

## RAMSEY GROWTH IN SOME NIP STRUCTURES

ARTEM CHERNIKOV<sup>1</sup>, SERGEI STARCHENKO<sup>2</sup> AND  
MARGARET E. M. THOMAS<sup>3</sup><sup>1</sup>*Department of Mathematics, University of California Los Angeles, Los Angeles,  
CA 90095-1555, USA (chernikov@math.ucla.edu)*<sup>2</sup>*Department of Mathematics, University of Notre Dame, Notre Dame, IN 46556,  
USA (Starchenko.1@nd.edu)*<sup>3</sup>*Zukunftskolleg, Department of Mathematics and Statistics, University of  
Konstanz, Box 216, 78457 Konstanz, Germany  
(margaret.thomas@uni-konstanz.de)*

(Received 15 May 2017; revised 18 January 2019; accepted 23 January 2019)

**Abstract** We investigate bounds in Ramsey’s theorem for relations definable in NIP structures. Applying model-theoretic methods to finitary combinatorics, we generalize a theorem of Bukh and Matousek (*Duke Mathematical Journal* **163**(12) (2014), 2243–2270) from the semialgebraic case to arbitrary polynomially bounded  $\mathcal{o}$ -minimal expansions of  $\mathbb{R}$ , and show that it does not hold in  $\mathbb{R}_{\text{exp}}$ . This provides a new combinatorial characterization of polynomial boundedness for  $\mathcal{o}$ -minimal structures. We also prove an analog for relations definable in  $P$ -minimal structures, in particular for the field of the  $p$ -adics. Generalizing Conlon *et al.* (*Transactions of the American Mathematical Society* **366**(9) (2014), 5043–5065), we show that in distal structures the upper bound for  $k$ -ary definable relations is given by the exponential tower of height  $k - 1$ .

**Keywords:**  $\mathcal{o}$ -minimality; NIP; Ramsey’s theorem;  $p$ -minimality; polynomially bounded

2010 *Mathematics subject classification:* Primary 03C45; 05C35; 05D10; 05C25

## 1. Introduction

We recall a fundamental theorem of Ramsey. Let  $X$  be a set and let  $E \subseteq X^k$  be a  $k$ -ary relation on  $X$ . We say that a sequence  $(a_i : 1 \leq i \leq m)$  of elements in  $X$  with  $m \geq k$  is  $E$ -indiscernible (also called ‘ $E$ -homogeneous’ in the literature) if either  $E$  holds on all  $k$ -tuples  $(a_{i_1}, \dots, a_{i_k})$  with  $1 \leq i_1 < \dots < i_k \leq m$ , or  $E$  does not hold on any  $k$ -tuple  $(a_{i_1}, \dots, a_{i_k})$  with  $1 \leq i_1 < \dots < i_k \leq m$ .

**Fact 1.1** (Ramsey [41]). *For every  $k, n \in \mathbb{N} = \{0, 1, \dots\}$  there is some number  $N \in \mathbb{N}$  such that if  $X$  is a set and  $E \subseteq X^k$  is a  $k$ -ary relation on  $X$ , then every sequence of elements of  $X$  of length  $N$  contains an  $E$ -indiscernible subsequence of length  $n$ .*

We denote the smallest such  $N$  by  $R_k(n)$ .

Establishing exact bounds for the asymptotics of  $R_k(n)$  is one of the central open problems in combinatorics, even in the case  $k = 2$ . We summarize briefly some of the known results.

- Fact 1.2.** (1) [21, 24]  $2^{\frac{n}{2}} < R_2(n) < 2^{2n}$  for all  $n > 2$ .  
 (2) [22, 23] There are positive constants  $c$  and  $c'$  such that  $2^{cn^2} < R_3(n) < 2^{2^{c'n}}$  for all sufficiently large  $n$ .  
 (3) [14, 26] For each  $k \geq 3$  there are positive constants  $c, c'$  such that  $\text{twr}_{k-1}(cn^2) \leq R_k(n) \leq \text{twr}_k(c'n)$  for all sufficiently large  $n$ , where the tower function  $\text{twr}_k(n)$  is defined recursively by  $\text{twr}_1(n) = n$  and  $\text{twr}_{i+1}(n) = 2^{\text{twr}_i(n)}$ .

Recently, this question was investigated in the context of semialgebraic relations, where stronger bounds were obtained. Recall that a set  $A \subseteq \mathbb{R}^d$  is *semialgebraic* if it is given by a finite Boolean combination of sets of the form  $\{x \in \mathbb{R}^d : f(x) \geq 0\}$ , where  $f(x)$  is a polynomial in  $d$  variables with coefficients in  $\mathbb{R}$ . We say that a semialgebraic set  $A$  has *description complexity* at most  $t$  if  $d \leq t$  and  $A$  can be written as a Boolean combination of such sets involving at most  $t$  different polynomials, each of degree at most  $t$ .

**Definition 1.3.** Let  $E \subseteq (\mathbb{R}^d)^k$  be a  $k$ -ary semialgebraic relation on  $\mathbb{R}^d$ . For  $n \in \mathbb{N}$ , we let  $R_E(n)$  be the smallest natural number  $N$  such that if  $(a_i : 1 \leq i \leq m)$ ,  $a_i \in \mathbb{R}^d$ , is a sequence of length  $m \geq N$ , then it contains an  $E$ -*indiscernible* subsequence of length  $n$ .

Let  $R_k^{d,t}(n)$  be the maximum of  $R_E(n)$ , where  $E$  varies over all  $k$ -ary semialgebraic relations on  $\mathbb{R}^d$  of description complexity at most  $t$ .

The case of binary relations ( $k = 2$ ) is addressed in the following theorem, which shows that  $R_2^{d,t}(n)$  can be bounded by a polynomial in  $n$  – as opposed to the necessarily exponential bound in the general case (Fact 1.2(1)). The following is proved in [2, Theorem 1.2] (it is only stated there for *symmetric* semialgebraic relations; the result for arbitrary semialgebraic relations follows easily from the symmetric case using that the lexicographic ordering on  $\mathbb{R}^d$  is semialgebraic – see the discussion after Definition 6.4).

**Fact 1.4** [2, Theorem 1.2]. For any  $d, t$  there is some  $c = c(d, t)$  such that  $R_2^{d,t}(n) \leq n^c$  for all sufficiently large  $n$ .

Based on this fact, [13] addresses the case of general  $k$ , establishing that  $R_k^{d,t}(n)$  can be bounded from above by an exponential tower of height  $k - 1$  (as opposed to  $k$  for general relations; Fact 1.2(3)).

**Fact 1.5** [13]. For any  $k \geq 2$  and  $d, t \geq 1$  there is some  $c = c(k, d, t)$  such that  $R_k^{d,t}(n) \leq \text{twr}_{k-1}(n^c)$  for all sufficiently large  $n$ .

In addition, matching lower bounds for semialgebraic relations were obtained in [13] and subsequently refined in [18].

**Fact 1.6.** (1) [13] For every  $k \geq 4$ , there exists  $d = d(k)$ ,  $t = t(k)$ ,  $c' = c'(k)$  and a  $k$ -ary semialgebraic relation  $E$  on  $\mathbb{R}^d$  of description complexity  $\leq t$  such that  $R_E(n) \geq \text{twr}_{k-1}(c'n)$  for all sufficiently large  $n$ .

(2) [18] In (1), one can take  $d = k - 3$ .

The dependence of the dimension  $d$  on the arity  $k$  of the relation  $E$  in Fact 1.6 is unavoidable, due to the following theorem of Bukh and Matousek.

**Fact 1.7** [6]. *For every  $k \in \mathbb{N}$  and every  $k$ -ary semialgebraic relation  $E$  on  $\mathbb{R}$  there is some  $c = c(E)$  such that  $R_E(n) \leq 2^{2^{cn}}$  for all sufficiently large  $n$ .*

That is, if we restrict to arbitrary  $k$ -ary semialgebraic relations on  $\mathbb{R}$  (as opposed to  $\mathbb{R}^d$  for some  $d > 1$ ), then  $R_E(n)$  is at most double exponential (rather than a tower of height  $k - 1$  as in Fact 1.5). The constant  $c$  given by the proof in [6] actually depends on the parameters of  $E$  (and not just on its description complexity, as in Fact 1.5); however this dependence can be eliminated (see Theorem 1.12).

In this paper we investigate a generalization from semialgebraic relations to relations definable in more general first-order structures, and the connection between Ramsey growth for relations definable in a structure and the model-theoretic tameness conditions that this structure satisfies.

**Definition 1.8.** Let  $\mathcal{M}$  be a first-order structure in a language  $\mathcal{L}$  (we denote its underlying set by  $M$ ). Let  $k \geq 1$  be an integer and let  $\varphi(x_1, \dots, x_k)$  be an  $\mathcal{L}(\mathcal{M})$ -formula (i.e., a formula with parameters from  $\mathcal{M}$ ) with its free variables partitioned into  $k$  groups of equal size, i.e.,  $|x_1| = \dots = |x_k| = d$ . Then  $\varphi$  defines a  $k$ -ary relation  $\varphi(M)$  on  $M^d$  (a definable subset of  $M^d$  in the case  $k = 1$ ), namely  $\varphi(M) = \{(a_1, \dots, a_k) \in (M^d)^k : \mathcal{M} \models \varphi(a_1, \dots, a_k)\}$ . The case  $|x_i| = 1$  for all  $i = 1, \dots, k$  will be referred to as ‘ $x_1, \dots, x_k$  singletons’.

We let  $R_\varphi(n)$  be the smallest natural number  $N$  such that any sequence  $(a_i : 1 \leq i \leq N)$ ,  $a_i \in M^d$ , of length  $N$  contains a  $\varphi(M)$ -indiscernible subsequence of length  $n$ .

Also, given an  $\mathcal{L}$ -formula  $\varphi(x_1, \dots, x_k; z)$ , where  $|x_1| = \dots = |x_k| = d$  and  $z$  is an additional tuple of free variables, we let  $R_\varphi^*(n) := \max\{R_{\varphi(x_1, \dots, x_k; b)}(n) : b \in M^{|z|}\}$  (or  $\infty$  if the maximum does not exist).

**Remark 1.9.** By Tarski’s quantifier elimination in the field of reals  $\mathcal{M} = (\mathbb{R}, <, +, \times, 0, 1)$ , given a formula  $\varphi(x; y)$ , all sets of the form  $\varphi(\mathbb{R}^{|x|}; b)$ ,  $b \in \mathbb{R}^{|y|}$ , are semialgebraic of description complexity  $\leq t$  for some  $t$  depending only on  $\varphi$ . Conversely, the family of all semialgebraic subsets of  $\mathbb{R}^{|x|}$  of description complexity  $\leq t$  is of the form  $\{\varphi(\mathbb{R}^{|x|}; b) : b \in \mathbb{R}^{|y|}\}$  for an appropriate choice of  $\varphi(x; y)$ . Hence  $R_k^{d,t}$  from Definition 1.3 is given by  $R_\varphi^*$  for an appropriate  $\varphi$  in the case of the field of reals.

We will restrict to the case of *NIP structures* (see § 2 for the definition; any structure which is not NIP codes arbitrary finite graphs in a definable way (see e.g., [11, Remark 4.12]), hence bounds in Fact 1.2 are optimal outside of the NIP context). First we give a brief overview of the relevant results in the model-theoretic literature indicating the relevance of NIP and its subclasses for the problem at hand.

The infinitary version of the problem of finding indiscernible subsequences was long known in model theory, under the name of the ‘*existence of indiscernibles*’ (starting with the work of Morley in the stable case, and later work of Shelah and others in general NIP [31, 44, 46, 48]).

The question of obtaining explicit bounds for  $R_\varphi(n)$  under some model-theoretic tameness assumptions on  $\mathcal{M}$  was first considered, it appears, in [19], where some quantitative improvements in the stable and NIP cases were obtained. In the case of a stable formula  $\varphi$ , a polynomial upper bound was established in [35].

**Fact 1.10** [35]. *Let  $\varphi(x_1, \dots, x_k; z)$  be a formula in a stable structure  $\mathcal{M}$  (or just assume that  $\varphi$  is a stable formula, relative to an arbitrary partition of its variables). Then there is some  $c = c(\varphi)$  such that  $R_\varphi^*(n) \leq n^c$  for all sufficiently large  $n$ .*

See also [10] for a different proof using the ‘non-standard’ method. Fact 1.4 was generalized to  $o$ -minimal structures (with some additional topological assumptions) in [4], and to *symmetric* relations in arbitrary *distal structures* in the following theorem (see Definition 6.7 for the definition of distality; examples of distal structures include arbitrary  $o$ -minimal structures and  $P$ -minimal structures, e.g., the fields  $\mathbb{Q}_p$  for  $p$  prime – see Definition 5.1).

**Fact 1.11** ([11, Theorem 3.6] + Remark 6.6). *Let  $\mathcal{M}$  be a reduct of a distal structure. Then for any formula  $\varphi(x_1, x_2; z)$  with  $|x_1| = |x_2|$  arbitrary and such that the relation defined by  $\varphi(x_1, x_2; b)$  is symmetric for any  $b \in M^{|z|}$ , there is some  $c = c(\varphi)$  such that  $R_\varphi^*(n) \leq n^c$  for all sufficiently large  $n$ .*

In this paper, we continue investigating the bounds for the functions  $R_\varphi(n)$  and  $R_\varphi^*(n)$  in various NIP structures. First, we consider an analog of the Bukh–Matousek theorem (Fact 1.7) in  $o$ -minimal structures. Recall that a structure  $\mathcal{M} = (M, <, \dots)$  is  *$o$ -minimal* if every definable subset of  $M$  is a finite union of singletons and intervals (with endpoints in  $M \cup \{\pm\infty\}$ ). From this assumption one obtains cell decomposition and other geometric information for definable subsets of  $M^n$ , for all  $n$ . The theory of  $o$ -minimal structures is rather well developed and has applications in other branches of mathematics (we refer to [16] for a detailed treatment of  $o$ -minimality, or to [43, §3] and references therein for a quick introduction). Examples of  $o$ -minimal structures include  $\bar{\mathbb{R}} = (\mathbb{R}, +, \times)$ ,  $\mathbb{R}_{\text{exp}} = (\mathbb{R}, +, \times, e^x)$ ,  $\mathbb{R}_{\text{an}} = (R, +, \times, f \upharpoonright_{[0,1]^k})$  for  $f$  ranging over all functions that are real-analytic on some neighborhood of  $[0,1]^k$ , or  $\mathbb{R}_{\text{an,exp}}$ , the combination of both these last two examples. An  $o$ -minimal structure  $\mathcal{M}$  is *polynomially bounded* if for every definable one-variable function  $f$ , there exists  $N \in \mathbb{N}$  such that  $|f(x)| \leq x^N$  for all sufficiently large positive  $x$ . So for example  $\bar{\mathbb{R}}$  and  $\mathbb{R}_{\text{an}}$  are polynomially bounded, but  $\mathbb{R}_{\text{exp}}$  is not. In §3 we generalize Fact 1.7 to arbitrary polynomially bounded  $o$ -minimal expansions of the field of reals  $\bar{\mathbb{R}}$ .

