ON LARGE EXTERNALLY DEFINABLE SETS IN NIP

MARTIN BAYS, OMER BEN-NERIA, ITAY KAPLAN, AND PIERRE SIMON

Abstract. We study cofinal systems of finite subsets of ω_1. We show that while such systems can be NIP, they cannot be defined in an NIP structure. We deduce a positive answer to a question of Chernikov and Simon from 2013: in an NIP theory, any uncountable externally definable set contains an infinite definable subset. A similar result holds for larger cardinals.

1. Introduction

Suppose that M is a structure and x a tuple of variables. Recall that a set $X \subseteq M^x$ is M**-definable** if there is some formula $\phi(x)$ over M such that $\phi(M) = X$. The set X is **externally definable** if there is some elementary extension $N \succ M$ and a formula $\psi(x)$ over N such that $X = \psi(M) = \{a \in M^x \mid N \vDash \psi(a)\}$.

When $\mathrm{Th}(M)$ is stable, all externally definable subsets are in fact M-definable (this is a characterization of stability: all types over any model are definable).

Let T be a theory. We consider the following natural question:

Question 1.1. Is there some (infinite) cardinal λ such that for any $M \vDash T$ and any externally definable set $X \subseteq M^k$, if $|X| \geq \lambda$ then X contains an infinite M-definable subset?

We cannot hope to say much about externally definable sets in arbitrary theories. In particular, supposing that T has a strong form of IP, the answer to Question 1.1 is negative (see Remark 5.5). On the other hand:

Fact 1.2. *[CS13, Corollary 1.12] Suppose T is NIP. Then the answer to Question 1.1 is positive: one can take $\lambda = \beth_\omega$.*

For a complete theory T, let $\mathrm{ext}(T)$ be the minimal λ as in Question 1.1 if such exists, and $\mathrm{ext}(T) = \infty$ otherwise. If T is NIP, Fact 1.2 shows that $\mathrm{ext}(T) \leq \beth_\omega$. We first observe that we cannot hope to improve this to $\mathrm{ext}(T) = \aleph_0$.

Example 1.3. [CS13, just above Question 1.13] Let M be the linear order $(\omega+\mathbb{Z}, <)$, whose theory is NIP (and even dp-minimal, see [Sim15, Proposition A.2]). Then, ω is externally definable (as is any cut), but no infinite subset of ω is M-definable, since $\mathrm{Th}(M)$ has quantifier elimination after adding the successor and predecessor functions. Thus, $\mathrm{ext}(\mathrm{Th}(M)) \geq \aleph_1$.

In their paper [CS13, Question 1.13], Chernikov and Simon posed the following question:

Is it true that $\mathrm{ext}(T) \leq \aleph_1$ whenever T is NIP?

In this paper we positively answer this question (see Main Theorem 1.1). We use the existence of honest definitions (see Definition 2.3 and Fact 2.4). Let

Bays was partially supported by DFG EXC 2044–390685587 and ANR-DFG AAPG2019 (Geomod). Ben-Neria was partially supported by the Israel Science Foundation, Grant 1832/19. Kaplan would like to thank the Israel Science Foundation (ISF) for their support of this research (grants no. 1254/18 and 804/22). Simon was partially supported by the NSF (grants no. 1665491 and 1848562).

$X = \phi(M, c)$ be externally definable and uncountable and let $\psi(x, z)$ be an honest definition for $\mathrm{tp}_{\phi^{\mathrm{opp}}}(c/M)$. This means that for every finite set $X_0 \subseteq X$, there is some $d \in M^z$ such that

$$X_0 = \phi(X_0, c) \subseteq \psi(M, d) \subseteq \phi(M, c) = X. \tag{*}$$

If one of these sets $Y_d := \psi(M, d)$ is infinite we are done, so assume for all d as in (*), Y_d is finite. We get a family of finite subsets of X which is cofinal as a subset of the partial order $\mathcal{P}^{<\omega}(X)$ of finite subsets of X. This raises the question:

Question 1.4. Suppose $\mathcal{F}$ is a cofinal family of finite subsets of $\aleph_1$. Can $\mathcal{F}$ have finite VC-dimension?

In other words, can the relation $\in|_{(\aleph_1 \times \mathcal{F})}$ be NIP?

In Theorem 3.8 we give a positive answer to Question 1.4. This means that the fact that the honest definition is NIP is not in itself a guarantee that X has an infinite M-definable subset (see Remark 5.2).

On the other hand, we prove that if $\mathcal{F}$ is a cofinal family of finite subsets of $\aleph_1$, then the two-sorted structure $(\aleph_1, \mathcal{F}; \in)$ has IP. We conclude (in Theorem 5.1):

Main Theorem 1.1. *For T NIP, ext$(T) \leq \aleph_1$.*

1.1. **A generalisation to arbitrary cardinals.** We also consider the following generalisation of Question 1.1.

Question 1.5. Let T be a theory and κ an infinite cardinal. Is there some cardinal λ such that for any $M \vDash T$ and any externally definable set $X \subseteq M^k$, if $|X| \geq \lambda$ then X contains an M-definable subset of size $\geq \kappa$?

For a complete theory T, let $\mathrm{ext}(T, \kappa)$ be the minimal λ as in Question 1.5 (if it does not exist, let $\mathrm{ext}(T, \kappa) = \infty$). So $\mathrm{ext}(T) = \mathrm{ext}(T, \aleph_0)$. The proof of [CS13, Corollary 1.12] can easily be adapted to show that if T is NIP then $\mathrm{ext}(T, \kappa) \leq \beth_\omega(\kappa)$.

The following slight adaptation of Example 1.3 gives us an NIP theory T (namely DLO) with $\mathrm{ext}(T, \kappa) \geq \kappa^+$ for $\kappa \geq \aleph_1$. Let I be an extension of the linear order $(\kappa, <)$ where between any two ordinals we put a copy of $\mathbb{Q}$. Let $M = \mathbb{Q} + I + \mathbb{Q}$. Then $\mathcal{M}$ is a dense linear order and thus $\mathrm{Th}(\mathcal{M})$ has quantifier elimination. The set I is externally definable but contains no M-definable subset of size κ.

In fact we prove the main theorem, Theorem 5.1, in this generality: if T is NIP then $\mathrm{ext}(T, \kappa) \leq \kappa^+$.

1.2. **Structure of the paper.** In Section 2, we give the necessary preliminaries on NIP and honest definitions. In Section 3 we discuss Question 1.4. In Section 4 we prove the technical lemmas needed to prove Theorem 5.1, which is proven in Section 5 and supplemented by some open questions.

