, 294 –318.

[14] Endre Szemerédi and William T Trotter Jr, *Extremal problems in discrete geometry*, Combinatorica **3** (1983), no. 3-4, 381–392.

Department of Mathematics, University of California Los Angeles, Los Angeles, CA 90095-1555

E-mail address: chernikov@math.ucla.edu

Department of Mathematics, University of Notre Dame, Notre Dame, IN 46556

E-mail address: dgalvin1@nd.edu

Department of Mathematics, University of Notre Dame, Notre Dame, IN 46556

E-mail address: Starchenko.1@nd.edu