**Theorem 1.12.** *Let  $\mathcal{M}$  be a polynomially bounded  $o$ -minimal expansion of  $\bar{\mathbb{R}}$ . Then for every  $k \in \mathbb{N}$  and every formula  $\varphi(x_1, \dots, x_k; z)$  with  $x_1, \dots, x_k$  singletons, there is some  $c = c(\varphi)$  such that  $R_\varphi^*(n) \leq 2^{2^{cn}}$  for all sufficiently large  $n$ .*

In particular this implies that in the semialgebraic case (Fact 1.7) the constant  $c$  only depends on the description complexity of the relation, and not on the magnitude of the parameters, which does not seem to have been noticed before. Our argument combines uniform definability of types over finite sets in NIP structures (see Definition 2.4), basic

properties of invariant types and a combinatorial lemma from [6]. On the other hand, in §4 we show that no analog of Theorem 1.12 can hold in  $\mathbb{R}_{\text{exp}}$ .

In this paper, ‘log’ always means logarithm with base 2, unless explicitly stated otherwise.

**Theorem 1.13.** *For every  $k \geq 3$  there are relations  $E_k(x_1, \dots, x_k)$  definable in  $\mathbb{R}_{\text{exp}}$  with  $x_1, \dots, x_k$  singletons, constants  $C_k > 0$  and  $n_k \in \mathbb{N}$  such that, for each  $n > n_k$ , there is a sequence  $\tilde{a}_n$  in  $\mathbb{R}$  of length  $n$  that does not contain an  $E_k$ -indiscernible subsequence of length greater than  $C_k \log \log \dots \log n$ , with  $k - 2$  iterations of log.*

By a theorem of Miller [37], if an  $o$ -minimal expansion of the field of real numbers is not polynomially bounded, then exponentiation is definable in it (i.e., the graph of the exponentiation function is a definable relation). Combining this with Theorems 1.12 and 1.13 we obtain a new combinatorial characterization of polynomial boundedness for  $o$ -minimal expansions of  $\mathbb{R}$ .

**Corollary 1.14.** *Let  $\mathcal{M}$  be an  $o$ -minimal expansion of  $\mathbb{R}$ . The following are equivalent.*

- (1)  $\mathcal{M}$  is polynomially bounded.
- (2) For every  $k \in \mathbb{N}$  and every formula  $\varphi(x_1, \dots, x_k; z)$  with  $x_1, \dots, x_k$  singletons, there is some  $c = c(\varphi)$  such that  $R_\varphi^*(n) \leq 2^{2^{cn}}$  for all sufficiently large  $n$ .
- (3) There is some  $h \in \mathbb{N}$  such that, for every  $k \in \mathbb{N}$  and every formula  $\varphi(x_1, \dots, x_k; z)$  with  $x_1, \dots, x_k$  singletons, there is some  $c = c(\varphi)$  such that  $R_\varphi^*(n) \leq \text{twr}_h(n^c)$  for all sufficiently large  $n$ .

Using the general method of the proof developed in §3, in §5 we apply it to prove an analog of Fact 1.7 in the fields of the  $p$ -adics  $\mathbb{Q}_p$ , for  $p$  prime, and many related structures (see §5 for the definition of  $P$ -minimality and related notions).

**Theorem 1.15.** *Let  $\mathcal{M}$  be a  $P$ -minimal expansion of a field with definable Skolem functions and the value group  $\mathbb{Z}$ . Then, for every  $k \in \mathbb{N}$  and every formula  $\varphi(x_1, \dots, x_k; z)$ , with  $x_1, \dots, x_k$  singletons, there is some  $c = c(\varphi)$  such that  $R_\varphi^*(n) \leq 2^{2^{cn}}$  for all sufficiently large  $n$ .*

This applies to the fields  $\mathbb{Q}_p$  for all primes  $p$ , their finite extensions, as well as expansions by the analytic structure – see §5 for the details. In fact, there are no known examples of  $P$ -minimal structures with value group  $\mathbb{Z}$  that do not satisfy Theorem 1.15 (note that the combinatorial conclusion obviously transfers to the reducts).

**Problem 1.16.** Do Theorems 1.12 and 1.15 hold in polynomially bounded  $o$ -minimal (respectively,  $P$ -minimal) theories that do not admit any archimedean models?

In §6 we consider the growth of  $R_\varphi^*(n)$  in NIP structures for definable relations of higher arity. Generalizing Fact 1.5, we show a definable stepping down lemma for NIP structures which implies the following.

**Theorem 1.17.** *Let  $\mathcal{M}$  be an NIP structure, and assume that for all formulas  $\varphi(x_1, x_2; z)$  we have  $R_\varphi^*(n) \leq n^c$  for some  $c = c(\varphi)$  and all  $n$  large enough. Then for all  $k \geq 3$  and all  $\varphi(x_1, \dots, x_k; z)$  we have  $R_\varphi^*(n) \leq \text{twr}_{k-1}(n^c)$  for some  $c = c(\varphi)$  and all  $n$  large enough.*

In Proposition 6.9 we generalize Fact 1.11 from symmetric binary formulas to arbitrary binary formulas, demonstrating that the assumption of Theorem 1.17 is satisfied in all reducts of distal structures (and it is satisfied in stable structures by Fact 1.10). We conjecture that it also holds in arbitrary NIP structures and discuss the connection to the Erdős–Hajnal conjecture (see e.g., [12]) for graphs definable in NIP structures.

## 2. Preliminaries on NIP

Vapnik–Chervonenkis dimension, or VC-dimension, is an important notion in combinatorics and statistical learning theory (see e.g., [36] for an exposition). Let  $X$  be a set, finite or infinite, and let  $\mathcal{F}$  be a family of subsets of  $X$ . Given  $A \subseteq X$ , we say that it is *shattered* by  $\mathcal{F}$  if for every  $A' \subseteq A$  there is some  $S \in \mathcal{F}$  such that  $A \cap S = A'$ . A family  $\mathcal{F}$  is a *VC-class* if there is some  $n < \omega$  such that no subset of  $X$  of size  $n$  is shattered by  $\mathcal{F}$ . In this case *the VC-dimension of  $\mathcal{F}$* , that we will denote by  $VC(\mathcal{F})$ , is the smallest integer  $n$  such that no subset of  $X$  of size  $n + 1$  is shattered by  $\mathcal{F}$ . For a set  $B \subseteq X$ , let  $\mathcal{F} \cap B = \{A \cap B : A \in \mathcal{F}\}$  and let  $\pi_{\mathcal{F}}(n) = \max\{|\mathcal{F} \cap B| : B \subseteq X, |B| = n\}$ .

**Fact 2.1** (Sauer–Shelah lemma [42, 45]). *If  $VC(\mathcal{F}) \leq d$  then for  $n \geq d$  we have  $\pi_{\mathcal{F}}(n) \leq \sum_{i \leq d} \binom{n}{i} = O(n^d)$ .*

The important class of NIP theories was introduced by Shelah in his work on the classification program [44]. It has attracted a lot of attention recently, both from the point of view of pure model theory and as a result of its applications in algebra and geometry (see e.g., [1, 50] for an introduction to the area). Examples of NIP structures are given by arbitrary stable structures, (weakly or quasi)  $\mathcal{O}$ -minimal structures, the field of  $p$ -adics for every prime  $p$  (along with its analytic expansion), as well as algebraically closed valued fields. As was observed in [33], the original definition of NIP is equivalent to the following one (see [3] for a more detailed account).

**Definition 2.2.** Let  $T$  be a complete theory and  $\varphi(x, y)$  a formula in  $T$ , where  $x, y$  are tuples of variables, possibly of different length. We say that *the formula  $\varphi(x, y)$  is NIP* if there is a model  $\mathcal{M}$  of  $T$  such that the family of definable sets  $\{\varphi(\mathcal{M}, a) : a \in M^{|y|}\}$  is a VC-class. In this case we define *the VC-dimension of  $\varphi(x, y)$*  to be the VC-dimension of this class. (It is easy to see that by elementary equivalence the above does not depend on the model  $\mathcal{M}$  of  $T$ .) A *theory  $T$  is NIP* if all formulas in  $T$  are NIP, and a structure  $\mathcal{M}$  is NIP if its complete theory  $Th(\mathcal{M})$  is NIP. That is, a structure  $\mathcal{M}$  is NIP if for every formula  $\varphi(x, y)$  the family of  $\varphi$ -definable sets  $\mathcal{F}_\varphi = \{\varphi(\mathcal{M}, a) : a \in M^{|y|}\}$  is a VC-class.

By a *partitioned* set of formulas  $\Delta(x, y)$ , where  $x$  and  $y$  are two groups of variables, we mean a set of formulas all of which are of the form  $\varphi(x, y) \in \mathcal{L}$ , i.e., have the same free variables partitioned into the same two groups. Given a (partitioned) set of formulas  $\Delta(x, y)$  and a set  $B \subseteq M^{|y|}$ , we say that  $\pi(x)$  is a  $\Delta$ -*type over  $B$*  if

$\pi(x) \subseteq \bigcup_{\varphi(x,y) \in \Delta, b \in B} \{\varphi(x,b), \neg\varphi(x,b)\}$  and there is some  $\mathcal{N} \succeq \mathcal{M}$  and some  $a \in N^{|x|}$  simultaneously satisfying all formulas from  $\pi(x)$ . By a *complete*  $\Delta$ -type over  $B$  we mean a maximal  $\Delta$ -type over  $B$ . We will denote by  $S_\Delta(B)$  the collection of all complete  $\Delta$ -types over  $B$ . If  $\Delta$  consists of a single formula  $\varphi(x, y)$ , we simply say  $\varphi$ -type and write  $S_\varphi(B)$ , and if  $\Delta$  consists of all formulas in the language, then we simply say ‘type’ and write  $S_x(B)$  for the space of complete types over  $B$ . In view of the remarks above, the following is an immediate corollary of the Sauer–Shelah lemma.

**Fact 2.3.** *A structure  $\mathcal{M}$  is NIP if and only if for any finite set of formulas  $\Delta(x, y)$  there is some  $d \in \mathbb{N}$  such that  $|S_\Delta(B)| = O(|B|^d)$  for any finite  $B \subseteq M^{|y|}$ .*

This result can be strengthened. The following definition is from [3, 27].

**Definition 2.4.** (1) Given a complete  $\varphi(x, y)$ -type  $q \in S_\varphi(B)$  for a set  $B \subseteq M^{|y|}$ , an  $L(M)$ -formula  $d\varphi(y)$  is said to *define*  $q$  if for all  $b \in B$  we have

$$\varphi(x, b) \in q \iff \mathcal{M} \models d\varphi(b).$$

- (2) We say that complete  $\varphi(x, y)$ -types are *uniformly definable over finite sets*, with  $m$  parameters, if there is a finite set of  $\mathcal{L}$ -formulas  $\Delta = (d\varphi_i(y; y_1, \dots, y_m) : i < k)$ , with  $|y_1| = |y|$  for all  $i < k$ , such that for every finite set  $B \subseteq M^{|y|}$  and every  $q \in S_\varphi(B)$  there are some  $b_1, \dots, b_m \in B$  and some  $i < k$  such that  $d\varphi_i(y; b_1, \dots, b_m)$  defines  $q$ . We call the set  $\Delta$  a *uniform definition* for  $\varphi$ -types over finite sets, with  $m$  parameters.
- (3) We say that  $T$  satisfies the *Uniform Definability of Types over Finite Sets*, or *UDTFS*, if for some (equivalently, any)  $\mathcal{M} \models T$ , complete  $\varphi$ -types are uniformly definable over finite sets for all formulas  $\varphi \in L$ .

**Fact 2.5** [8]. *Every NIP theory satisfies UDTFS.*

This result can be viewed as a model-theoretic version of the Warmuth conjecture on the existence of compression schemes for VC-families, which was later established in [39]. Special cases of Fact 2.5 were proved earlier for some subclasses of NIP theories including stable [44],  $o$ -minimal [30], and dp-minimal [27] theories. Note that this implies Fact 2.3 since, under UDTFS, for every finite set of formulas  $\Delta$ , every  $\Delta$ -type over a finite set  $B$  is determined by fixing a definition for each  $\varphi \in \Delta$  with parameters from  $B$ , of which there are only polynomially many choices. Explicit bounds on the number of parameters needed are given in [3] for some cases considered in this article.

**Fact 2.6.** (1) [3, § 6.1] *Let  $\mathcal{M}$  be a (weakly or quasi)  $o$ -minimal structure. Then  $\varphi(x, y)$ -types are uniformly definable over finite sets using  $|x|$  parameters, for all formulas  $\varphi \in L$ . In particular this applies to Presburger arithmetic  $(\mathbb{Z}, +, <)$ .*

- (2) [3, § 7.2] *Let  $\mathcal{M}$  be the field of  $p$ -adics. Then  $\varphi(x, y)$ -types are uniformly definable over finite sets using  $2|x|$  parameters, for all formulas  $\varphi \in L$ .*

Finally, we recall global invariant types and their products. We will use some standard model-theoretic notation, e.g.,  $\mathbb{M} \succ \mathcal{M}$  will be a saturated elementary extension, and,



given a set  $A \subseteq \mathbb{M}$ ,  $\text{dcl}(A)$  will denote the model-theoretic algebraic closure of  $A$ ,  $A$  will be called *small* if its cardinality is smaller than the saturation of  $\mathbb{M}$ , etc. Given a tuple of variables  $x$ , we call complete types in  $S_x(\mathbb{M})$  *global*, and we say that a global type  $p(x)$  is *M-invariant* if it is  $\text{Aut}(\mathbb{M}/M)$ -invariant (meaning that, for every automorphism  $\sigma$  of  $\mathbb{M}$  fixing  $M$  pointwise, for every  $L(\mathbb{M})$ -formula  $\varphi(x, a)$ , we have  $\varphi(x, a) \in p \iff \varphi(x, \sigma(a)) \in p$ ).

**Definition 2.7.** Given a set of formulas  $\Delta$ ,  $d \in \mathbb{N}$ , a set of parameters  $A \subseteq \mathbb{M}$  and an arbitrary linear order  $I$ , we say that a sequence  $(a_i : i \in I)$  of tuples from  $\mathbb{M}^d$  is  $\Delta$ -*indiscernible over  $A$*  if it is  $E$ -indiscernible for every relation  $E$  of the form  $\varphi(x_1, \dots, x_n; b)$  with  $\varphi(x_1, \dots, x_n; z) \in \Delta$ ,  $|x_i| = d$  for all  $1 \leq i \leq n$  and  $b \in A^{|z|}$ .

If  $\Delta$  consists of all formulas, we simply say that the sequence is *indiscernible over  $A$* , and if  $A = \emptyset$ , we say that the sequence is  $\Delta$ -indiscernible.

**Fact 2.8** (See e.g., [29, § 2] or [50]). *Let  $p$  be a global  $M$ -invariant type. Let the sequence  $(c_i : i \in \mathbb{N})$  in  $\mathbb{M}$  be such that  $c_i \models p|_{M c_{<i}}$  (such a sequence is called a Morley sequence in  $p$  over  $M$ ). Then the sequence  $(c_i : i \in \mathbb{N})$  is indiscernible over  $M$  and  $\text{tp}((c_i : i \in \mathbb{N})/M)$  does not depend on the choice of  $(c_i)$ . Call this type  $p^{(\omega)}|_M$ , and let  $p^{(n)}|_M := \text{tp}(c_1, \dots, c_n/M)$ .*

### 3. Bukh–Matousek theorem in polynomially bounded $o$ -minimal expansions of $\mathbb{R}$

First we prove a general lemma about NIP structures, which is a finitary version of Shelah’s ‘shrinking of indiscernibles’ [47].