In a previous version of this paper there was a mistake in the proof of Theorem 3.8 (pointed out to us by George Peterzil). The old proof involved the construction of well orders of order type ω on countable ordinals which agree up to finite sets. Since this result may be of independent interest, we put it in Appendix A.

1.3. **Acknowledgements.** The work on this project started while the third author was a visiting researcher at the 2021 Thematic Program on Trends in Pure and Applied Model Theory at the Fields Institute, and he would like to thank the Fields Institute for their hospitality.

We would like to thank George Peterzil for pointing out an error in the proof of Theorem 3.8 that appeared in a previous version. We would also like to thank the anonymous referee for their meticulous reading and useful comments.

2. Preliminaries

2.1. **Notations.** Our notation is standard. We use $\mathcal{L}$ to denote a first order language and $\phi(x,y)$ to denote a formula ϕ with a partition of (perhaps a superset of) its free variables. Let ϕ^{opp} be the partitioned formula $\phi(y,x)$ (it is the same formula with the partition reversed).

T will denote a complete theory in $\mathcal{L}$, and $\mathcal{U} \vDash T$ will be a monster model (a sufficiently large saturated model).

When x is a tuple of variables and A is a set contained in some structure (perhaps in a collection of sorts), we write A^x to denote the tuples of the sort of x (and of length $|x|$) of elements from A; alternatively, one may think of A^x as the set of assignments of the variables x to A. If M is a structure and $A \subseteq M^x$, $b \in M^y$, then $\phi(A,b) = \{a \in A \mid M \vDash \phi(a,b)\}$.

When $B \subseteq \mathcal{U}$, $\mathcal{L}(B)$ is the language $\mathcal{L}$ augmented with constants for elements from B so that a set is B-definable if it is definable in $\mathcal{L}(B)$.

For an $\mathcal{L}$-formula $\phi(x,y)$, an *instance* of ϕ over $B \subseteq \mathcal{U}$ is a formula $\phi(x,b)$ where $b \in B^y$, and a (complete) ϕ-*type* over B is a maximal partial type consisting of instances and negations of instances of ϕ over B. We write $S_\phi(B)$ for the space of ϕ-types over B in x (in this notation we keep in mind the partition (x,y), and x is the first tuple there). We also use the notation $\phi^1 = \phi$ and $\phi^0 = \neg\phi$. For $a \in \mathcal{U}^x$, we write $\mathrm{tp}_\phi(a/B) \in S_\phi(B)$ for its ϕ-type over B.

2.2. **VC-dimension and NIP.**

Definition 2.1 (VC-dimension)**.** Let X be a set and $\mathcal{F} \subseteq \mathcal{P}(X)$. The pair $(X,\mathcal{F})$ is called a **set system**. We say that $A \subseteq X$ is **shattered** by $\mathcal{F}$ if for every $S \subseteq A$ there is $F \in \mathcal{F}$ such that $F \cap A = S$. A family $\mathcal{F}$ is said to be a **VC-class** on X 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** $\mathcal{F}$, denoted by $\mathrm{VC}(\mathcal{F})$, is the smallest integer n such that no subset of X of size $n+1$ is shattered by $\mathcal{F}$.

If no such n exists, we write $\mathrm{VC}(\mathcal{F}) = \infty$.

Definition 2.2. Suppose T is an $\mathcal{L}$-theory and $\phi(x,y)$ is a formula. Say $\phi(x,y)$ is **NIP** if for some/every $M \vDash T$, the family $\{\phi(M^x,a) \mid a \in M^y\}$ is a VC-class. Otherwise, ϕ is **IP** (IP stands for "Independence Property" while NIP stands for "Not IP").

The theory T is **NIP** if all formulas are NIP. A structure M is **NIP** if $\mathrm{Th}(M)$ is NIP.

Definition 2.3. [Sim15, Definition 3.16 and Remark 3.14] Suppose T is an $\mathcal{L}$-theory and $M \vDash T$. Suppose that $\phi(x,y)$ is a formula, $A \subseteq M^x$ is some set and $b \in \mathcal{U}^y$. Say that an $\mathcal{L}$-formula $\psi(x,z)$ (with z a tuple of variables each of the same sort as x) is an **honest definition** of $\mathrm{tp}_{\phi^{\mathrm{opp}}}(b/A)$ if for every finite $A_0 \subseteq A$ there is some $c \in A^z$ such

$$\phi(A_0,b) \subseteq \psi(A,c) \subseteq \phi(A,b).$$

In other words, for all $a \in A$, if $\psi(a,c)$ holds then so does $\phi(a,b)$ and for all $a \in A_0$ the other direction holds: if $\phi(a,b)$ holds then $\psi(a,c)$ holds.

The existence of honest definitions for NIP theories was first proved in [CS13]. This was improved in [CS15] to get *uniformity* of the honest definitions assuming that T is NIP. This was subsequently improved to:

Fact 2.4. *[BKS21, Corollary 5.23] If $\phi(x,y)$ is NIP then there is a formula $\psi(x,z)$ that serves as an honest definition for any ϕ^{opp}-type over any set A of size ≥ 2.*

We also recall the Shelah expansion.

Definition 2.5. For a structure M, the **Shelah expansion** M^{Sh} of M is given by: for any formula $\phi(x,y)$ and any $b \in \mathcal{U}^y$, add a new relation $R_{\phi(x,b)}(x)$ interpreted as $\phi(M,b)$.

Fact 2.6. *[She09] If T is NIP then for any $M \vDash T$, M^{Sh} is NIP.*

3. The VC-dimension of cofinal families of finite subsets of an uncountable set

The goal of this section is to answer Question 1.4.

Definition 3.1. We say that a set $\mathcal{F}$ of subsets of a set X is **ω-cofinal** if every finite subset of X is contained in some element of $\mathcal{F}$. (In the case that $\mathcal{F}$ consists of finite subsets of X, we omit "ω-".)

We start with an easy observation.

Remark 3.2. If $\mathcal{F}$ is an ω-cofinal family of subsets of an infinite set X such that if $s \in \mathcal{F}$ then $|s| < |X|$, then $\in |_{(X \times \mathcal{F})}$ is **unstable**: there exist $(x_i, s_i)_{i \in \omega}$ such that $x_i \in s_j$ iff $i \leq j$. Indeed, inductively choose $x_i \in X$ and $s_i \in \mathcal{F}$ such that $x_i \notin \bigcup_{j<i} s_j$ and s_i contains $\{x_j \mid j \leq i\}$.

The proof of [CS13, Corollary 1.12(2)] can be adapted to say that if $|X| \geq \beth_\omega$ and $\mathcal{F}$ is a cofinal family of finite subsets of X then $\mathcal{F}$ is not a VC-class (i.e., it has IP). In fact, one can also make a connection between the VC-dimension of $\mathcal{F}$ and the cardinality of X (via the alternation rank of the appropriate relation). The next proposition replaces $\beth_\omega$ with $\aleph_\omega$, and gives a precise lower bound on the VC-dimension in terms of the cardinality of X.