**Lemma 3.1.** *Let  $\mathcal{M}$  be an NIP structure, and let  $\varphi(x_1, \dots, x_n; y)$  be a formula with  $|x_1| = \dots = |x_n| = d$ . Then there are some  $k, l \in \mathbb{N}$  and a finite set of formulas  $\Delta$  in the variables  $x_1, \dots, x_l$  with  $|x_i| = d$  such that for any finite  $\Delta$ -indiscernible sequence  $(a_i)_{i < N}$  in  $M^d$  and any  $b \in M^{|y|}$  there are  $0 = j_0 < j_1 < \dots < j_{k'} = N - 1$  with  $k' \leq k$  such that for every  $s \in \{0, \dots, k' - 1\}$  the sequence  $(a_i : j_s < i < j_{s+1})$  is  $\varphi(x_1, \dots, x_n, b)$ -indiscernible.*

*In particular, for any  $N$  large enough and any  $b \in M^{|y|}$ , any finite  $\Delta$ -indiscernible sequence of elements in  $M^d$  of length  $N$  contains a  $\varphi(x_1, \dots, x_n, b)$ -indiscernible subsequence of length at least  $\frac{N - (k+1)}{k}$ .*

**Proof.** To simplify the notation we assume  $d = 1$ .

By UDIFS (Fact 2.5) applied to the formula  $\varphi^{\text{op}}(y; x_1, \dots, x_n) := \varphi(x_1, \dots, x_n; y)$ , there is a finite set of formulas  $\Delta(x_1, \dots, x_n; \bar{x}_1, \dots, \bar{x}_m)$  with  $|\bar{x}_i| = n$  such that, for any finite set  $A \subseteq M$  and  $b \in M^{|y|}$ , the  $\varphi^{\text{op}}$ -type of  $b$  over  $A^n$  is definable by an instance of some  $\psi \in \Delta$  with parameters from  $A^n$ . That is, there are some  $\bar{c}_1, \dots, \bar{c}_m \in A^n$ , such that, for all  $a_1, \dots, a_n \in A$ , we have  $\models \varphi(a_1, \dots, a_n; b)$  if and only if  $\models \psi(a_1, \dots, a_n; \bar{c}_1, \dots, \bar{c}_m)$ .

Writing each  $n$ -tuple  $\bar{x}_i$ ,  $i = 1, \dots, m$ , as  $n$  single variables in every  $\psi \in \Delta$ , we can view  $\Delta$  as a finite set of formulas in the variables  $x_1, \dots, x_l$ , where  $l = n + mn$ .



Let  $(a_i)_{i < N}$  be a finite  $\Delta$ -indiscernible sequence,  $b \in M^{|\mathcal{V}|}$ , and  $A = \{a_i : i < N\}$ . We choose  $\psi \in \Delta$  and  $c_{n+1}, \dots, c_l \in A$  such that for all  $c_1, \dots, c_n \in A$  we have  $\mathcal{M} \models \varphi(c_1, \dots, c_n; b)$  if and only if  $\mathcal{M} \models \psi(c_1, \dots, c_n, c_{n+1}, \dots, c_l)$ .

We choose  $0 = j_0 < j_1 < \dots < j_{k'} = N - 1$  with  $k' \leq (l - n) + 2 = mn + 2$  so that  $\{a_{j_s} : s = 0, \dots, k'\} = \{c_i : i = n + 1, \dots, l\} \cup \{a_0, a_{N-1}\}$ .

Since  $(a_i)_{i < N}$  is  $\psi$ -indiscernible, it follows that for any  $0 \leq i_1 < \dots < i_n < N$  the truth value of  $\psi(a_{i_1}, \dots, a_{i_n}; c_{n+1}, \dots, c_l)$ , and so of  $\varphi(a_{i_1}, \dots, a_{i_n}; b)$ , is determined by the quantifier-free order type of  $(i_1, \dots, i_n)$  over  $\{j_s : s = 0, \dots, k'\}$ . The conclusion of the lemma follows taking  $k := mn + 2$ .  $\square$

From now on we work in a polynomially bounded  $\mathcal{o}$ -minimal expansion  $\mathcal{R} = \langle \mathbb{R}, <, \dots \rangle$  of the field of real numbers. Let  $T = Th(\mathcal{R})$  and let  $\mathbb{M} \succ \mathcal{R}$  be a big saturated model.

As  $T$  has Skolem functions (see e.g., [16]), it follows that for all  $\mathcal{M} \prec \mathbb{M}$  and  $\bar{a} \in \mathbb{M}^n$ , the set

$$M(\bar{a}) = \{f(\bar{a}) : f(x) \text{ is an } M\text{-definable function}\}$$

is an elementary substructure of  $\mathbb{M}$ .

Let  $\tilde{p}(x) \in S_1(\mathbb{M})$  be the global type of ‘ $+\infty$ ’, i.e.,  $\tilde{p}$  is the unique complete global type such that  $\tilde{p} \vdash x > m$  for every  $m \in \mathbb{M}$  (uniqueness is by  $\mathcal{o}$ -minimality). It is invariant over  $\emptyset$  (as the set of formulas  $\{m < x : m \in \mathbb{M}\}$  is clearly  $\text{Aut}(\mathbb{M}/\emptyset)$ -invariant).

The following fact is obvious.

**Fact 3.2.** *For every  $\mathcal{M} \prec \mathbb{M}$ , an element  $\alpha \in \mathbb{M}$  realizes  $\tilde{p}(x)|_{\mathcal{M}}$  if and only if  $\alpha > m$  for every  $m \in M$ .*

Since polynomial boundedness is preserved under elementarily equivalence (see [38, Theorems A and B]) we have the following fact.

**Fact 3.3.** *If  $\mathcal{M} \prec \mathbb{M}$  and  $\alpha \models \tilde{p}|_{\mathcal{M}}$ , then the set  $\{\alpha^n : n \in \mathbb{N}\}$  is cofinal in  $M\langle\alpha\rangle$ , i.e., for every  $m \in M\langle\alpha\rangle$  there is some  $n \in \mathbb{N}$  such that  $m < \alpha^n$ .*

**Lemma 3.4.** *Let  $\mathcal{M} \prec \mathbb{M}$  and  $\alpha_1, \dots, \alpha_n \in \mathbb{M}$ . Then  $(\alpha_1, \dots, \alpha_n)$  realizes  $\tilde{p}^{(n)}|_{\mathcal{M}}$  if and only if  $\alpha_1 > m$  for all  $m \in M$  and  $\alpha_{i+1} > \alpha_i^k$  for all  $k \in \mathbb{N}$  and  $i = 1, \dots, n - 1$ .*

**Proof.** Let  $M_0 = M$ , and for  $i = 1, \dots, n - 1$  let  $M_i = M_{i-1}\langle\alpha_i\rangle$ .

Obviously for any  $A \subset \mathbb{M}$  an element  $\alpha \in \mathbb{M}$  realizes  $\tilde{p}|_A$  if and only if it realizes  $\tilde{p}|_{\text{dcl}(A)}$ . Thus  $(\alpha_1, \dots, \alpha_n)$  realizes  $\tilde{p}^{(n)}|_{\mathcal{M}}$  if and only if  $\alpha_{i+1}$  realizes  $\tilde{p}|_{M_i}$  for  $i = 0, \dots, n - 1$ , and the lemma follows from Facts 3.2 and 3.3.  $\square$

In view of the above lemma, we define ‘finitary’ approximations to a realization of  $\tilde{p}^{(n)}|_{\mathcal{R}}$ .

**Definition 3.5** [6, Definition 2.1]. Let  $h > 2$  be a real number. A sequence  $\vec{a} = (a_1, \dots, a_n)$  in  $\mathcal{R}$  is called  *$h$ -growing* if  $a_1 \geq h$  and  $a_{i+1} \geq a_i^h$  for  $i = 1, \dots, n - 1$ .

Notice that any subsequence of an  $h$ -growing sequence is  $h$ -growing as well.

**Lemma 3.6.** *For any finite set of formulas  $\Delta(x_1, \dots, x_l)$  with parameters from  $\mathbb{R}$  there is some  $h \in \mathbb{R}$  such that any  $h$ -growing sequence  $(a_i : i = 1, \dots, N)$  of elements in  $\mathbb{R}$  is  $\Delta$ -indiscernible.*

**Proof.** Consider the (partial) type

$$\Sigma(x_1, \dots, x_{2l}) = \left\{ x_1 > n \wedge \bigwedge_{i=1}^{2l-1} (x_{i+1} > x_i^n) : n \in \mathbb{N} \right\}.$$

By Fact 2.8, for any  $N \in \mathbb{N}$ , if  $(a_1, \dots, a_N) \models \tilde{p}^{(N)}|_M$ , then the sequence  $(a_1, \dots, a_N)$  is indiscernible. Together with Lemma 3.4 this implies that

$$\Sigma(x_1, \dots, x_{2l}) \vdash \psi(x_1, \dots, x_l) \leftrightarrow \psi(x_{i_1}, \dots, x_{i_l})$$

for any  $1 \leq i_1 < i_2 < \dots < i_l \leq 2l$  and  $\psi \in \Delta$ . By compactness, this holds with  $\Sigma$  replaced by some finite subset  $\Sigma_0$ . But then, if  $a_1, \dots, a_N$  is an  $h$ -growing sequence and  $h$  is larger than the largest  $n$  appearing in  $\Sigma_0$ , then every increasing  $2l$ -tuple from  $a_1, \dots, a_N$  satisfies  $\Sigma_0$ , hence  $a_1, \dots, a_N$  is  $\Delta(x_1, \dots, x_l)$ -indiscernible.  $\square$

Combining Lemma 3.6 with Lemma 3.1 we can allow additional parameters in  $\Delta$ .

**Corollary 3.7.** *For any finite set of formulas  $\Delta(x_1, \dots, x_l; y)$  with parameters from  $\mathbb{R}$  there is some  $h \in \mathbb{R}$  and  $m \in \mathbb{N}$  such that, for any  $h$ -growing sequence of elements  $\vec{a} = (a_i : i = 1, \dots, N)$  in  $\mathbb{R}$  with  $N$  large enough and for any  $b \in \mathbb{R}^{|y|}$ ,  $\vec{a}$  contains a  $\Delta(x_1, \dots, x_l; b)$ -indiscernible subsequence of length  $\frac{N}{m}$ .*

**Proof.** For every  $\varphi(x_1, \dots, x_l; y) \in \Delta$ , let  $k_\varphi \in \mathbb{N}$  and the finite set of formulas  $\Delta_\varphi$  be as given by Lemma 3.1 for  $\varphi$ , and let  $\Delta' = \bigcup_{\varphi \in \Delta} \Delta_\varphi$  and  $k = \max\{k_\varphi : \varphi \in \Delta\}$ . Now by Lemma 3.6 there is some  $h$  such that every  $h$ -growing sequence  $\vec{a} = (a_1, \dots, a_N)$  of elements from  $\mathbb{R}$  is  $\Delta'$ -indiscernible. By Lemma 3.1, for any  $b \in \mathbb{R}^{|y|}$  we can find an interval  $[i^0, i^1]$  in  $[1, N]$  of length at least  $\frac{N - (k|\Delta| - 2)}{k|\Delta|}$  such that the sequence  $(a_i : i^0 \leq i \leq i^1)$  is  $\Delta(x_1, \dots, x_l; c)$ -indiscernible. We can take  $m = 2k|\Delta|$ .  $\square$

Finally, the following combinatorial lemma is from [6] (namely, Proposition 2.4 combined with Definition 2.3 there).

**Fact 3.8.** *For every  $n$  and  $h \geq h_0$ , where  $h_0$  is a certain absolute constant, there exists  $N \leq 2^{h^{2n}}$  such that for any sequence  $\vec{a}$  of length  $N$  there is an  $h$ -growing sequence  $\vec{b}$  of length  $n$  and  $A, B \in \mathbb{R}$  such that one of the following sequences is a subsequence of  $\vec{a}$ .*

- (1)  $A + Bb_i, i = 1, \dots, n.$
- (2)  $A + \frac{B}{b_i}, i = 1, \dots, n.$
- (3)  $A + Bb_i, i = n, \dots, 1.$
- (4)  $A + \frac{B}{b_i}, i = n, \dots, 1.$

(Note: the order in (3) and (4) is reversed.)

We are ready to prove the main result of the section, generalizing [6, Proposition 1.6].

**Theorem 3.9.** *Let  $\mathcal{R}$  be a polynomially bounded  $o$ -minimal expansion of the real field. Then for any formula  $\varphi(x_1, \dots, x_r; z)$  with parameters from  $\mathbb{R}$ , with all  $x_i$  singletons, there is a constant  $C = C(\varphi)$  such that*

$$R_\varphi^*(n) \leq 2^{2^{Cn}},$$

for all sufficiently large  $n$ .

**Proof.** Let  $\Delta(x_1, \dots, x_r; y_1, y_2, z)$  consist of the formulas

$$\begin{aligned} \varphi_1(x_1, \dots, x_r; y_1, y_2, z) &= \varphi(y_1 + y_2 x_1, \dots, y_1 + y_2 x_r; z), \\ \varphi_2(x_1, \dots, x_r; y_1, y_2, z) &= \varphi\left(y_1 + \frac{y_2}{x_1}, \dots, y_1 + \frac{y_2}{x_r}; z\right), \\ \varphi_3(x_1, \dots, x_r; y_1, y_2, z) &= \varphi(y_1 + y_2 x_r, \dots, y_1 + y_2 x_1; z), \\ \varphi_4(x_1, \dots, x_r; y_1, y_2, z) &= \varphi\left(y_1 + \frac{y_2}{x_r}, \dots, y_1 + \frac{y_2}{x_1}; z\right), \end{aligned}$$

and let  $h$  and  $m$  be as given by Corollary 3.7 for  $\Delta$ . Now assume that  $\vec{a}$  is an arbitrary sequence of singletons of length  $N = 2^{h^{2mn}}$  (which is bounded by  $2^{2^{Cn}}$  for an appropriate constant  $C$  depending just on  $m, h$ ), and let  $d \in \mathbb{R}^{|z|}$  be an arbitrary tuple of elements.

By Fact 3.8, there is some  $h$ -growing sequence  $\vec{b} = (b_i : 1 \leq i \leq mn)$  and some  $A, B \in \mathbb{R}$  such that one of the corresponding sequences given by (1)–(4) in Fact 3.8 is a subsequence of  $\vec{a}$ . By Corollary 3.7,  $\vec{b}$  contains a  $\Delta(x_1, \dots, x_r; A, B, d)$ -indiscernible subsequence of length  $n$ . But by the choice of  $\Delta$ , the corresponding subsequence of  $\vec{a}$  must be  $\varphi(x_1, \dots, x_r; d)$ -indiscernible.  $\square$

## 4. Counterexample in $\mathbb{R}_{\exp}$

### 4.1. Preliminaries

We work in the structure  $\mathcal{M} := \mathbb{R}_{\exp}$  in the language  $\mathcal{L} := (<, +, \times, 0, 1, \exp(x))$ , i.e, the expansion of the field of reals with the exponential function. It is well known to be  $o$ -minimal [51].

Instead of tower notations we use iterated  $\log$  and  $\exp$ . By induction on  $n$  we define functions  $e_n(x)$  and  $l_n(x)$  as

$$e_0(x) = x, e_{n+1}(x) = 2^{e_n(x)}; \text{ and } l_0(x) = x, l_{n+1}(x) = \log(l_n(x)),$$

where by  $\log$  we always mean  $\log_2$ . Obviously  $l_n(x)$  is defined for large enough  $x$  and it is the compositional inverse of  $e_n(x)$ .

Our goal is to prove the following theorem.

**Theorem 4.1.** *For every  $k \geq 3$  there is a relation  $E_k(x_1, \dots, x_k)$  definable in  $\mathbb{R}_{\exp}$ , with  $x_1, \dots, x_k$  all singletons, and  $c_k > 0$  such that  $R_{E_k}(n) \geq e_{k-2}(c_k n)$  for all sufficiently large  $n$ .*

The proof of the above theorem closely follows [13, proof of Theorem 1.2] (see also [18, Theorem 1.3]). In general, the so-called stepping-up lemma of Erdős and Hajnal [14, 26]

gives a lower bound for  $(k+1)$ -ary relations which is exponentially larger than the one for  $k$ -ary relations. In [13] it is demonstrated that the stepping-up lemma can be carried out ‘semialgebraically’ (the  $k$ -ary semialgebraic relations that they construct live on  $\mathbb{R}^d$ , and  $d$  grows with  $k$ , see Fact 1.6). We show that in the structure  $\mathbb{R}_{\text{exp}}$  the stepping-up approach can be implemented definably without increasing the dimension (i.e., our  $k$ -ary relations all live on  $\mathbb{R}$ ). But first we discuss some preliminaries.

#### 4.2. Robustness

We will use the notion of robustness from [18] (that was originally called ‘depth’ in [13]).

**Definition 4.2.** Let  $\varphi(x_1, \dots, x_k)$  be an  $\mathcal{L}$ -formula and let  $\vec{a} = (a_1, \dots, a_n)$  be a sequence of real numbers. We say that  $\varphi$  is *robust* on  $\vec{a}$  if there is  $\varepsilon > 0$  such that, for all  $1 \leq i_1 < \dots < i_k \leq n$  and all real numbers  $a'_1, \dots, a'_k$  with  $|a_{i_j} - a'_j| < \varepsilon$  for each  $j = 1, \dots, k$ , we have

$$\models \varphi(a'_1, \dots, a'_k) \leftrightarrow \varphi(a_{i_1}, \dots, a_{i_k}).$$

#### 4.3. $\log_T$ -transformations

**Definition 4.3.** Let  $\varphi(x_1, \dots, x_r)$  be an  $\mathcal{L}$ -formula. Let  $T > 0$  be a real number. For a formula  $\psi(y_1, \dots, y_s)$  we say that  $\psi$  is a  $\log_T$ -transformation of  $\varphi$  if it is obtained from  $\varphi$  by replacing **every** free variable  $x_i$  in  $\varphi$  by an expression of the form  $\log_T(u_i - v_i)$  with  $u_i, v_i \in \{y_1, \dots, y_s\}$ .

**Definition 4.4.** We say that an  $\mathcal{L}$ -formula  $\varphi(x_1, \dots, x_r)$  is an *rd-formula* if it depends only on the ratios of differences of its variables, i.e., it is equivalent to a formula of the form

$$\psi\left(\frac{x_{i_1} - x_{j_1}}{x_{p_1} - x_{q_1}}, \dots, \frac{x_{i_s} - x_{j_s}}{x_{p_s} - x_{q_s}}\right)$$

for some  $\psi(y_1, \dots, y_s) \in \mathcal{L}$ , where  $i_t, j_t, p_t, q_t \in \{1, \dots, r\}$  for all  $t = 1, \dots, s$  (and there are no other free variables in  $\psi$ ).

**Claim 4.5.** Let  $T > 0$ . A  $\log_T$ -transformation of an rd-formula  $\varphi(x_1, \dots, x_r)$ , is also an rd-formula, and it is also a  $\log_2$ -transformation of  $\varphi$ .

**Proof.** In a  $\log_T$ -transformation of  $\varphi$  an expression of the form  $\frac{x_i - x_j}{x_p - x_q}$  is replaced by an expression of the form

$$\frac{\log_T(u_i - v_i) - \log_T(u_j - v_j)}{\log_T(u_p - v_p) - \log_T(u_q - v_q)},$$

which is equivalent to

$$\frac{\log_T \frac{u_i - v_i}{u_j - v_j}}{\log_T \frac{u_p - v_p}{u_q - v_q}}.$$

Since the ratio of two logarithms does not depend on the base, it is also equivalent to

$$\frac{\log \frac{u_i - v_i}{u_j - v_j}}{\log \frac{u_p - v_p}{u_q - v_q}}.$$

Thus, a  $\log_7$ -transformation of an rd-formula  $\varphi$  is again an rd-formula that is also a  $\log_2$ -transformation of  $\varphi$ .  $\square$

#### 4.4. Proof of Theorem 4.1

For a formula  $\varphi(x_1, \dots, x_k) \in \mathcal{L}$  with  $|x_1| = \dots = |x_k| = d$  and an integer  $n$  we will denote by  $R_\varphi^+(n)$  the smallest integer  $N$  such that any *increasing* sequence  $a_1 < \dots < a_N$  contains a  $\varphi$ -indiscernible subsequence of length  $n$ .

Obviously for any formula  $\varphi(x_1, \dots, x_k)$  with  $|x_i| = 1$  we have  $R_\varphi^+(n) \leq R_\varphi(n)$ .

Thus Theorem 4.1 follows from the following refined version.

**Theorem 4.6.** *For every  $k \geq 3$  there are an rd-formula  $E_k(x_1, \dots, x_k) \in \mathcal{L}$  with  $x_1, \dots, x_k$  all singletons and a constant  $C_k > 1$  such that, for all real  $0 < c < 1$  and for all large enough  $n \in \mathbb{N}$ , there is an increasing sequence of natural numbers  $\vec{a}^n$  of length at least  $e_{k-2}(cn)$  such that  $E_k$  is robust on  $\vec{a}^n$ , and  $\vec{a}^n$  does not contain an  $E_k$ -indiscernible subsequence of length  $C_k n$ .*

**Proof of Theorem 4.6  $\Rightarrow$  Theorem 4.1.** Fix  $0 < c < 1$  and set  $c_k = \frac{c}{C_k}$ , for each  $k \geq 3$ , where  $C_k$  is the constant given by 4.6. We then have, for the rd-formula  $E_k(x_1, \dots, x_k)$  given to us by 4.6 and for all large enough  $n$ , an increasing sequence of natural numbers  $\vec{a}^n$  of length at least  $e_{k-2}(c_k n)$  such that  $\vec{a}^n$  does not contain an  $E_k$ -indiscernible subsequence of length  $n$ . Thus  $R_\varphi^+(n) \geq e_{k-2}(c_k n)$ , and hence  $R_\varphi(n) \geq e_{k-2}(c_k n)$  by the preceding remark.  $\square$

**Remark 4.7.** To prove Theorem 4.6 it is enough to construct formulas  $E_k(x_1, \dots, x_k)$  whose truth values are well defined only on increasing sequences of real numbers  $r_1 < \dots < r_k$ . (The formula  $\log(x_2 - x_1) > \log(x_3 - x_2)$  is an example of a formula that we will use often.)

We proceed by induction on  $k$ .

#### 4.5. The base case $k = 3$

For the following claim see [13, §3.1].

**Claim 4.8.** Let  $E_3(x_1, x_2, x_3)$  be the formula  $x_1 + x_3 - 2x_2 \geq 0$ . Then for any  $n \geq 1$  the sequence  $1, 2, 3, \dots, 2^n$  does not contain an  $E_3$ -indiscernible subsequence of length  $n + 2$ .

It is not hard to see that  $E_3$  is equivalent to an rd-formula. Indeed we can rewrite  $E_3$  as  $x_3 - x_2 \geq x_2 - x_1$ , which on increasing sequences is equivalent to  $\frac{x_3 - x_2}{x_2 - x_1} \geq 1$ .

We also need  $E_3$  to be robust on  $\vec{a}^n$ . It is not hard to see that  $E_3$  is not robust on the sequence  $1, 2, \dots, 2^n$ , since  $1 + 3 - 2 \cdot 2 = 0$  and the truth of  $E_3$  can change even if

we perturb the first 3 elements of the sequence by arbitrarily small positive amounts. It is however also easy to see that  $E_3$  is robust on any sequence that does not contain any terms  $a < b < c$  with  $a + c - 2b = 0$ , i.e., it is robust on any sequence that does not contain a non-trivial 3-term arithmetic progression. To get such a sequence we use Behrend's Theorem (see [5]).

**Theorem 4.9** (Behrend's Theorem). *There is a constant  $D > 0$  such that for all natural numbers  $m$  there exists a set  $X \subseteq \{1, \dots, m\}$  with  $|X| \geq \frac{m}{2^{D\sqrt{\log m}}}$  not containing any non-trivial 3-term arithmetic progressions.*

For any  $0 < c < 1$  and for all  $n$  large enough,  $2^{n-D\sqrt{\log 2^n}} > 2^{cn}$ . Therefore, for all large enough  $n$ , the sequence  $1, 2, \dots, 2^n$  contains a subsequence of length  $2^{cn}$  that does not contain a non-trivial 3-term arithmetic progression.

This finishes the case  $k = 3$ , and we can take  $C_3 := 2^\eta$ , for any  $\eta > 0$  (as then  $C_3 n \geq n + 2$  for all large enough  $n$ ).

#### 4.6. Inductive step

Assume we have an rd-formula  $E_k(x_1, \dots, x_k)$  as in Theorem 4.6. To complete the inductive step it is enough to construct an rd-formula  $E_{k+1}(x_1, \dots, x_{k+1})$  satisfying the following for any  $N \in \mathbb{N}$ :

Let  $\vec{a}$  be an increasing sequence of natural numbers of length  $N$  such that  $E_k$  is robust on  $\vec{a}$ , and  $\vec{a}$  does not contain an  $E_k$ -indiscernible subsequence of length  $n$ . Then there is an increasing sequence of natural numbers  $\vec{b}$  of length  $2^N$  such that  $E_{k+1}$  is robust on  $\vec{b}$  and  $\vec{b}$  does not contain an  $E_{k+1}$ -indiscernible subsequence of length  $2n + k - 4$ .

(We are then done taking  $C_k := 2^{k-3+\eta}$  for all  $k \geq 3$ , where  $\eta > 0$  was fixed in the base case.)

Let  $\vec{a} = (a_1, \dots, a_N)$  be an increasing sequence of natural numbers such that  $E_k$  is robust on  $\vec{a}$ , and  $\vec{a}$  does not contain an  $E_k$ -indiscernible sequence of length  $n$ .

Let  $T$  be a very large integer, specified later (in terms of  $\vec{a}$ ).

Consider the set

$$B_T = \left\{ \sum_{i=1}^N \beta_i T^{a_i} : \beta_i \in \{0, 1\} \right\}.$$

Since  $T$  is large enough, any  $b \in B_T$  can be written uniquely as  $b = \sum_{i=1}^N b(i) T^{a_i}$  with  $b(i) \in \{0, 1\}$ . Obviously  $B_T$  has size  $2^N$  and we construct the sequence  $\vec{b}_T$  by taking the increasing enumeration of  $B_T$ .

For  $b, c \in B_T$  with  $b \neq c$ , let  $\Delta(b, c) := \max\{i : b(i) \neq c(i)\}$ . It is easy to see that, when  $T$  is large enough, for  $b, c \in B_T$  with  $b \neq c$  and  $i := \Delta(b, c)$  we have  $b < c \Leftrightarrow b(i) < c(i)$ . It follows then that

$$b < c < d \in B_T \Rightarrow \Delta(b, c) \neq \Delta(c, d). \quad (4.1)$$

Finally for  $b \neq c \in B_T$  let  $\delta(b, c) := a_{\Delta(b, c)}$ .

We will now construct the step-up relation  $E_k^\uparrow(x_1, \dots, x_{k+1})$  (not definable in  $\mathbb{R}_{\text{exp}}$ ) on increasing  $(k+1)$ -tuples of elements of  $B_T$  (we do not care how it is defined on the other elements).

Let  $b_1 < b_2 < \dots < b_{k+1}$  be elements of  $B_T$  and for  $i = 1, \dots, k$  let  $\delta_i := \delta(b_{i+1}, b_i)$ . Notice that  $\delta_i$  is an element of  $\vec{a}$ .

We define  $E_k^\uparrow(b_1, \dots, b_{k+1})$  to be true if and only if

$$\begin{aligned} & E_k(\delta_1, \dots, \delta_k) \quad \text{and} \quad \delta_1 < \delta_2 < \dots < \delta_k, \\ & \quad \text{or} \\ & E_k(\delta_k, \dots, \delta_1) \quad \text{and} \quad \delta_1 > \delta_2 > \dots > \delta_k, \\ & \quad \text{or} \\ & \delta_1 < \delta_2 \quad \text{and} \quad \delta_2 > \delta_3. \end{aligned}$$

**Claim 4.10.** The sequence  $\vec{b}_T$  does not contain an  $E_k^\uparrow$ -indiscernible subsequence of length  $2n + k - 4$ .

**Proof.** We repeat the Erdős–Hajnal argument (see [13, Lemma 3.1]).

Assume, toward getting a contradiction, that  $\vec{b}_T$  contains an  $E_k^\uparrow$ -indiscernible subsequence  $c_1 < c_2 < \dots < c_{2n+k-4}$ . Let  $\delta'_i := \delta(c_{i+1}, c_i)$ .

Assume first that there exists  $j$  such that  $\delta'_j, \delta'_{j+1}, \dots, \delta'_{j+n-1}$  is a monotone sequence. Then by (4.1) this sequence must be strictly monotone. From the definition of  $E_k^\uparrow$  it follows then that the sequence  $\delta'_j, \delta'_{j+1}, \dots, \delta'_{j+n-1}$  is  $E_k$ -indiscernible – a contradiction.

Thus neither of the sequences  $\delta'_1, \dots, \delta'_n$  or  $\delta'_{n-1}, \dots, \delta'_{2n-2}$  is monotone. Hence each of them contains either a local maximum, i.e.,  $\delta'_{j-1} < \delta'_j > \delta'_{j+1}$ , or a local minimum, i.e.,  $\delta'_{j-1} > \delta'_j < \delta'_{j+1}$ . Since between two local minima there is a local maximum and vice versa, the sequence  $\delta'_1, \dots, \delta'_{2n-2}$  contains both a local maximum and a local minimum. But then, by the definition of  $E_k^\uparrow$ , the sequence  $c_1 < c_2 < \dots < c_{2n+k-4}$  cannot be  $E_k^\uparrow$ -indiscernible. A contradiction.  $\square$

**4.6.1. Definability.** Now, as in [13], for  $b > c$  we define

$$\bar{\delta}_T(b, c) = \log_T(b - c).$$

It is not hard to see that for any fixed  $\varepsilon > 0$ , if  $T$  is large enough, then for all  $b > c \in B_T$  we have  $|\delta(b, c) - \log_T(b - c)| < \varepsilon$ .

Since  $E_k$  is robust on  $\vec{a}$ , choosing a very large integer  $T$  and considering the relation  $E_k^{\uparrow T}(x_1, \dots, x_{k+1})$  obtained from  $E_k^\uparrow$  by replacing  $\delta_i$  by  $\bar{\delta}_T(b_{i+1}, b_i)$  for all  $i$ , we obtain that for  $b_1, \dots, b_{k+1} \in B_T$  with  $b_1 < \dots < b_{k+1}$  we have  $E_k^\uparrow(b_1, \dots, b_{k+1})$  if and only if  $E_k^{\uparrow T}(b_1, \dots, b_{k+1})$ . Hence  $\vec{b}_T$  does not contain an  $E_k^{\uparrow T}$ -indiscernible subsequence of length  $2n + k - 4$ .