Proposition 3.3. *If $|X| \geq \aleph_n$, then any cofinal system $\mathcal{F}$ of finite subsets of X has VC-dimension $> n$. So any cofinal set system of finite sets on a set of size $\geq \aleph_\omega$ has IP.*

Proof. We may assume $X = \aleph_n$.

For finite subsets $A, B \subseteq X$, write $A \vdash B$ to mean that if $D \in \mathcal{F}$ contains A then $D \cap B \neq \emptyset$, and write $A \nvdash B$ for the negation of this.

Observation 3.4. We have $A \nvdash \emptyset$ for any A, since $\mathcal{F}$ is cofinal.

Observation 3.5. If $A \nvdash B$, then there are only finitely many $c \in X$ such that $A \vdash B \cup \{c\}$. (Indeed, if $D \in \mathcal{F}$ witnesses $A \nvdash B$, then we must have $c \in D$.)

Observation 3.6. If $A' \subseteq A$ and $A \nvdash B$ then $A' \nvdash B$.

We find $c_i \in \aleph_i$ for $0 \leq i \leq n$ by downwards induction such that for $k = n, ..., 0, -1$:

$$(+)_k \begin{cases} \text{for any partition } c_{>k} = A \cup B \text{ where } A, B \text{ are disjoint and any} \\ b_i \in \aleph_i \text{ for } i \leq k, \\ \qquad\qquad b_{\leq k} \cup A \nvdash B. \end{cases}$$

$(+)_n$ holds by Observation 3.4.

$(+)_{-1}$ means that $\{c_i \mid i < n+1\}$ is shattered, from which we conclude.

Suppose $(+)_k$ holds, we choose $c_k \in \aleph_k$ such that $(+)_{k-1}$ holds.

Such a c_k exists because for each of the $\aleph_{k-1}$ choices for $b_{<k}$ and A, there are only finitely many choices to rule out. More explicitly, for every choice of $b_{<k}$ as above, and any $A \subseteq c_{>k}$ such that $b_{<k} \cup A \nvdash c_{>k} \setminus A$, let $s_{b_{<k},A} = \{c \in \aleph_k \mid b_{<k} \cup A \vdash c_{>k} \setminus A \cup \{c\}\}$. By Observation 3.5, $s_{b_{<k},A}$ is finite for each such $b_{<k}, A$, and let $c_k \in \aleph_k \setminus (\bigcup \{s_{b_{<k},A} \mid b_{<k} \cup A \nvdash c_{>k} \setminus A\} \cup c_{>k})$.

$(+)_{k-1}$ holds: if $c_k \in A$ then we are done by induction, and otherwise $c_k \in B$ and this follows from Observation 3.6 and the choice of c_k. $\square$

Remark 3.7. With the same proof *mutatis mutandis* one can see that if $\mathcal{F}$ is an ω-cofinal family of subsets of X, each of size $< \aleph_\alpha$, and if $|X| \geq \aleph_{\alpha+n}$, then $\mathcal{F}$ has VC-dimension $> n$.

Theorem 3.8. *There is a cofinal family $\mathcal{F}$ of finite subsets of $\aleph_1$ of VC-dimension 2.*

Proof. Let $\delta \leq \omega_1$. Suppose that $\mathcal{C} = (<^\alpha)_{\alpha<\delta}$ is a sequence of linear orders, where $<^\alpha$ is a linear order on α. We define the following relation on triples $\alpha, \beta, \gamma < \delta$: $\alpha, \beta \vdash_{\mathcal{C}} \gamma$ iff $\beta, \gamma < \alpha$ and $\gamma <^\alpha \beta$. Say $B \subseteq \delta$ is $\vdash_{\mathcal{C}}$**-closed** if for any $\alpha, \beta \in B$, if $\alpha, \beta \vdash_{\mathcal{C}} \gamma$ then $\gamma \in B$.

We inductively define well-orders $<^\alpha$ on $\alpha < \omega_1$ such that
$(*)_\alpha$ any finite subset $A \subseteq \alpha$ extends to a finite subset $A \subseteq B \subseteq \alpha$ such that $B \cup \{\alpha\}$ is $\vdash_{\mathcal{C}_\alpha}$-closed for $\mathcal{C}_\alpha := (<^\beta)_{\beta\leq\alpha}$.

$(*)_0$ holds with $<^0$ the empty order.

Suppose $(*)_\alpha$ holds. Let $<^{\alpha+1}$ be the order obtained from $<^\alpha$ by putting α at the start: $<^{\alpha+1} = <^\alpha \cup \{(\alpha, \beta) \mid \beta < \alpha\}$. Let $A \subseteq \alpha + 1$. By $(*)_\alpha$, let $B \subseteq \alpha$ be a finite set containing $A \setminus \{\alpha\} \subseteq \alpha$ such that $B' := B \cup \{\alpha\}$ is $\vdash_{\mathcal{C}_\alpha}$-closed. Then it follows from the definition of $<^{\alpha+1}$ that also $B' \cup \{\alpha + 1\}$ is $\vdash_{\mathcal{C}_{\alpha+1}}$-closed. Since B' is finite and contains A, we conclude that $(*)_{\alpha+1}$ holds.

Suppose that $\eta < \omega_1$ is a limit ordinal and $(*)_\alpha$ holds for all $\alpha < \eta$. Note that for $\alpha < \beta < \eta$ and any $B \subseteq \alpha$, B is $\vdash_{\mathcal{C}_\alpha}$-closed iff B is $\vdash_{\mathcal{C}_\beta}$-closed.

Since η is countable, it follows that $\eta = \bigcup_{n\in\omega} S_n$ where for each $n < \omega$, S_n is finite, $S_n \subseteq S_{n+1}$ and S_n is $\vdash_{\mathcal{C}_\alpha}$-closed for any (some) $\alpha < \eta$ such that $S_n \subseteq \alpha$. (In the construction, given S_n, let $S'_n = S_n \cup \{\beta_n\}$ where $(\beta_n)_{n<\omega}$ enumerates η and let S_{n+1} be finite and $\vdash_{\mathcal{C}_\alpha}$-closed containing S'_n for $\alpha < \eta$ such that $S'_n \subseteq \alpha$.) We define $<^\eta$ to be of order type ω in such a way that each S_n is an initial segment. Then $(*)_\eta$ holds: if A is a finite subset of η, then A is contained in some S_n which is finite and $\vdash_{\mathcal{C}_\alpha}$-closed for any α large enough, and since S_n is an initial segment of $<^\eta$, $S_n \cup \{\eta\}$ is $\vdash_{\mathcal{C}_\eta}$-closed.