Notice that  $E_k^{\uparrow T}$  is definable in  $\mathbb{R}_{\text{exp}}$  and for  $b_1 < b_2 < \dots < b_{k+1}$  we have that  $E_k^{\uparrow T}(b_1, \dots, b_{k+1})$  holds if and only if



$$\begin{array}{ccc}
E_k(\bar{\delta}_T(b_2, b_1), \dots, \bar{\delta}_T(b_{k+1}, b_k)) & \text{and} & \bigwedge_{i=1}^{k-1} \bar{\delta}_T(b_{i+1}, b_i) < \bar{\delta}_T(b_{i+2}, b_{i+1}) \\
& \text{or} & \\
E_k(\bar{\delta}_T(b_{k+1}, b_k), \dots, \bar{\delta}_T(b_2, b_1)) & \text{and} & \bigwedge_{i=1}^{k-1} \bar{\delta}_T(b_{i+1}, b_i) > \bar{\delta}_T(b_{i+2}, b_{i+1}) \\
& \text{or} & \\
\bar{\delta}_T(b_2, b_1) < \bar{\delta}_T(b_3, b_2) & \text{and} & \bar{\delta}_T(b_3, b_2) > \bar{\delta}_T(b_4, b_3).
\end{array}$$

**Claim 4.11.**  $E_k^{\uparrow T}$  is equivalent to an rd-formula and does not depend on  $T$ .

**Proof.** By definition,  $E_k^{\uparrow T}$  is a Boolean combination of  $\log_T$ -transformations of  $E_k$  and formulas of the form  $\log_T(y - x) > \log_T(u - v)$ .

By Claim 4.5, a  $\log_T$ -transformation of an rd-formula is an rd-formula that does not depend on  $T$ , and so we only need to check that  $\log_T(y - x) > \log_T(u - v)$  is equivalent to an rd-formula that does not depend on  $T$ . Indeed,  $\log_T(y - x) - \log_T(u - v) > 0$  is equivalent to  $\frac{y-x}{u-v} > 1$ , which is an rd-formula.  $\square$

Using Claim 4.11, we define  $E_{k+1}$  to be  $E_k^{\uparrow 2}$ . We can write a more explicit definition of  $E_{k+1}$ . It is the disjunction of three formulas  $\varphi_1 \vee \varphi_2 \vee \varphi_3$ , where

$$\begin{aligned}
\varphi_1 & \text{ is } E_k\left(\log(x_2 - x_1), \dots, \log(x_{k+1} - x_k)\right) \wedge \bigwedge_{i=1}^{k-1} \left(\frac{x_{i+1} - x_i}{x_{i+2} - x_{i+1}} < 1\right), \\
\varphi_2 & \text{ is } E_k\left(\log(x_{k+1} - x_k), \dots, \log(x_2 - x_1)\right) \wedge \bigwedge_{i=1}^{k-1} \left(\frac{x_{i+1} - x_i}{x_{i+2} - x_{i+1}} > 1\right),
\end{aligned}$$

and

$$\varphi_3 \text{ is } \frac{x_2 - x_1}{x_3 - x_2} < 1 \wedge \frac{x_3 - x_2}{x_4 - x_3} > 1.$$

It remains to show that for large enough  $T$  the relation  $E_{k+1}$  is robust on  $\vec{b}_T$ .

**4.6.2. Robustness.** It is not hard to see that since  $E_k$  is robust on  $\vec{a}$  and  $\log_T$  is continuous, both  $E_k(\log_T(x_2 - x_1), \dots, \log_T(x_{k+1} - x_k))$  and  $E_k(\log_T(x_{k+1} - x_k), \dots, \log_T(x_2 - x_1))$  are robust on  $\vec{b}_T$ , and we only need to check that all of the formulas  $x_{i+1} - x_i < x_{i+2} - x_{i+1}$  and  $x_{i+1} - x_i > x_{i+2} - x_{i+1}$  are robust on  $\vec{b}_T$ , i.e., for  $b < c < d$  in  $B_T$  we do not have  $c - b = d - c$ . It is easy to check that there are no such  $b, c, d$  in  $B_T$ .

## 5. Bukh–Matousek in expansions of the $p$ -adics

In this section we give an analog of Theorem 3.9 for relations definable in the fields of the  $p$ -adic numbers  $\mathbb{Q}_p$  for  $p$  prime and many of their expansions. We begin by recalling the relevant definitions and facts.

Let  $\mathcal{L}_p$  be the Macintyre language for the  $p$ -adics [34], i.e.,  $\mathcal{L}_p$  consists of

- (a) the language of rings:  $(0, 1, +, -, \cdot, {}^{-1})$ ;
- (b) a unary predicate  $V$ ;
- (c) a unary predicate  $P_n$  for each  $n \in \mathbb{N}$ ;

with the usual interpretations in  $\mathbb{Q}_p$ :  $V(\mathbb{Q}_p) = \mathbb{Z}_p$  and  $P_n(\mathbb{Q}_p) = \{x \in \mathbb{Q}_p : \exists y \, x = y^n\}$ . We will denote by  $T_p$  the complete theory  $Th(\mathbb{Q}_p)$ . Given  $a \in \mathbb{Q}_p$ , we will write  $v(a)$  to denote the  $p$ -adic valuation of  $a$ ; note that the relation  $v(x) < v(y)$  is definable in  $\mathcal{L}_p$ .

By a result of Macintyre (see [34]), the theory  $T_p$  eliminates quantifiers in the language  $\mathcal{L}_p$ . Similarly to the  $o$ -minimal case, there is a notion of minimality for expansions of  $p$ -valued fields. Recall that a  $p$ -valued field  $\mathcal{K}$  is a valued field of characteristic 0 with the residue field of characteristic  $p$ , and such that  $\mathcal{O}/p\mathcal{O}$  has finite dimension as a vector space over  $\mathbb{F}_p$ , where  $\mathcal{O}$  is the valuation ring of  $\mathcal{K}$ .

**Definition 5.1.** [28] Let  $\mathcal{K}$  be a  $p$ -valued field, viewed as a structure in the language  $\mathcal{L}_p$ . An expansion  $\mathcal{M}$  of  $\mathcal{K}$  in a language  $\mathcal{L} \supseteq \mathcal{L}_p$  is *P-minimal* if, in every model of  $Th(\mathcal{M})$ , every definable subset in one variable is quantifier-free definable just using the language  $\mathcal{L}_p$ .

**Example 5.2.** Important examples of  $P$ -minimal structures are given by:

- (1) for each prime  $p$ , the field  $\mathbb{Q}_p$  (by Macintyre's theorem);
- (2) any finite extension of  $\mathbb{Q}_p$  [40];
- (3) given a finite extension of  $\mathbb{Q}_p$ , the expansion obtained by adding a new function symbol for every restricted analytic function [17].

**Fact 5.3** [28]. *Every  $P$ -minimal field  $\mathcal{K}$  is  $p$ -adically closed, i.e., it is henselian and its value group is elementarily equivalent to  $\mathbb{Z}$  as an ordered group.*

*In particular, if the value group is  $\mathbb{Z}$ , then  $\mathcal{K}$  is a finite extension of  $\mathbb{Q}_p$ , hence the residue field is finite.*

In this section we will prove the following.

**Theorem 5.4.** *Let  $\mathcal{M}$  be a  $P$ -minimal expansion of a field, and assume that  $\mathcal{M}$  has definable Skolem functions and the value group of  $\mathcal{M}$  (i.e., the value group of the underlying  $p$ -adically closed field) is  $\mathbb{Z}$ . Then for any formula  $\varphi(x_1, \dots, x_r; z) \in \mathcal{L}(\mathcal{M})$ , with all  $x_i$  singletons, there is a constant  $C = C(\varphi)$  such that*

$$R_\varphi^*(n) \leq 2^{2^{Cn}}$$

for all sufficiently large  $n \in \mathbb{N}$ .

It is well known that all of the structures in Example 5.2 satisfy the assumptions of the theorem.

Given a  $P$ -minimal expansion of a field  $\mathcal{M}$ , we will write  $\Gamma_{\mathcal{M}}$  to denote its value group. It is well known that  $\Gamma_{\mathcal{M}}$  is interpretable in  $\mathcal{M}$ .

**Definition 5.5** [32, Definition 4.4]. A  $P$ -minimal structure  $\mathcal{M}$  is *uniformly polynomially bounded* if, for all definable sets  $X$ ,  $W$  and every definable family of functions  $f : X \times W \rightarrow M$ , there is some  $n \in \mathbb{N}$  and a definable function  $a : W \rightarrow \Gamma_{\mathcal{M}}$  such that for each  $w \in W$  we have  $v(f_w(x)) > nv(x)$  for all  $x \in X$  with  $v(x) < a(w)$ .

The next fact is immediate from [15, Lemma 4.3] (as their ‘Extreme Value Property’ holds in every  $P$ -minimal expansion of a field elementarily equivalent to one with the value group  $\mathbb{Z}$ ; see the discussion in [15, p. 123]) and compactness.

**Fact 5.6.** *Let  $\mathcal{M}$  be a  $P$ -minimal expansion of a field elementarily equivalent to a structure with the value group  $\mathbb{Z}$ . Then  $\mathcal{M}$  is uniformly polynomially bounded.*

From now on, we fix a  $P$ -minimal expansion  $\mathcal{M}_0$  of a field with the value group  $\Gamma_{\mathcal{M}_0} = \mathbb{Z}$  and definable Skolem functions in a language  $\mathcal{L}$ . We will denote by  $T$  the complete theory of  $\mathcal{M}_0$ , and we also fix a large sufficiently saturated and homogeneous model  $\mathbb{M}$  of  $T$ . We are following the same strategy as in §3. First we isolate some sufficiently representative global invariant types (in the  $o$ -minimal case, working with a single type of ‘ $+\infty$ ’ was sufficient).

**Proposition 5.7.** *Let  $\mathcal{M} \prec \mathbb{M}$  be a small model of  $T$  and let  $\alpha_1, \alpha_2 \in \mathbb{M}$  be singletons with  $\alpha_1 \equiv_{\emptyset} \alpha_2$  and  $v(\alpha_l) > v(m)$  for  $l = 1, 2$  and every  $m \in M$ . Then  $\alpha_1 \equiv_M \alpha_2$ .*

**Proof.** Since  $\alpha_1$  and  $\alpha_2$  are singletons, by  $P$ -minimality, we need to show the following:

- (1)  $p(\alpha_1) = 0$  if and only if  $p(\alpha_2) = 0$  for any polynomial  $p(x) \in M[x]$ ;
- (2)  $\models V(p(\alpha_1)/q(\alpha_1))$  if and only if  $\models V(p(\alpha_2)/q(\alpha_2))$ , for any  $p(x), q(x) \in M[x]$ ;
- (3)  $\models P_n(p(\alpha_1)/q(\alpha_1))$  if and only if  $\models P_n(p(\alpha_2)/q(\alpha_2))$ , for any  $n \geq 2$  and  $p(x), q(x) \in M[x]$ .

Now (1) holds since the assumption implies that both  $\alpha_1$  and  $\alpha_2$  are transcendental over  $M$ : if  $p(\alpha_l) = 0$  for some  $p(x) \in M[x]$ , then, as  $\mathcal{M}$  is a model and  $p(x) = 0$  is an  $M$ -definable algebraic set, necessarily  $\alpha_l \in M$ , but  $v(\alpha_l) \not\geq v(\alpha_l)$ , contradicting the assumption. And (2) is equivalent to:

$$v(p(\alpha_1)) \geq v(q(\alpha_1)) \text{ if and only if } v(p(\alpha_2)) \geq v(q(\alpha_2)),$$

for any non-zero  $p(x), q(x) \in M[x]$ . Let  $p(x) = a_0 + a_1x + \dots + a_kx^k$  and  $q(x) = b_0 + b_1x + \dots + b_sx^s$ . Let  $i$  be minimal with  $a_i \neq 0$  and  $j$  be minimal with  $b_j \neq 0$ . Then, for  $l = 1, 2$  we have  $v(p(\alpha_l)) = v(a_i\alpha_l^i) = v(a_i) + iv(\alpha_l)$ , and  $v(q(\alpha_l)) = v(b_j\alpha_l^j) = v(b_j) + jv(\alpha_l)$ . Thus  $v(p(\alpha_l)) \geq v(q(\alpha_l))$  if and only if  $i > j$ , or  $i = j$  and  $v(a_i) \geq v(b_j)$ . The latter condition is independent of  $l$ .

Finally, we demonstrate (3). It is easy to see that  $\models P_n(p(\alpha_l)/q(\alpha_l))$  if and only if  $\models P_n(p(\alpha_l)q^{n-1}(\alpha_l))$ . Thus we need to show that  $\models P_n(p(\alpha_1))$  if and only if  $\models P_n(p(\alpha_2))$ , for any  $p(x) \in M[x]$ .

We will need the following fact that follows easily from henselianity of  $\mathcal{M}$  (which holds by Fact 5.3).

**Fact 5.8.** *If  $\varepsilon \in \mathbb{M}$  satisfies  $v(\varepsilon) > k$  for all  $k \in \mathbb{N}$  then for any  $n \in \mathbb{N}$  the element  $1 + \varepsilon$  has  $n$ -th root.*

Let  $p(x) = a_0 + a_1x + \dots + a_kx^k$  be a non-zero polynomial over  $M$  and choose minimal  $i$  such that  $a_i \neq 0$ . Then, for  $l = 1, 2$  we have  $p(\alpha_l) = a_i\alpha_l^i(1 + \varepsilon_l)$  with  $v(\varepsilon_l) > k$  for all

$k \in \mathbb{N}$ , and  $p(\alpha_l)$  has  $n$ -th root if and only if  $a_i \alpha_l^i$  had  $n$ -th root. We can find  $b_i \in M$  and  $c \in \mathbb{Z}$  such that  $a_i = b^n c$ . Hence  $a_i \alpha_l^i$  has  $n$ -th root if and only if  $c \alpha_l^i$  does. Since  $\mathbb{Z}$  is in the definable closure of  $\emptyset$ , we have  $\alpha_1 \equiv_{\mathbb{Z}} \alpha_2$  and  $P_n(c\alpha_1)$  if and only if  $P_n(c\alpha_2)$ .  $\square$

**Lemma 5.9.** *Let  $\mathcal{M}$  be a small model of  $T$  and let  $\alpha \in \mathbb{M}$  satisfy  $v(\alpha) > v(m)$  for every  $m \in M$ . Then the sequence  $\{nv(\alpha) : n \in \mathbb{N}\}$  is cofinal in the value group of  $M\langle\alpha\rangle$ , where  $M\langle\alpha\rangle$  is a prime model over  $M \cup \{\alpha\}$  (i.e.,  $M\langle\alpha\rangle = \text{dcl}(M \cup \{\alpha\})$ ).*

**Proof.** Let  $\gamma \in \Gamma_{M\langle\alpha\rangle}$  be arbitrary. As  $\gamma \in \text{dcl}(M \cup \{\alpha\})$ , we have that  $\gamma = v(f_m(\alpha))$  for some  $\emptyset$ -definable family of functions  $f$  and some tuple  $m$  in  $M$ . Consider the  $\emptyset$ -definable family of functions  $g : X \times \mathbb{M} \rightarrow \mathbb{M}$  with  $X = \mathbb{M} \setminus \{0\}$  given by  $g_m(x) := \frac{1}{f_m(\frac{1}{x})}$ .