Finally, let $\mathcal{C} = (<^\alpha)_{\alpha<\omega_1}$ and $\vdash = \vdash_{\mathcal{C}}$. Let $\mathcal{F}$ be the family of finite subsets of ω_1 which are $\vdash$-closed. By the above construction, $\mathcal{F}$ is cofinal. As for any triple $\alpha_0, \alpha_1, \alpha_2 < \omega_1$ of distinct ordinals there is some permutation σ of 3 such that $\alpha_{\sigma(0)}, \alpha_{\sigma(1)} \vdash \alpha_{\sigma(2)}$, $\mathcal{F}$ does not shatter any set of size 3. □

Corollary 3.9. *The following statement is independent of ZFC: there is an NIP cofinal family of finite subsets of $2^{\aleph_0}$.*

Proof. On the one hand CH is consistent with ZFC (by Gödel's theorem, see e.g., [Jec03, Theorem 13.20]), and on the other hand it is consistent with ZFC that $\aleph_\omega < 2^{\aleph_0}$ (using Cohen forcing, see e.g., [Jec03, Chapter 15, "Cohen Reals"]). Thus, the statement follows from Proposition 3.3 and Theorem 3.8. □

Question 3.10. Is there a cofinal family of finite subsets of $\aleph_2$ of VC-dimension 3? More generally: is the bound in Proposition 3.3 tight, or can we improve $\aleph_\omega$ to a smaller cardinal?

4. NIP and cofinal families of finite subsets of an uncountable set

This section is devoted to proving the following theorem.

Theorem 4.1. *Suppose that κ is an infinite cardinal, $|X| \geq \kappa^+$, and $\mathcal{F}$ is an ω-cofinal family of subsets of X, each of size $< \kappa$. Then $(X, \mathcal{F}; \in)$ has IP (as a two-sorted structure whose only relation is $\in \subseteq X \times \mathcal{F}$).*

The proof relies upon the following lemma.

Lemma 4.2. *Let κ be any infinite cardinal. Assume that:*

(1) $|X| \geq \kappa^+$.
(2) $R \subseteq X^n$ *and* $1 \leq n$.
(3) *For every* $a_1, \ldots, a_{n-1} \in X$, $|\{a_0 \in X \mid R(a_0, a_1, \ldots, a_{n-1})\}| < \kappa$.
(4) *For every set* $A \subseteq X$ *of size* $|A| = n$, *for some* $a \in A$ *and some tuple* $\bar{a} \in (A \setminus a)^{n-1}$, $R(a, \bar{a})$ *holds.*

Then, there is some partition of $\{1, \ldots, n-1\}$ *into nonempty disjoint sets* u, v *such that, letting* $x := (x_i)_{i \in u \cup \{0\}}$ *and* $y := (x_i)_{i \in v}$, *the partitioned formula* $\phi(x, y) := R(x_0, x_1, \ldots, x_{n-1})$ *has IP.*

Remark 4.3. Lemma 4.2 does not hold if we replace κ^+ by κ in (1). Indeed, let $X = \omega$ and let $R(x, y) = (x < y)$. Then (2)–(4) hold for $\kappa = \aleph_0$ and $n = 2$ but R is NIP (by e.g., [Sim15, Proposition A.2]).

Remark 4.4. Note that conditions (1)–(4) imply that $n > 2$. If $n = 1$ then by (3), R defines a set of size $< \kappa$, but by (4), R contains X, contradicting (1). Suppose that $n = 2$ and for $a \in X$ let $s_a = \{b \in X \mid R(b, a)\}$. Let $X_0, X_1 \subseteq X$ be such that $X_0 \cap X_1 = \emptyset$, $|X_0| = \kappa$ and $|X_1| = \kappa^+$. Let $S = \bigcup\{s_a \mid a \in X_0\}$. As $|S| \leq \kappa$, there must be some $b \in X_1 \setminus S$. As $|s_b| < \kappa$, there must be some $a \in X_0 \setminus s_b$. Then $a \notin s_b$ and $b \notin s_a$, contradicting (4).

The following example shows that the conditions of Lemma 4.2 can hold when $n = 3$.

Example 4.5. Suppose that for each $\alpha < \omega_1$, $<^\alpha$ is a well order on α of order type ω. For $\alpha, \beta, \gamma < \omega_1$, let $R(\gamma, \beta, \alpha)$ hold iff $\gamma, \beta < \alpha$ and $\gamma <^\alpha \beta$. Then R satisfies the conditions of Lemma 4.2 with $\kappa = \aleph_0$.

Remark 4.6. In essence, the proof of Lemma 4.2 is an induction on n, with Remark 4.4 as the base case. However, we need to keep track of sets witnessing IP ($D^{k,j,\bar{c}}_{\bar{A}}$ in the proof below), which substantially complicates the proof.

Proof of Lemma 4.2. Assume not, i.e., that

(5) for any partition of $\{1, \ldots, n-1\}$ into nonempty disjoint sets u, v, letting $x := (x_i)_{i \in u \cup \{0\}}$ and $y := (x_i)_{i \in v}$ the partitioned formula $\phi(x, y) := R(x_0, x_1, \ldots, x_{n-1})$ is NIP.

Define $R' \subseteq X^n$ by $R'(a_0, \ldots, a_{n-1})$ iff for some tuple $\bar{a} \in \{a_1, \ldots, a_{n-1}\}^{n-1}$, $R(a_0, \bar{a})$ holds. Note that R' satisfies (2)–(5) above (it satisfies (5) as a finite disjunction of NIP relations; in fact (5) can now be simplified by saying that $R'(x_0, \ldots, x_{k-1}; x_k, \ldots, x_{n-1})$ is NIP for any $1 < k < n$). Thus, we can replace R with R' and assume in addition that

(6) For any tuple $\bar{a} \in X^{n-1}$ and for any permutation $\bar{a}'$ of $\bar{a}$, $R(X, \bar{a}) = R(X, \bar{a}')$.

For any nonempty set $t \subseteq X$ of size $\leq n-1$, let $s_t = R(X, \bar{a})$ where $\bar{a}$ is any enumeration of t of length $n-1$; this is well-defined by (6). We can then restate (4) as:

$\boxtimes$ For every t of size n, for some $a \in t$, $a \in s_{t \setminus \{a\}}$.

We may assume that $|X| = \kappa^+$, and even that $X = \kappa^+$. When we say that a subset of X is "cofinal" or "contains an end segment of some cofinal set", we mean with respect to the canonical order on κ^+. For a cofinal set $D \subseteq X$ and some property $P \subseteq X$, write $\forall^* x \in D\ P(x)$ to mean that P contains an end segment of D. By downwards induction on $k \in [2, n]$ (note that by Remark 4.4, $n > 2$, so this range for k makes sense), we will find:

(A) $m_k < \omega$, and
(B) a cofinal set $D^{k,j,\bar{c}}_{\bar{A}} \subseteq X$ for every $j \in [k, n)$ and $\bar{c} \in X^{j-k}$ and $\bar{A} \in \prod_{i \in [k,j]} \mathcal{P}(m_i)$,

such that $\boxplus_k$ and $\boxtimes_k$ below hold. To state these conditions, we first introduce some additional notation:

- For l, j with $k \le l \le j \le n$, let $M_l^j := \prod_{i \in [l,j]} \mathcal{P}(m_i)$. We denote elements of M_l^j by $(j+1-l)$-tuples $\bar{A}$. For $j < l$, set $M_l^j := \{\emptyset\}$.
- For $j \in [k,n]$, $t \subseteq X$ of size $k-1$, $\bar{c} \in X^{j-k}$, and $\bar{A} \in M_k^{j-1}$, define sets $s_t^{k,j,\bar{c},\bar{A}}$ as follows. We set $s_t^{k,n,\bar{c},\bar{A}} = s_{t \cup \bar{c}}$, and then define recursively for $j \in [k,n)$:
$$s_t^{k,j,\bar{c},\bar{A}} = \{a \in X \mid \exists A \subseteq m_j \ \forall^* c \in D_{\bar{A}A}^{k,j,\bar{c}} \ a \in s_t^{k,j+1,\bar{c}c,\bar{A}A}\}.$$
- For $t \subseteq X$ of size $k-1$, let $s_t^k = s_t^{k,k,\emptyset,\emptyset}$.

Now we can state the conditions to be satisfied by our inductive construction:

$\boxplus_k$ For every $j \in (k,n)$ and every $c\bar{c} \in X^{j-k}$ and $A\bar{A} \in M_k^j$, $D_{A\bar{A}}^{k,j,c\bar{c}} \subseteq D_{\bar{A}}^{k+1,j,\bar{c}}$ (for $k \ge n-1$, this condition holds trivially).

$\boxtimes_k$ For all $t \subseteq X$ of size $|t| = k$, for some $a \in t$, $a \in s_{t \setminus \{a\}}^k$.

Note that each $|s_t^{k,j,\bar{c},\bar{A}}| < \kappa$ by downwards induction on j: for $j = n$ this is clear, and suppose that $|s_t^{k,j+1,\bar{c}c,\bar{A}A}| < \kappa$ for all $c \in X$ and $A \subseteq m_j$. Towards a contradiction, assume that $s_t^{k,j,\bar{c},\bar{A}}$ contains a set F of size κ. We may assume that for some $A \subseteq m_j$ and all $a \in F$, $\forall^* c \in D_{\bar{A}A}^{k,j,\bar{c}} \ a \in s_t^{k,j+1,\bar{c}c,\bar{A}A}$. For any $a \in F$ there is an end segment F_a of $D_{\bar{A}A}^{k,j,\bar{c}}$ such that for any $c \in F_a$, $a \in s_t^{k,j+1,\bar{c}c,\bar{A}A}$. Since $D_{\bar{A}A}^{k,j,\bar{c}}$ is cofinal in κ^+ (which is a regular cardinal), $\bigcap_{a \in F} F_a$ contains an end segment of $D_{\bar{A}A}^{k,j,\bar{c}}$, and in particular is nonempty. Let $c \in \bigcap_{a \in F} F_a$. Then $F \subseteq s_t^{k,j+1,\bar{c}c,\bar{A}A}$, contradicting the induction hypothesis.

Note also that for $t \subseteq X$ of size $n-1$, $s_t = s_t^n$.

We now proceed with the inductive construction of the m_k and $D_{\bar{A}}^{k,j,\bar{c}}$.

For $k = n$, let $m_n = 0$. Then $\boxplus_n$ holds trivially and $\boxtimes_n$ holds by $\boxtimes$ above.

Assume that $2 \le k < n$ and we found $m_{k'}$ and cofinal sets $D_{\bar{A}}^{k',j,\bar{c}}$ such that $\boxplus_{k'}$ and $\boxtimes_{k'}$ hold for all $k' > k$. We want to find m_k and sets $D_{\bar{A}}^{k,j,\bar{c}}$ such that $\boxplus_k$ and $\boxtimes_k$ hold.

For $m < \omega$ we let $\otimes_m$ be the following statement: there are

- cofinal sets $D_{\bar{A}}^{k,j,\bar{c}}$ for $j \in [k,n)$ and $\bar{c} \in X^{j-k}$ and $\bar{A} \in (\mathcal{P}(m) \times M_{k+1}^j)$, and
- subsets $t_i \subseteq X$ of size k for $i < m$,

such that:

(I) $\boxplus_k$ holds with m playing the role of m_k;

(II) if $B \ne A$ are subsets of m, and $i := \min(B \triangle A) \in B$, then for some enumeration $\bar{a}$ of t_i, the following hold:

$\oplus$ for all $c_k \in D_{(B)}^{k,k,\emptyset}$, for some $A_{k+1} \subseteq m_{k+1}$ and all $c_{k+1} \in D_{(B,A_{k+1})}^{k,k+1,(c_k)}$, $\dots$, for some $A_{n-1} \subseteq m_{n-1}$ and all c_{n-1} in $D_{(B,\dots,A_{n-1})}^{k,n-1,(c_k,\dots,c_{n-2})}$,
$$R(\bar{a}, c_k, \dots, c_{n-1});$$

$\ominus$ for all $c_k \in D_{(A)}^{k,k,\emptyset}$, for all $A_{k+1} \subseteq m_{k+1}$ and all $c_{k+1} \in D_{(A,A_{k+1})}^{k,k+1,(c_k)}$, $\dots$, for all $A_{n-1} \subseteq m_{n-1}$ and all c_{n-1} in $D_{(A,\dots,A_{n-1})}^{k,n-1,(c_k,\dots,c_{n-2})}$,
$$\neg R(\bar{a}, c_k, \dots, c_{n-1}).$$

If $\otimes_m$ holds for all $m < \omega$, we get IP as we now explain. Consider the formula
$$\phi(\bar{x}, \bar{y}) = \bigvee_{\bar{A} \in M_{k+1}^{n-1}} R(x_0, \dots, x_{k-1}; y_k, y_{k+1}^{\bar{A}}, y_{k+2}^{\bar{A}}, \dots, y_{n-1}^{\bar{A}}).$$