Let  $n \in \mathbb{N}$  and the definable map  $a : M \rightarrow \Gamma_M$  be given by Definition 5.5 for the family  $g$  using Fact 5.6. Then  $-a(m) \in \Gamma_M$ , and so  $v(\alpha) > -a(m)$  by assumption. Hence  $v\left(\frac{1}{\alpha}\right) = -v(\alpha) < a(m)$  and so we have

$$-v(f_m(\alpha)) = v\left(\frac{1}{f_m(\alpha)}\right) = v\left(g_m\left(\frac{1}{\alpha}\right)\right) > nv\left(\frac{1}{\alpha}\right) = -nv(\alpha).$$

Hence  $\gamma = v(f_m(\alpha)) < nv(\alpha)$ , as wanted.  $\square$

**Lemma 5.10.** *Let  $p \in S_1(\emptyset)$  be arbitrary.*

- (1) *There is at most one global type  $\tilde{p} \in S_1(\mathbb{M})$  such that  $\tilde{p} \supseteq p \cup \{v(x) > v(m) : m \in \mathbb{M}\}$ , and  $\tilde{p}$  is  $\emptyset$ -invariant.*
- (2) *Assume  $\tilde{p}$  as in (1) exists. Let  $\mathcal{M} \prec \mathbb{M}$  and  $\alpha_1, \dots, \alpha_n \in \mathbb{M}$ . Then  $(\alpha_1, \dots, \alpha_n)$  realizes  $\tilde{p}^{(n)}|M$  if and only if each  $\alpha_i$  realizes  $p$ ,  $v(\alpha_1) > v(m)$  for all  $m \in M$  and  $v(\alpha_{i+1}) > kv(\alpha_i)$  for all  $k \in \mathbb{N}$  and  $i = 1, \dots, n-1$ .*

**Proof.** Part (1) follows from Proposition 5.7 and  $P$ -minimality.

Part (2) follows by the same argument as in Lemma 3.4 using Lemma 5.9.  $\square$

**Definition 5.11.** For an integer  $n > 0 \in \mathbb{N}$ , we say that a sequence  $a_i, i = 1, \dots, L$ , of elements of  $\mathcal{M}_0$  is *linearly  $n$ -growing* if  $v(a_0) > n$  and  $v(a_{i+1}) > nv(a_i)$  for all  $i$ .

Notice that a subsequence of a linearly  $n$ -growing sequence is also linearly  $n$ -growing.

**Lemma 5.12.** *For any finite set of formulas  $\Delta(x_1, \dots, x_k)$  with parameters from  $\mathcal{M}_0$  there are  $n \in \mathbb{N}$  and  $d_0 \in \mathbb{N}$  such that any linearly  $n$ -growing sequence of elements  $a_i \in \mathcal{M}_0$  of length  $N$  contains a  $\Delta$ -indiscernible subsequence of length at least  $\frac{N}{d_0}$ .*

**Proof.** Let  $\Sigma(x_1, \dots, x_{2k})$  be the partial type that is the union of

$$\Sigma_1 = \left\{ \bigwedge_{1 \leq i < j \leq 2k} \varphi(x_i) \leftrightarrow \varphi(x_j) : \varphi(x) \text{ is an } \mathcal{L}\text{-formula over } \emptyset \right\}$$

and

$$\Sigma_2 = \left\{ (x_1 > n) \wedge \bigwedge_{i=1}^{2k-1} v(x_{i+1}) > nv(x_i) : n \in \mathbb{N} \right\}.$$

Let  $(a_i : 1 \leq i \leq N)$  be any sequence of elements in  $\mathbb{M}$  such that all of the  $a_i$ 's have the same type over the empty set. Assume that  $v(a_0) > \mathbb{N}$  and  $v(a_{i+1}) > nv(a_i)$  for every  $i = 1, \dots, N-1$  and  $n \in \mathbb{N}$ . Then, by Lemma 5.10, the sequence  $(a_i : 1 \leq i \leq N)$  realizes  $\tilde{p}^{(N)}|_{\mathcal{M}_0}$  for  $p = \text{tp}(a_1/\emptyset)$ , and so is indiscernible over  $\mathcal{M}_0$  by Fact 2.8. It follows that

$$\Sigma(x_1, \dots, x_{2k}) \vdash \psi(x_1, \dots, x_k) \leftrightarrow \psi(x_{i_1}, \dots, x_{i_k})$$

for any  $1 \leq i_1 < i_2 < \dots < i_k \leq 2k$  and  $\psi \in \Delta$ .

By compactness, there are finite subsets  $\Sigma_1^0 \subseteq \Sigma_1$  and  $\Sigma_2^0 \subseteq \Sigma_2$  such that

$$\Sigma_1^0 \cup \Sigma_2^0 \vdash \psi(x_1, \dots, x_k) \leftrightarrow \psi(x_{i_1}, \dots, x_{i_k})$$

for any  $1 \leq i_1 < i_2 < \dots < i_k \leq 2k$  and  $\psi \in \Delta$ .

Let  $\varphi_1, \dots, \varphi_s$  be all  $\mathcal{L}$ -formulas over  $\emptyset$  appearing in  $\Sigma_1^0$ , and let  $n \in \mathbb{N}$  be maximal such that the condition  $v(x_{i+1}) > nv(x_i)$  appears in  $\Sigma_2^0$ .

Let  $d_0 = 2^s$ . Now any linearly  $n$ -growing sequence of length  $N$  contains a subsequence of length at least  $\frac{N}{d_0}$  satisfying the same  $\varphi_1, \dots, \varphi_s$ , and this subsequence is  $\Delta$ -indiscernible.  $\square$

As in the  $o$ -minimal case, combining Lemma 3.1 with Lemma 5.12 we can also allow additional parameters in  $\Delta$ .

**Corollary 5.13.** *For any finite set of  $\mathcal{L}$ -formulas  $\Delta(x_1, \dots, x_k; y)$  with parameters from  $\mathcal{M}_0$  there are  $n \in \mathbb{N}$  and  $d \in \mathbb{N}$  such that for all sufficiently large  $N$ , for any  $c \in \mathcal{M}_0^{|y|}$ , any linearly  $n$ -growing sequence  $\vec{a} = (a_1, a_2, \dots, a_N)$  of elements from  $\mathcal{M}_0$  contains a  $\Delta(x_1, \dots, x_k; c)$ -indiscernible subsequence of length at least  $\frac{N}{d}$ .*

It remains to establish an analog of Fact 3.8 in the  $p$ -adic case, demonstrating that there are ‘enough’ linearly  $n$ -growing sequences.

For elements  $\alpha \in \mathcal{M}_0$  and  $r$  in the value group of  $\mathcal{M}_0$ , we will denote by  $B(\alpha, r)$  the closed ball in  $\mathcal{M}_0$  of (valuational) radius  $r$  centered at  $\alpha$ , i.e.,

$$B(\alpha, r) = \{a \in \mathcal{M}_0 : v(a - \alpha) \geq r\}.$$

**Remark 5.14.** Since the value group  $\Gamma_{\mathcal{M}_0} = \mathbb{Z}$  is discrete by assumption, and the residue field is  $\mathbb{F}_q$  with  $q = p^t$  for some prime  $p$  and  $t \in \mathbb{N}$  by Fact 5.3, it is easy to see that every ball  $B(\alpha, r)$  is given by a disjoint union of  $q$  balls  $B(\alpha_i, r+1)$ ,  $i = 0, \dots, q-1$  with  $\alpha_i = \alpha + i\beta$ , where  $\beta \in \mathcal{M}_0$  is arbitrary with  $v(\beta) = r$ .

**Lemma 5.15.** *Let  $A \subseteq \mathcal{M}_0$  be a finite non-empty set with  $|A| \geq 2$  and  $A \subseteq B(\alpha, r)$  for some  $\alpha \in \mathcal{M}_0$  and  $r \in \mathbb{Z}$ . Then there is  $\alpha' \in B(\alpha, r)$  and  $r' > r$  such that  $\frac{1}{q}|A| \leq |A \cap B(\alpha', r')| < |A|$ .*

**Proof.** Let  $r_1 \in \mathbb{Z}$  be maximal such that some ball  $B(\alpha_1, r_1)$  with  $\alpha_1 \in B(\alpha, r)$  contains  $A$  (so  $r_1 \geq r$ ). As remarked above, the ball  $B(\alpha_1, r_1)$  is the union of  $q$  balls of radius  $r' = r_1 + 1$ . Hence for at least one of these  $q$  balls, say  $B(\alpha', r')$  with  $\alpha' \in B(\alpha_1, r_1)$ , we have  $\frac{1}{q}|A| \leq |A \cap B(\alpha', r')| < |A|$  (the last inequality is by maximality of  $r_1$ ).  $\square$

**Proposition 5.16.** *Let  $k \in \mathbb{N}$  be positive. For every finite  $A \subseteq \mathcal{M}_0$  with  $|A| \geq 2q^{k-1}$  there is  $\alpha \in \mathcal{M}_0$  and elements  $a_1, \dots, a_k \in A$  such that the valuations of  $\alpha - a_i, i = 1, \dots, k$ , are pairwise distinct.*

**Proof.** Let  $A \subseteq \mathcal{M}_0$  be a finite set with  $|A| \geq 2q^{k-1}$ . We set  $A_0 = A$ , and also choose  $\alpha_0 \in \mathcal{M}_0$  and  $r_0 \in \mathbb{Z}$  so that  $A \subseteq B(\alpha_0, r_0)$ .

Using Lemma 5.15, by induction on  $i = 1, \dots, k$  we construct finite sets  $A_0 \supsetneq A_1 \supsetneq \dots \supsetneq A_k$ , elements  $\alpha_1, \dots, \alpha_k \in \mathcal{M}_0$  and integers  $r_1 < r_2 < \dots < r_k$  such that:

- $A_i = B(\alpha_i, r_i) \cap A$ ;
- $|A_i| \geq \frac{1}{q^i} |A|$ ;
- $\alpha_i \in B(\alpha_{i-1}, r_{i-1})$  for  $i = 1, \dots, k$ .

We take  $\alpha = \alpha_k$ , and for  $i = 1, \dots, k$  we let  $a_i \in A_{i-1} \setminus A_i$  be arbitrary. Then  $a_i \in B(\alpha, r_{i-1}) \setminus B(\alpha, r_i)$ , hence  $r_{i-1} \leq v(\alpha - a_i) < r_i$ .  $\square$

We will use the following combinatorial facts.

**Fact 5.17** (Erdős–Szekeres Theorem [20]). *For any  $r, s \in \mathbb{N}$ , any sequence of pairwise distinct real numbers of length at least  $(r-1)(s-1)+1$  contains an increasing subsequence of length  $r$  or a decreasing subsequence of length  $s$ .*

**Fact 5.18** [6, Lemma 4.1]. *Given  $n \in \mathbb{N}$ , every strictly increasing sequence of real numbers of length  $4^n$  contains a subsequence  $(b_1, \dots, b_n)$  of length  $n$  such that either  $b_2 - b_1 \geq 2$  and  $b_{i+1} - b_i \geq 2(b_i - b_1)$  for all  $i = 2, \dots, n-1$ , or  $b_n - b_{n-1} \geq 2$  and  $b_n - b_{n-(i+1)} \geq 2(b_n - b_{n-i})$  for all  $i = 1, \dots, n-2$ .*

**Proposition 5.19.** *There are finitely many functions  $F_1(x, \bar{y}), \dots, F_s(x, \bar{y})$  definable with parameters from  $\mathcal{M}_0$  such that for any  $n \in \mathbb{N}$  there is a constant  $C > 0$  such that for any  $k \in \mathbb{N}$  the following holds. For any  $K \geq 2^{2^{Ck}}$  and any sequence  $\vec{a} = (a_1, \dots, a_K)$  in  $\mathcal{M}_0$  there are a linearly  $n$ -growing sequence  $\vec{b} = (b_1, \dots, b_k)$  of elements in  $\mathcal{M}_0$ ,  $\bar{c} \in \mathcal{M}_0^{|\bar{y}|}$  and  $i \in \{1, \dots, s\}$  such that one of the sequences*

$$F_i(\vec{b}, \bar{c}) := (F_i(b_1, \bar{c}), F_i(b_2, \bar{c}), \dots, F_i(b_k, \bar{c}))$$

or

$$F_i(\vec{b}, \bar{c}) := (F_i(b_k, \bar{c}), F_i(b_{k-1}, \bar{c}), \dots, F_i(b_1, \bar{c}))$$

is a subsequence of  $\vec{a}$ .

**Proof.** As usual, for a real number  $R$  we will denote by  $\lceil R \rceil$  the smallest integer  $N$  satisfying  $N \geq R$ .

First, notice that it is sufficient to prove the proposition for  $n = 2$ . Indeed if  $a_1, a_2, \dots$ , is a linearly 2-growing sequence of length  $N$ , then for a given  $n$ , taking  $l = \lceil \log_2 n \rceil$ , the sequence  $a_l, a_{2l}, \dots$  is a linearly  $n$ -growing sequence of length at least  $\frac{N}{l}$ .

Assume  $k$  is given and  $K \geq 2^{2^{Ck}}$ , where a suitable constant  $C$  will be determined in the proof. Let  $a_1, \dots, a_K$  be a sequence of elements of  $\mathcal{M}_0$ .

*Case 1.* The sequence  $a_1, \dots, a_K$  contains at least  $\sqrt{K}$  equal elements.

Let us call this repeated element  $a'$ . Then the conclusion of the proposition holds as we can map any linearly 2-growing sequence of length  $\lceil \sqrt{K} \rceil$  onto a subsequence  $\vec{a}$  using the constant map  $F(x, a') := a'$ .

*Case 2.* The sequence  $a_1, \dots, a_K$  does not contain  $\lceil \sqrt{K} \rceil$  equal elements. Then it contains at least  $K_1 := \sqrt{K}$  pairwise distinct elements.

Using Proposition 5.16 we can find an element  $\alpha \in \mathcal{M}_0$  and a subsequence  $\vec{a}^1 = (a_1^1, \dots, a_{K_2}^1)$  of  $\vec{a}$  with  $K_2 := \lceil \log_q(\frac{1}{2}K_1) + 1 \rceil$  such that the valuations  $v(\alpha - a_i^1)$  are pairwise distinct for all  $1 \leq i \leq K_2$ .

Thus, using the map  $F(x, \alpha) = x + \alpha$  we can find a sequence  $\vec{b} = (b_1, \dots, b_{K_2})$  such that  $F(\vec{b}, \alpha) = (F(b_1, \alpha), \dots, F(b_{K_2}, \alpha))$  is a subsequence of  $\vec{a}$  and all of the valuations  $v(b_i)$  are pairwise distinct.

By the Erdős–Szekeres Theorem (Fact 5.17), the sequence  $\vec{b}$  contains a subsequence  $\vec{b}^1 = (b_1^1, \dots, b_{K_3}^1)$  with  $K_3 := \lceil \sqrt{K_2} \rceil$  such that the corresponding sequence of valuations  $(v(b_1^1), \dots, v(b_{K_3}^1))$  is either increasing or decreasing. Using the function  $F(x) = x^{-1}$  if needed, we can assume that the sequence is increasing.

By Fact 5.18, there is a subsequence  $\vec{b}^2 = (b_1^2, \dots, b_{K_4}^2)$  of  $\vec{b}^1$  with  $K_4 := \lceil \frac{1}{2} \log_2 K_3 \rceil$  such that either for the sequence

$$\vec{v} = (v(b_2^2) - v(b_1^2), v(b_3^2) - v(b_1^2), \dots, v(b_{K_4}^2) - v(b_1^2))$$

we have  $v_1 \geq 2, v_{i+1} \geq 2v_i$ , or for the sequence

$$\vec{v}' = (v(b_{K_4}^2) - v(b_{K_4-1}^2), v(b_{K_4}^2) - v(b_{K_4-2}^2), \dots, v(b_{K_4}^2) - v(b_1^2))$$

we have  $v_1 \geq 2, v_{i+1} \geq 2v_i$ .

In the first case, the sequence  $\vec{b}^3 = (b_2^2/b_1^2, \dots, b_{K_4}^2/b_1^2)$  is linearly 2-growing and can be embedded into  $\vec{b}^2$  via the transformation  $x \mapsto b_1^2 x$ .

In the second case, the sequence  $\vec{b}^4 = (b_{K_4}^2/b_{K_4-1}^2, \dots, b_{K_4}^2/b_1^2)$  is linearly 2-growing, and its reverse sequence  $\vec{b}^4$  can be embedded into  $\vec{b}^2$  via the transformation  $x \mapsto x^{-1}/b_{K_4}^2$ .

Hence we have demonstrated that every sequence of length  $K$  contains a subsequence of length  $\min\{\lceil \sqrt{K} \rceil, K_4\}$  with the desired property. Going backwards through the proof we have  $K_3 \leq 4^{K_4}$ ,  $K_2 \leq K_3^2$ ,  $K_1 \leq 2q^{K_2-1}$  and  $K \leq K_1^2$ . As  $q$  is fixed, an easy calculation shows that for any sufficiently large  $C \in \mathbb{R}$ , we have  $K \leq 2^{2^{CK_4}}$  for all sufficiently large  $K_4 \in \mathbb{N}$ , and taking  $K_4 = k$  we can conclude the result.  $\square$

Combining Corollary 5.13 and Proposition 5.19 exactly as in the  $\mathcal{o}$ -minimal case (see Theorem 3.9), we obtain Theorem 5.4.

## 6. Ramsey growth in NIP

In this section we consider Ramsey numbers for definable relations of higher arity. We fix a structure  $\mathcal{M}$  in a language  $\mathcal{L}$ , and by a ‘formula’ we always mean an  $\mathcal{L}$ -formula. Following the method of [13] for the semialgebraic case, we obtain the following recursive bound for higher arity Ramsey numbers in arbitrary NIP structures.

**Theorem 6.1.** *Let  $\mathcal{M}$  be an NIP structure,  $k \geq 3$  and  $\varphi(x_1, \dots, x_k; z)$  a formula with  $|x_1| = \dots = |x_k| = d$ . Then, defining the formula  $\psi(x_1, \dots, x_{k-1}; z') := \varphi(x_1, \dots, x_{k-1}; x_k, z)$  and taking  $m := R_\psi^*(n-1)$ , for all large enough  $n$  we have*

$$R_\varphi^*(n) \leq 2^{Cm \log m}$$

for some constant  $C = C(\varphi)$ .

**Proof.** We are generalizing the argument from [13, Theorem 2.2].

Let  $e \in M^{|z|}$  be arbitrary,  $\psi(x_1, \dots, x_{k-1}; z') = \varphi(x_1, \dots, x_{k-1}; x_k, z)$ ,  $n \in \mathbb{N}$  large enough (to be determined in the proof), and  $m = R_\psi^*(n-1)$ . Let  $\vec{a} = (a_1, \dots, a_N)$  be a sequence of elements in  $M^d$  with  $N \geq 2^{Cm \log m}$ , where  $C = C(\varphi)$  is a constant to be specified later. We need to find a  $\varphi(x_1, \dots, x_k; e)$ -indiscernible subsequence of  $\vec{a}$  of length  $n$ .

Let  $E \subseteq (M^d)^k$  be the  $k$ -ary relation on  $M^d$  defined by  $\varphi(x_1, \dots, x_k; e)$ , i.e.,  $E = \{(t_1, \dots, t_k) \in (M^d)^k : \mathcal{M} \models \varphi(t_1, \dots, t_k; e)\}$ .

The idea is to find a subsequence  $\vec{b} = (b_1, \dots, b_{m+1})$  of  $\vec{a}$  such that for all  $1 \leq i_1 < \dots < i_{k-1} \leq m$ , either  $(b_{i_1}, \dots, b_{i_{k-1}}, b_i) \in E$  for all  $i_{k-1} < i \leq m+1$  or  $(b_{i_1}, \dots, b_{i_{k-1}}, b_i) \notin E$  for all  $i_{k-1} < i \leq m+1$ .

To build a sequence  $\vec{b}$  as above, we recursively choose elements  $b_r$  in  $\vec{a}$  and also subsequences  $\vec{c}_r$  of  $\vec{a}$  for  $r = k-2, k-1, \dots, m+1$  with  $\vec{c}_{r+1} \subset \vec{c}_r$  so that the following holds.

- (1) For every  $(k-1)$ -subsequence  $(b_{i_1}, \dots, b_{i_{k-1}})$  of  $(b_1, \dots, b_{r-1})$  with  $i_1 < \dots < i_{k-1}$ , either  $(b_{i_1}, \dots, b_{i_{k-1}}, b) \in E$  for every  $b \in \{b_j : i_{k-1} < j \leq r\} \cup \vec{c}_r$  or  $(b_{i_1}, \dots, b_{i_{k-1}}, b) \notin E$  for every  $b \in \{b_j : i_{k-1} < j \leq r\} \cup \vec{c}_r$ .
- (2)  $|\vec{c}_r| \geq \frac{N}{C_1^r C_2^{r-1}}$ , where  $C_1, C_2$  are some constants depending just on  $\varphi$ .
- (3) The subsequence  $(b_1, \dots, b_r)$  appears in  $\vec{a}$  in front of the subsequence  $\vec{c}_r$ , i.e.,  $(b_1, \dots, b_r) \hat{\ } \vec{c}_r$  is a subsequence of  $\vec{a}$ .

We start with  $r = k-2$  by taking  $(b_1, \dots, b_{k-2}) = (a_1, \dots, a_{k-2})$  and  $\vec{c}_{k-2} = (a_{k-1}, \dots, a_N)$ . Assume we have obtained  $(b_1, \dots, b_r)$  and  $\vec{c}_r$  satisfying (1)–(3) above, and we define  $b_{r+1}$  and  $\vec{c}_{r+1}$  as follows.

Let  $b_{r+1}$  be the first element in  $\vec{c}_r$  and let  $\vec{c}_r^*$  be the sequence  $\vec{c}_r$  with the first element removed. Let  $\theta(x_k; u)$  be the partitioned formula obtained from  $\varphi(x_1, \dots, x_{k-1}, x_k, z)$  by partitioning its variables into two groups  $x_k$  and  $u = x_1, \dots, x_{k-1}, z$ . As the formula  $\theta$  is NIP, by Fact 2.3 the number of complete  $\theta(x_k; u)$ -types over an arbitrary finite set  $D \subseteq M^{|z|+(k-1)d}$  of parameters is bounded by  $C_3|D|^{C_4}$  for some constants  $C_3, C_4$  depending just on  $\varphi$ .

Let  $D = \{(b_{i_1}, \dots, b_{i_{k-1}}, e) : 1 \leq i_1 < \dots < i_{k-1} \leq r+1\}$ . Obviously  $|D| \leq (r+1)^{k-1}$ .

It follows by the pigeonhole principle that there is some complete  $\theta$ -type  $p(x_k) \in S_\theta(D)$  such that the number of elements in  $\vec{c}_r^*$  realizing  $p(x)$  is at least  $\frac{|\vec{c}_r^*|}{C_3|D|^{C_4}} \geq \frac{|\vec{c}_r|-1}{C_3|(r+1)|^{(k-1)C_4}} \geq \frac{|\vec{c}_r|}{2C_3|(r+1)|^{(k-1)C_4}}$ , provided  $|\vec{c}_r| \geq 2$ .



We take  $\vec{c}_{r+1}$  to be the subsequence of elements of  $\vec{c}_r$  realizing  $p$ . For  $C_1 = 2C_3$  and  $C_2 = (k-1)C_4$  (again, both  $C_1$  and  $C_2$  only depend on  $\varphi$ ), using the inductive lower bound for the length of  $\vec{c}_r$  and calculating, we obtain  $|\vec{c}_{r+1}| \geq \frac{N}{C_1^{(r+1)}(r+1)^{rC_2 \cdot (r+1)}}$ , i.e., (2) is satisfied.

Now for any subsequence  $(b_{i_1}, \dots, b_{i_{k-1}})$  of  $(b_1, \dots, b_{r+1})$ , we have that either  $(b_{i_1}, \dots, b_{i_{k-1}}, b) \in E$  for all  $b \in \vec{c}_{r+1}$  or  $(b_{i_1}, \dots, b_{i_{k-1}}, b) \notin E$  for all  $b \in \vec{c}_{r+1}$ . Together with the inductive assumption this implies that (1) is satisfied by  $(b_1, \dots, b_{r+1})$  and  $\vec{c}_{r+1}$ . Finally, (3) is clear from the construction.

For  $\vec{c}_m$  to be non-empty (in which case we would have constructed our sequence  $(b_1, \dots, b_{m+1})$ ), by (2) we need  $\frac{N}{C_1^m m^{C_2 \cdot m}} \geq 1$ , i.e.,  $N \geq C_1^m m^{C_2 \cdot m}$ . It is not hard to find a constant  $C$ , depending on  $C_1, C_2$  only, so that the condition  $N \geq 2^{Cm \log m}$  is sufficient.  $\square$

**Remark 6.2.** The constant  $C_4$  in the above proof depends just on the VC-density of  $\varphi$  (with a corresponding partition of the variables). By Fact 2.6, in the case of  $\mathcal{o}$ -minimal theories we can take  $C_4 = d$ .

By a repeated application of Theorem 6.1 we have an improved bound on Ramsey numbers for relations of higher arities.

**Theorem 6.3.** *Let  $\mathcal{M}$  be an NIP structure, and assume that for all  $\psi(x_1, x_2; z)$  we have  $R_\varphi^*(n) \leq n^c$  for some  $c = c(\psi)$  and all  $n$  large enough. Then for all  $\varphi(x_1, \dots, x_k; z')$  we have  $R_\varphi^*(n) \leq \text{twr}_{k-1}(n^c)$  for some  $c = c(\varphi)$  and all  $n$  large enough.*

Now we discuss the connection of the assumption of Theorem 6.3 with the (strong) Erdős–Hajnal property for graphs definable in  $\mathcal{M}$ .

**Definition 6.4** [25].

- (1) Let  $\mathcal{G}$  be a class of finite graphs (i.e., the edge relation is assumed to be symmetric and irreflexive). We say that  $\mathcal{G}$  has the *Erdős–Hajnal property*, or the *EH property*, if there is  $\delta > 0$  such that every  $G = (V, E) \in \mathcal{G}$  has a homogeneous subset  $V_0$  of size  $|V_0| \geq |V|^\delta$  (i.e., either  $(a, b) \in E$  for all  $a \neq b \in V_0$ , or  $(a, b) \notin E$  for all  $a \neq b \in V_0$ ).
- (2) Let  $\mathcal{G}$  be a class of finite binary relations, i.e., every member of  $\mathcal{G}$  is of the form  $(E, V_1, V_2)$ , where  $E \subseteq V_1 \times V_2$  with  $V_1, V_2$  finite sets (not necessarily disjoint). We say that  $\mathcal{G}$  has the *strong EH property* if there is  $\delta > 0$  such that for every  $(E, V_1, V_2) \in \mathcal{G}$  there are subsets  $V'_i \subseteq V_i$  with  $|V'_i| \geq \delta|V_i|$  for  $i = 1, 2$  such that the pair of sets  $V'_1, V'_2$  is homogeneous (i.e., either  $V'_1 \times V'_2 \subseteq E$  or  $V'_1 \times V'_2 \cap E = \emptyset$ ).
- (3) A family of finite graphs  $\mathcal{G}$  has the strong EH property if the family of finite binary relations  $\{(E, V, V) : (E, V) \in \mathcal{G}\}$  has the strong EH property.

We recall that a famous conjecture of Erdős and Hajnal [12] says that for every finite graph  $H$ , the family of all finite graphs not containing an induced copy of  $H$  has the EH property.

**Definition 6.5.** Let  $\mathcal{M}$  be a first-order structure, and  $\varphi(x_1, x_2; z)$  a formula with  $|x_1| = |x_2| = d$ . Let  $\mathcal{G}_\varphi$  be the family of all finite binary relations  $(E, V_1, V_2)$  with  $V_1, V_2 \subseteq M^d$  finite and  $E = (V_1 \times V_2) \cap \varphi(M, b)$  for some  $b \in M^{|z|}$ . Let  $\mathcal{G}_\varphi^{\text{sym}}$  be the family of all finite graphs  $(V, E)$  with  $V \subseteq M^d$  and  $E = (V \times V) \cap \varphi(M, b) \setminus \Delta$  for some  $b \in M^z$  such that  $E$  is symmetric (where  $\Delta = \{(v, v) : v \in V\}$  is the diagonal). We say that  $\varphi$  satisfies the EH property (respectively strong EH property) if the family  $\mathcal{G}_\varphi^{\text{sym}}$  (respectively  $\mathcal{G}_\varphi$ ) does.

If this holds for all formulas  $\varphi$  in  $\mathcal{M}$ , we say that  $\mathcal{M}$  satisfies the (strong) EH property.

**Remark 6.6.** It is shown in [2] that if a family of finite graphs  $\mathcal{G}$  has the strong EH property and is closed under taking induced subgraphs then it has the EH property. In particular, this applies to every family of the form  $\mathcal{G}_\varphi^{\text{sym}}$  as in Definition 6.5.

Hence,  $\mathcal{M}$  satisfies the EH property precisely when the assumption of Theorem 6.3 holds for all *symmetric* definable relations. By the results in [11] we know that this property holds in arbitrary reducts of distal structures.

**Definition 6.7.** A structure  $\mathcal{M}$  is *distal* if the following holds.

For every formula  $\varphi(x, y)$  there is a formula  $\theta(x, y_1, \dots, y_n)$  with  $|y_1| = \dots = |y_n| = |y|$  such that: for any finite  $B \subseteq M^{|y|}$  with  $|B| \geq 2$  and any  $a \in M^{|x|}$ , there are  $b_1, \dots, b_n \in B$  such that  $\mathcal{M} \models \theta(a, b_1, \dots, b_n)$  and for any  $b \in B$ , either  $\varphi(M, b) \subseteq \theta(M, b_1, \dots, b_n)$  or  $\varphi(M, b) \cap \theta(M, b_1, \dots, b_n) = \emptyset$ .

Distality was introduced in [49], the equivalence of the original definition and the combinatorial definition above is from [8], and the connection to combinatorics is from [11] (see also [7]). Important examples of distal structures are given by arbitrary (weakly) *o*-minimal and *P*-minimal structures. We refer to the introduction of [11] for a detailed discussion of distality.

**Fact 6.8** [8]. *If  $\mathcal{M}$  is a reduct of a distal structure then it satisfies the strong EH property (and so the EH property as well, by Remark 6.6).*

In the next proposition, we demonstrate that in any structure satisfying the strong EH property, all (not necessarily symmetric) definable binary relations also satisfy a polynomial Ramsey bound.

**Proposition 6.9.** *Let  $\mathcal{M}$  be a structure satisfying the strong EH property. Then for all  $\varphi(x, y; z)$  with  $|x| = |y|$  we have  $R_\varphi^*(n) \leq n^c$  for some  $c = c(\varphi)$  and all  $n$  large enough.*

**Proof.** Let  $d := |x| = |y|$ , and let  $E(x, y) \subseteq M^d \times M^d$  be a definable relation given by  $\varphi(x, y; b)$  for some parameter  $b \in M^{|z|}$ . We want to show that there is some real  $c = c(\varphi) > 0$  such that every finite sequence from  $M^d$  of length  $n$  contains an  $E$ -indiscernible subsequence of length  $n^c$ . By Fact 6.8 we know that it is true in the case when  $E$  is symmetric.