Fix some $m < \omega$. For $A \subseteq m$, we define a $\bar{y}$-tuple $\bar{c}^A$ as follows. Let $c_k^A \in D_{(A)}^{k,k,\emptyset}$ and for $j \in [k+1, n)$ and $(A_{k+1}, ..., A_{n-1}) \in M_{k+1}^{n-1}$, inductively let $c_j^{A,\bar{A}} \in D_{(A,A_{k+1},...,A_j)}^{k,j,(c_k^A, c_{k+1}^{A,\bar{A}},...,c_{j-1}^{A,\bar{A}})}$. Then, by (II) we get that if $B \neq A$ and $i = \min(B \triangle A) \in B$, then for some tuple $\bar{a}$ enumerating t_i, $\phi(\bar{a}, \bar{c}^B)$ holds, while $\phi(\bar{a}, \bar{c}^A)$ does not. Let E be the set of all $\bar{x}$-tuples $\bar{a}$ enumerating t_i for all $i < m$. We get that the number of ϕ-types in $\bar{y}$ over E is exponential in m (at least 2^m). However, $|E| \leq mk!$. By Sauer-Shelah ([Sim15, Lemma 6.4]), we get that $\phi(\bar{x}, \bar{y})$ has IP. As NIP formulas are closed under Boolean combinations, we get that $R(x_0, \ldots, x_{k-1}; y_k, \ldots, y_{n-1})$ has IP, contradicting (5).

We first show that $\otimes_0$ holds. Let $D_{(\emptyset)}^{k,k,\emptyset} = X$ and for $j > k$, $\bar{A} \in M_{k+1}^j$, $\bar{c} \in X^{j-(k+1)}$ and $c \in X$, let $D_{\emptyset\bar{A}}^{k,j,c\bar{c}} = D_{\bar{A}}^{k+1,j,\bar{c}}$. Then (I) is immediate, and (II) is trivially satisfied.

Let m_k be maximal such that $\otimes_{m_k}$ holds, witnessed by $D_{\bar{A}}^{k,j,\bar{c}}$ and t_i. We claim that this m_k and $D_{\bar{A}}^{k,j,\bar{c}}$ satisfy $\boxplus_k$ and $\boxtimes_k$. $\boxplus_k$ is satisfied by (I), so we are left to check $\boxtimes_k$.

Assume that $\boxtimes_k$ does not hold. Then there is some $t \subseteq X$ of size k witnessing this: for all $a \in t$, $a \notin s_{t\setminus\{a\}}^k$. We will show that letting $t_{m_k} := t$, we can find new $D_{\bar{A}}^{k,j,\bar{c}}$ for $\bar{A} \in \mathcal{P}(m_k+1) \times M_{k+1}^j$ and $\bar{c} \in X^{j-k}$ witnessing $\otimes_{m_k+1}$. We will construct two sequences of cofinal sets, $E_{\bar{A}}^{k,j,\bar{c}}$ and $F_{\bar{A}}^{k,j,\bar{c}}$, that will then be used to find suitable D's.

Let $A_k \subseteq m_k$. Let $E_{(A_k)}^{k,k,\emptyset} = D_{(A_k)}^{k,k,\emptyset} \setminus (s_t^{k+1} \cup t)$. Since s_t^{k+1} has size $< \kappa$, $E_{(A_k)}^{k,k,\emptyset}$ is still cofinal. Let $c_k \in E_{(A_k)}^{k,k,\emptyset}$. By $\boxtimes_{k+1}$ applied to $t \cup \{c_k\}$, and as $c_k \notin s_t^{k+1}$, for some $a_{c_k,A_k} \in t$, we have $a_{c_k,A_k} \in s_{\{c_k\}\cup t\setminus\{a_{c_k,A_k}\}}^{k+1}$. As t is finite, by reducing $E_{(A_k)}^{k,k,\emptyset}$, we may assume that there is some $a_{A_k} \in t$ such that $a_{A_k} \in s_{\{c_k\}\cup t\setminus\{a_{A_k}\}}^{k+1}$ for any $c_k \in E_{(A_k)}^{k,k,\emptyset}$.

Let $j \in (k, n)$, $\bar{A} \in M_k^j$, and $\bar{c} \in X^{j-k}$. Write $\bar{c} = c_k\bar{c}'$ and $\bar{A} = A_k\bar{A}'$, so $\bar{A}' \in M_{k+1}^j$. We define cofinal sets $E_{\bar{A}}^{k,j,\bar{c}}$ as follows

- If $\forall^* c \in D_{\bar{A}'}^{k+1,j,\bar{c}'}$ $a_{A_k} \in s_{\{c_k\}\cup t\setminus\{a_{A_k}\}}^{k+1,j+1,\bar{c}'c,\bar{A}'}$ then let $S \subseteq D_{\bar{A}'}^{k+1,j,\bar{c}'}$ be an end segment witnessing this, and set $E_{\bar{A}}^{k,j,\bar{c}} = S \cap D_{\bar{A}}^{k,j,\bar{c}}$. Note that $E_{\bar{A}}^{k,j,\bar{c}}$ is cofinal as $D_{\bar{A}}^{k,j,\bar{c}} \subseteq D_{\bar{A}'}^{k+1,j,\bar{c}'}$ by $\boxplus_k$.
- Otherwise, let $E_{\bar{A}}^{k,j,\bar{c}} = D_{\bar{A}}^{k,j,\bar{c}}$.

By (upwards) induction on $j \in [k, n)$ one proves that:

($\dagger_j$) For any $A_k \subseteq m_k$ and any $c_k \in E_{(A_k)}^{k,k,\emptyset}$ there is some $A_{k+1} \subseteq m_{k+1}$ such that for any $c_{k+1} \in E_{(A_k,A_{k+1})}^{k,k+1,(c_k)}$ there is some $A_{k+2} \subseteq m_{k+2}$ such that ... for any $c_j \in E_{(A_k,...,A_j)}^{k,j,(c_k,...,c_{j-1})}$, $a_{A_k} \in s_{\{c_k\}\cup t\setminus\{a_{A_k}\}}^{k+1,j+1,(c_{k+1},...,c_j),(A_{k+1},...,A_j)}$.

Now for $A_k \subseteq m_k$, let $F_{(A_k)}^{k,k,\emptyset} \subseteq D_{(A_k)}^{k,k,\emptyset}$ be a cofinal set such that $a_{A_k} \notin s_{t\setminus\{a_{A_k}\}}^{k,k+1,(c),(A_k)}$ for any $c \in F_{(A_k)}^{k,k,\emptyset}$; such a set exists since $a_{A_k} \notin s_{t\setminus\{a_{A_k}\}}^k = s_{t\setminus\{a_{A_k}\}}^{k,k,\emptyset,\emptyset}$.