Let  $\vec{a} = (a_1, \dots, a_n)$  be a sequence in  $M^d$ . For simplicity we will assume that all of the  $a_i$ 's are pairwise distinct. We can always achieve it by taking a subsequence of length  $\sqrt{n}$ .

We will also assume that  $\models \neg E(x, x)$ , i.e.,  $E$  is irreflexive (replacing  $E(x, y)$  by  $E'(x, y) := E(x, y) \wedge x \neq y$ , any  $E'$ -indiscernible subsequence of  $\vec{a}$  is also  $E$ -indiscernible).

Consider the relation  $E_0(x, y) = E(x, y) \vee E(y, x)$ . It is symmetric. Hence  $\vec{a}$  contains an  $E_0$ -indiscernible subsequence of length  $n^{c_1}$ , with  $c_1 = c_1(\varphi) > 0$ . If  $\neg E_0(x, y)$  holds on every increasing pair of elements in this subsequence then we are done. Otherwise, replacing  $\vec{a}$  with this subsequence, we may assume that  $E_0(x, y)$  holds on  $\vec{a}$ .

Now consider the relation  $E_1(x, y) = E(x, y) \wedge E(y, x)$ . Again it is symmetric, so  $\vec{a}$  contains an  $E_1$ -indiscernible subsequence of polynomial length. If  $E_1(x, y)$  holds on this subsequence then we are done. Assume otherwise, then again replacing  $\vec{a}$  with this subsequence we may assume that  $\neg E_1(x, y)$  holds on  $\vec{a}$ .

Let  $A = \{a_i : i \in \{1, \dots, n\}\}$ . We have that for  $a \neq b \in A$  exactly one of  $E(a, b)$  or  $E(b, a)$  holds, and we also have that  $\neg E(a, a)$  holds for all  $a \in A$ . Hence  $E$  is a *tournament* on  $A$ .

Our goal is to show that for some  $A_0 \subseteq A$  of size  $n^{c_2}$ , with  $c_2 = c_2(\varphi) > 0$ ,  $E$  restricted to  $A_0$  defines a linear order. Then, by the Erdős–Szekeres Theorem (Fact 5.17), a subsequence corresponding to  $A_0$  would contain an  $E$ -monotone subsequence of length  $\sqrt{|A_0|}$  and we would be done.

For an integer  $m \leq n$ , let us denote by  $f(m)$  the maximal  $k$  such that every subset  $A' \subseteq A$  of size  $m$  contains a linearly ordered subset of size  $k$ . Obviously, we have  $f(m) \geq 1$  for all  $m \geq 1$ .

Now we use the strong EH property. We know that there is  $0 < \alpha < 1$ , with  $\alpha = \alpha(\varphi)$ , such that for any  $B \subseteq A$  there are disjoint subsets  $B_0, B_1 \subseteq B$  with  $|B_0|, |B_1| \geq \alpha|B|$  that are  $E$ -homogeneous. If  $C_0 \subseteq B_0, C_1 \subseteq B_1$  are subsets linearly ordered by  $E$ , then by  $E$ -homogeneity  $C_0 \cup C_1$  is also linearly ordered by  $E$ .

This implies that  $f(m) \geq 2f(\alpha m)$  and for any  $s \in \mathbb{N}$  we get  $f(m) \geq 2^s f(\alpha^s m)$ .

Recall that  $|A| = n$ . We choose the maximal  $s$  such that  $\alpha^s n \geq 1$ . Up to taking the integer part, we have

$$s \log(\alpha) + \log(n) \geq 0, \text{ i.e., we have } s \geq \frac{-\log n}{\log \alpha}.$$

Then we get

$$f(n) \geq 2^{\frac{-\log n}{\log \alpha}} = n^{-\frac{1}{\log \alpha}},$$

and taking  $c := -\frac{1}{2 \log \alpha} > 0$  we can conclude the result.  $\square$

Hence the assumption of Theorem 6.3 is satisfied in reducts of distal structures by Fact 1.11 and Proposition 6.9. As every distal structure is NIP, and every reduct of an NIP structure is NIP, applying Theorem 6.1 we get the following.

**Corollary 6.10.** *Let  $\mathcal{M}$  be a reduct of a distal structure. Then for any  $\varphi(x_1, \dots, x_k; z)$  we have  $R_\varphi^*(n) \leq \text{twr}_{k-1}(n^c)$  for some  $c = c(\varphi)$  and all  $n$  large enough.*

Note that the assumption of Theorem 6.3 is also trivially satisfied in stable structures by Fact 1.10. We conjecture that it holds in arbitrary NIP structures.

**Conjecture 6.11.** *If  $\mathcal{M}$  is an NIP structure and  $\varphi(x_1, x_2; z)$  is a formula, then  $R_\varphi^*(n) \leq n^c$  for some  $c = c(\varphi, \mathcal{M})$  and all sufficiently large  $n$ .*

This conjecture, in the case of a symmetric formula, is equivalent to saying that all graphs definable in NIP structures satisfy the Erdős–Hajnal property. We refer the reader to [9, 11] for further discussion.

**Acknowledgements.** We would like to thank Martin Hils for encouraging us to present the results in §5 in full generality, Pablo Cubides Kovacsics for a very helpful discussion on polynomial boundedness in  $P$ -minimal structures, and the anonymous referee for some suggestions on improving the presentation.

Chernikov was supported by the NSF Research Grant DMS-1600796; by the NSF CAREER grant DMS-1651321, and by an Alfred P. Sloan Fellowship.

Starchenko was supported by the NSF Research Grant DMS-1500671.

Thomas was supported by the DFG Research Grant TH 1781/2-1; by the Zukunftskolleg, University of Konstanz and by the NSERC Discovery Grant RGPIN 261961.

This work was finished during the ‘Model Theory, Combinatorics and Valued Fields’ trimester program at Institut Henri Poincaré. We thank IHP for their hospitality.

## References

1. H. ADLER, An introduction to theories without the independence property, *Arch. Math. Logic* **5** (2008), <http://www.logic.univie.ac.at/~adler/docs/nip.pdf>.
2. N. ALON, J. PACH, R. PINCHASI, R. RADOIČIĆ AND M. SHARIR, Crossing patterns of semi-algebraic sets, *J. Combin. Theory Ser. A* **111**(2) (2005), 310–326.
3. M. ASCHENBRENNER, A. DOLICH, D. HASKELL, D. MACPHERSON AND S. STARCHENKO, Vapnik–Chervonenkis density in some theories without the independence property, I, *Trans. Amer. Math. Soc.* **368**(8) (2016), 5889–5949.
4. S. BASU, Combinatorial complexity in o-minimal geometry, *Proc. Lond. Math. Soc.* **100**(2) (2010), 405–428.
5. F. A. BEHREND, On sets of integers which contain no three terms in arithmetical progression, *Proc. Natl. Acad. Sci. USA* **32** (1946), 331–332.
6. B. BUKH AND J. MATOUSEK, Erdős–Szekeres-type statements: Ramsey function and decidability in dimension 1, *Duke Math. J.* **163**(12) (2014), 2243–2270.
7. A. CHERNIKOV, D. GALVIN AND S. STARCHENKO, Cutting lemma and Zarankiewicz’s problem in distal structures, Preprint, 2016, [arXiv:1612.00908](https://arxiv.org/abs/1612.00908).
8. A. CHERNIKOV AND P. SIMON, Externally definable sets and dependent pairs II, *Trans. Amer. Math. Soc.* **367**(7) (2015), 5217–5235.
9. A. CHERNIKOV AND S. STARCHENKO, Definable regularity lemmas for NIP hypergraphs, Preprint, 2016, [arXiv:1607.07701](https://arxiv.org/abs/1607.07701).
10. A. CHERNIKOV AND S. STARCHENKO, A note on the Erdős–Hajnal property for stable graphs, *Proc. Amer. Math. Soc.* **146**(2) (2018), 785–790.
11. A. CHERNIKOV AND S. STARCHENKO, Regularity lemma for distal structures, *J. Eur. Math. Soc. (JEMS)* **20**(10) (2018), 2437–2466.
12. M. CHUDNOVSKY, The Erdős–Hajnal Conjecture: A survey, *J. Graph Theory* **75**(2) (2014), 178–190.
13. D. CONLON, J. FOX, J. PACH, B. SUDAKOV AND A. SUK, Ramsey-type results for semi-algebraic relations, *Trans. Amer. Math. Soc.* **366**(9) (2014), 5043–5065.

14. D. CONLON, J. FOX AND B. SUDAKOV, Hypergraph Ramsey numbers, *J. Amer. Math. Soc.* **23**(1) (2010), 247–266.
15. L. DARNIERE AND I. HALUPCZOK, Cell decomposition and classification of definable sets in  $p$ -optimal fields, *J. Symbolic Logic* **82**(1) (2017), 120–136.
16. L. VAN DEN DRIES, *Tame Topology and o-minimal Structures* vol. 248, London Mathematical Society Lecture Note Series, Volume 248 (Cambridge University Press, Cambridge, 1998). x+180 pp.
17. L. VAN DEN DRIES, D. HASKELL AND D. MACPHERSON, One-dimensional  $p$ -adic subanalytic sets, *J. Lond. Math. Soc.* **59**(1) (1999), 1–20.
18. M. ELIAS, J. MATOUSEK, E. ROLDÁN-PENSADO AND Z. SAFERNOVÁ, Lower bounds on geometric Ramsey functions, *SIAM J. Discrete Math.* **28**(4) (2014), 1960–1970.
19. D. ENSLEY AND R. GROSSBERG, Ramsey’s theorem in stable structures, manuscript (1997). URL: <http://www.math.cmu.edu/~rami/finite.pdf>.
20. P. ERDŐS AND G. SZEKERES, A combinatorial problem in geometry, *Compos. Math.* **2** (1935), 463–470.
21. P. ERDŐS, Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.* **53**(4) (1947), 292–294.
22. P. ERDŐS, A. HAJNAL AND R. RADO, Partition relations for cardinal numbers, *Acta Math. Hungar.* **16**(1–2) (1965), 93–196.
23. P. ERDŐS AND R. RADO, Combinatorial theorems on classifications of subsets of a given set, *Proc. Lond. Math. Soc.* **3**(1) (1952), 417–439.
24. P. ERDŐS AND G. SZEKERES, A combinatorial problem in geometry, *Compos. Math.* **2** (1935), 463–470.
25. J. FOX AND J. PACH, Erdős-Hajnal-type results on intersection patterns of geometric objects, in *Horizons of Combinatorics*, Bolyai Soc. Math. Stud., Volume 17, pp. 79–103 (Springer, Berlin, 2008).
26. R. L. GRAHAM, B. L. ROTHSCILD AND J. H. SPENCER, *Ramsey Theory*, Wiley-Interscience Series in Discrete Mathematics. A Wiley-Interscience Publication (John Wiley & Sons, Inc., New York, 1980). ix+174 pp.
27. V. GUINGONA, On uniform definability of types over finite sets, *J. Symbolic Logic* **77**(02) (2012), 499–514.
28. D. HASKELL AND D. MACPHERSON, A version of  $o$ -minimality for the  $p$ -adics, *J. Symbolic Logic* **62**(04) (1997), 1075–1092.
29. E. HRUSHOVSKI AND A. PILLAY, On NIP and invariant measures, *J. Eur. Math. Soc. (JEMS)* **13**(4) (2011), 1005–1061.
30. H. JOHNSON AND M. LASKOWSKI, Compression schemes, stable definable families, and  $o$ -minimal structures, *Discrete Comput. Geom.* **43**(4) (2010), 914–926.
31. I. KAPLAN AND S. SHELAH, A dependent theory with few indiscernibles, *Israel J. Math.* **202**(1) (2014), 59–103.
32. P. C. KOVACSICS AND F. DELON, Definable functions in tame expansions of algebraically closed valued fields, Preprint, 2018, [arXiv:1802.03323](https://arxiv.org/abs/1802.03323).
33. M. LASKOWSKI, Vapnik-Chervonenkis classes of definable sets, *J. Lond. Math. Soc.* **2**(2) (1992), 377–384.
34. A. MACINTYRE, On definable subsets of  $p$ -adic fields, *J. Symbolic Logic* **41**(3) (1976), 605–610.
35. M. MALLIARIS AND S. SHELAH, Regularity lemmas for stable graphs, *Trans. Amer. Math. Soc.* **366**(3) (2014), 1551–1585.
36. J. MATOUSEK, *Lectures on Discrete Geometry*, Graduate Texts in Mathematics, Volume 212 (Springer, New York, 2002).
37. C. MILLER, Exponentiation is hard to avoid, *Proc. Amer. Math. Soc.* **122**(1) (1994),

- 257–259.
38. C. MILLER AND S. STARCHENKO, A growth dichotomy for o-minimal expansions of ordered groups, *Trans. Amer. Math. Soc.* **350**(9) (1998), 3505–3521.
  39. S. MORAN AND A. YEHUDAYOFF, Sample compression schemes for VC classes, *J. ACM (JACM)* **63**(3) (2016), 21.
  40. A. PRESTEL AND P. ROQUETTE, *Peter Formally p-Adic Fields*, Lecture Notes in Mathematics, Volume 1050, (Springer, Berlin, 1984). v+167 pp.
  41. F. P. RAMSEY, On a problem of formal logic, *Proc. Lond. Math. Soc.* **2**(1) (1930), 264–286.
  42. N. SAUER, On the density of families of sets, *J. Combin. Theory Ser. A* **13**(1) (1972), 145–147.
  43. T. SCANLON, O-minimality as an approach to the André–Oort conjecture, in *Around the Zilber–Pink conjecture/Autour de la conjecture de Zilber–Pink*, Panor. Synthèses, Volume 52, pp. 111–165 (Soc. Math. France, Paris, 2017).
  44. S. SHELAH, Classification theory and the number of nonisomorphic models, in *Studies in Logic and the Foundations of Mathematics* vol. 92, 2nd edn (North-Holland Publishing Co., Amsterdam, 1990).
  45. S. SHELAH, A combinatorial problem; stability and order for models and theories in infinitary languages, *Pacific J. Math.* **41**(1) (1972), 247–261.
  46. S. SHELAH, *Around Classification Theory of Models*, Lecture Notes in Mathematics, Volume 1182 (Springer, Berlin, 1986). viii+279 pp.
  47. S. SHELAH, Classification theory for elementary classes with the dependence property—a modest beginning, *Sci. Math. Jpn.* **59**(2) (2004), 265–316. Special issue on set theory and algebraic model theory.
  48. S. SHELAH, Strongly dependent theories, *Israel J. Math.* **204**(1) (2014), 1–83.
  49. P. SIMON, Distal and non-distal NIP theories, *Ann. Pure Appl. Logic* **164**(3) (2013), 294–318.
  50. P. SIMON, *A Guide to NIP Theories*, Lecture Notes in Logic, Volume 44 (Association for Symbolic Logic, Chicago, IL; Cambridge Scientific Publishers, Cambridge, 2015). vii+156 pp.
  51. A. J. WILKIE, Model completeness results for expansions of the ordered field of real numbers by restricted Pfaffian functions and the exponential function, *J. Amer. Math. Soc.* **9**(4) (1996), 1051–1094.