Then for any $j \in (k, n)$, any $\bar{c} \in X^{j-k}$ and any $\bar{A} = (A_k, ..., A_j) \in M_k^j$, we similarly define cofinal sets $F_{\bar{A}}^{k,j,\bar{c}}$ as follows.

- If $a_{A_k} \in s_{t\setminus\{a_{A_k}\}}^{k,j,\bar{c},(A_k,...,A_{j-1})}$, let $F_{\bar{A}}^{k,j,\bar{c}} = D_{\bar{A}}^{k,j,\bar{c}}$, and
- if $a_{A_k} \notin s_{t\setminus\{a_{A_k}\}}^{k,j,\bar{c},(A_k,...,A_{j-1})}$, let $F_{\bar{A}}^{k,j,\bar{c}} \subseteq D_{\bar{A}}^{k,j,\bar{c}}$ be cofinal such that $a_{A_k} \notin s_{t\setminus\{a_{A_k}\}}^{k,j+1,\bar{c}c,\bar{A}}$ for any $c \in F_{\bar{A}}^{k,j,\bar{c}}$.

Recall that by choice of t, $a \notin s^k_{t\setminus\{a\}}$ for all $a \in t$. By (upwards) induction on $j \in [k, n]$ one proves that:

($\$_j$) If $(A_k, ..., A_{j-1}) \in M^{j-1}_k$ and $(c_k, \dots, c_{j-1}) \in X^{j-k}$ are such that $c_i \in F^{k,i,(c_k,\dots,c_{i-1})}_{(A_k,\dots,A_i)}$ for every $i \in [k, j)$, then for every $i \in [k, j]$,

$$a_{A_k} \notin s^{k,i,(c_k,\dots,c_{i-1}),(A_k,\dots,A_{i-1})}_{t\setminus\{a_{A_k}\}}.$$

Now, for any $A\bar{A} \in \mathcal{P}(m_k + 1) \times M^j_{k+1}$, let $G^{k,j,\bar{c}}_{A\bar{A}} := F^{k,j,\bar{c}}_{A\bar{A}}$ if $m_k \notin A$, else let $G^{k,j,\bar{c}}_{A\bar{A}} := E^{k,j,\bar{c}}_{(A\cap m_k)\bar{A}}$. Now we show that $(t_i)_{i\leq m_k}$ and these $G^{k,j,\bar{c}}_{A\bar{A}}$ witness $\otimes_{m_k+1}$. Note that $G^{k,j,\bar{c}}_{A\bar{A}}$ is a cofinal subset of $D^{k,j,\bar{c}}_{(A\cap m_k)\bar{A}}$. Hence $\boxplus_k$ still holds, establishing (I). For (II), let $B \neq A$ be subsets of $m_k + 1$. If $i = \min(B \triangle A) \in B$ and $i < m_k$, then $i = \min((B \cap m_k) \triangle (A \cap m_k))$ and so $\oplus$ and $\ominus$ still hold (using $G^{k,j,\bar{c}}_{A\bar{A}} \subseteq D^{k,j,\bar{c}}_{(A\cap m_k)\bar{A}}$). If not, then $B = A \cup \{m_k\}$. Let $\bar{a}$ be an enumeration of t_{m_k} starting with a_A. Then $\oplus$ follows from $(\dagger_{n-1})$ and $\ominus$ follows from $(\$_n)$, as required.

This completes the construction of m_k and $D^{k,j,\bar{c}}_{\bar{A}}$ as in (A), (B) above for all $k \in [2, n]$.

Finally, $\boxtimes_2$ yields a contradiction by the argument of Remark 4.4. Indeed, by $\boxtimes_2$ we get that for any distinct $a, b \in X$, either $a \in s^2_{\{b\}}$ or $b \in s^2_{\{a\}}$. Let $X_0, X_1 \subseteq X$ be such that $X_0 \cap X_1 = \emptyset$, $|X_0| = \kappa$ and $|X_1| = \kappa^+$. Let $S = \bigcup\{s^2_{\{a\}} \mid a \in X_0\}$. As $|S| \leq \kappa$, there must be some $b \in X_1 \setminus S$. As $|s^2_{\{b\}}| < \kappa$, there must be some $a \in X_0 \setminus s^2_{\{b\}}$. But then $a \notin s^2_{\{b\}}$ and $b \notin s^2_{\{a\}}$ — contradiction. □

Proof of Theorem 4.1. Suppose that $|X| \geq \kappa^+$ and that $\mathcal{F}$ is a cofinal family of subsets of X, each of size $< \kappa$. Suppose that $\mathrm{VC}(\mathcal{F}) = n$.

For any $0 \leq k \leq n$ and any $m \leq k$, let $R_{m,k}(x_0, \dots, x_k)$ be the relation defined by:

$$[\exists t \in \mathcal{F} \bigwedge_{1\leq i\leq k} (x_i \in t)^{(i\leq m)}] \wedge [\forall t \in \mathcal{F} ((\bigwedge_{1\leq i\leq k} (x_i \in t)^{(i\leq m)}) \to x_0 \in t)].$$

(If $k = 0$ the conjunction is empty and thus holds trivially, meaning that $R_{0,0}(x_0) = \forall t \in \mathcal{F}\ x_0 \in t$.)

Let $R(x_0, x_1, \dots, x_n) = \bigvee_{m\leq k\leq n} R_{m,k}(x_0, \dots, x_k)$. We claim that R satisfies the conditions of Lemma 4.2 on X. Conditions (1) and (2) are trivial, condition (3) follows from the fact that it is true for each m, k separately and that each $t \in \mathcal{F}$ has size $< \kappa$ (using the existential clause of the definition of $R_{m,k}$).

We show condition (4). Suppose that $A \subseteq X$ has size $n + 1$. Since $\mathrm{VC}(\mathcal{F}) = n$, $\mathcal{F}$ does not shatter A. Let $B \subseteq A$ be of minimal size such that $\mathcal{F}$ does not shatter B. Note that B is nonempty and let $k = |B| - 1$. Since B is not shattered, there is some $B_0 \subseteq B$ such that for no $t \in \mathcal{F}$, $t \cap B = B_0$. Note that $B_0 \neq B$ since $\mathcal{F}$ is ω-cofinal (and B is finite). Let $m = |B_0|$. Let $a_0 \in B \setminus B_0$ and let $a_1, \dots, a_k$ enumerate $B \setminus \{a_0\}$ such that $a_i \in B_0$ iff $i \leq m$. It follows that $R_{m,k}(a_0, a_1, \dots, a_k)$ holds: the first clause holds by the minimality of B (any proper subset is shattered), and the second clause follows by the choice of B_0.

By Lemma 4.2, for some permutation σ of $\{1, \dots, n\}$ and some $1 < k < n$ the partitioned formula $R(x_0, x_{\sigma(1)}, \dots, x_{\sigma(k-1)}; x_{\sigma(k)}, \dots, x_{\sigma(n-1)})$ has IP and we are done. □

Question 4.7. Let $R(x, y, z)$ be the relation from Example 4.5. The proof of Lemma 4.2 yields that $R(x, y; z)$ has IP. Could the relation $R(x; y, z)$ be NIP? Note that $\{R(\omega_1; \beta, \alpha) \mid \beta, \alpha\}$ is a cofinal family of finite subsets of ω_1 (see Theorem 3.8).

Similarly, we do not know whether the formula $\phi(x, z; y) = R(x, y, z)$ has IP.

5. Conclusion and final thoughts

We conclude with the final theorem, i.e., the generalisation of Main Theorem 1.1 discussed in Section 1.1.

Theorem 5.1. *Let κ be an infinite cardinal. If $T = \mathrm{Th}(M)$ is NIP and $X \subseteq M^k$ is externally definable of size $\geq \kappa^+$, then X contains an M-definable subset of size $\geq \kappa$. In other words, ext$(T, \kappa) \leq \kappa^+$.*

Proof. Suppose that X is defined by $\phi(x, c)$ for some formula $\phi(x, y)$. Let $\psi(x, z)$ be an honest definition for $\mathrm{tp}_{\phi^{\mathrm{opp}}}(c/M)$. This means that for every finite set $X_0 \subseteq X$, there is some $d \in M^z$ such that

$$X_0 = \phi(X_0, c) \subseteq \psi(M, d) \subseteq \phi(M, c) = X.$$

Let $Y = \{d \in M^z \mid \psi(M, d) \subseteq X\}$. Note that Y is definable in M^{Sh}. If for some $d \in Y$, $\psi(M, d)$ has size $\geq \kappa$, we are done, so assume for all $d \in Y$, $|\psi(M, d)| < \kappa$. Let $\mathcal{F} = \{\psi(M, d) \mid d \in Y\}$. Then $N := (X, \mathcal{F}; \in)$ is interpretable in M^{Sh}. By Fact 2.6 it follows that N is NIP, contradicting Theorem 4.1 as required. □

Remark 5.2. Note that the above proof implies that in an NIP theory, if $X = \phi(M, c)$ is externally definable of size κ^+, then an instance of the honest definition of $\mathrm{tp}_{\phi^{\mathrm{opp}}}(c/M)$ has size $\geq \kappa$. By Fact 2.4, we know that the existence of an honest definition $\psi(x, z)$ only requires ϕ to be NIP. We do not know if ψ itself can be chosen to be NIP (this is open even for the finite case, see [EK20, Question 22]). However, even if it were NIP, we cannot get a contradiction as in the proof above due to Theorem 3.8.

Question 5.3. Suppose that M is a structure and $X = \phi(M, c)$ is externally definable of size $\geq \aleph_1$. Suppose that ϕ is NIP. Does it follow that X contains an infinite definable subset?

Note that when T eliminates the quantifier $\exists^\infty$, the answer is "yes" (even just assuming that X is infinite), as in [CS13, Corollary 1.12(1)]: by Fact 2.4 (or [Sim15, Theorem 3.13]), $\mathrm{tp}_{\phi^{\mathrm{opp}}}(c/M)$ has an honest definition ψ, and so as above some instance $\psi(M, d) \subseteq X$ contains a finite subset X_0 large enough that, by elimination of $\exists^\infty$, the instance must be infinite.

Refining Question 5.3, we can define ext(T, ϕ, κ) as the minimal λ (if exists) such that whenever $M \models T$ and $X \subseteq M^k$ is externally definable by $\phi(x, c)$ for some $c \in \mathcal{U}$, then X contains an M-definable subset of size $\geq \kappa$. By Theorem 5.1, if T is NIP then ext$(T, \phi, \kappa) \leq \kappa^+$. If the honest definition of ϕ is NIP then by Remark 3.7, if $\kappa = \aleph_\alpha$, ext$(T, \phi, \kappa) \leq \aleph_{\alpha+\omega}$. If we assume only that ϕ is NIP, it is not even clear that ext$(T, \phi, \aleph_0)$ exists.

Question 5.4. What is ext(T, ϕ, κ) when ϕ is NIP?

Remark 5.5. Let T be a complete theory. Suppose that there is some infinite $\emptyset$-definable set Z of x-tuples such that $\phi(x, y)$ is **random** on Z: for any finite disjoint sets $A, B \subseteq Z$ there is some y-tuple d such that $A = \phi(A \cup B, d)$ (for the notations, see Section 2.1). This is a strong negation of NIP and happens e.g., in the case of the random graph. Then every subset of Z is externally definable by compactness. Let T^{Sk} be a Skolemization of T. Let λ be any infinite cardinal and let $I = (a_i)_{i<\lambda}$ be an indiscernible sequence (in the sense of T^{Sk}) contained in Z (in some model of T), and let $N = \mathrm{Sk}(I)$ (the Skolem hull of I). Then $X := \{a_i \mid i \text{ even}\}$ is a subset of N which is externally definable by an instance of ϕ, but which does not contain an infinite N-definable subset (even in $\mathcal{L}^{\mathrm{Sk}}$). Hence ext$(T, \phi, \aleph_0) = \infty$, and in particular ext$(T, \aleph_0) = \infty$.

Question 5.6. Does $\mathrm{ext}(T, \aleph_0) = \infty$ hold whenever T is IP? That is, does every IP theory have a model containing an uncountable externally definable set which contains no infinite definable set?

Appendix A. Almost agreeing orders on the countable ordinals

In this appendix, we show how to construct on each countable ordinal an order of order type ω, in such a way that any two of the orders agree up to a finite set. This result is not used in the paper. It formed part of our first attempt to prove Theorem 3.8, but in the end turned out not to provide a route to proving that theorem. We nonetheless present the result in this appendix, in the hope that it may be of interest in its own right.

Definition A.1. Let X be a set. Say two orders $<^1$ and $<^2$ on X **almost agree**, and write $<^1 \sim <^2$, if there is a finite subset $X_0 \subseteq X$ such that $<^1|_{(X\setminus X_0)} = <^2|_{(X\setminus X_0)}$.

Note that $\sim$ is an equivalence relation.

If $(X, <)$ has order type ω, we call $<$ an ω-order on X.

Theorem A.2. *There are ω-orders $<^\alpha$ on each α for $\omega \leq \alpha < \omega_1$ such that $<^\beta \sim <^\alpha|_\beta$ whenever $\omega \leq \beta < \alpha$.*

Before proving Theorem A.2, we establish a pair of lemmas.

Lemma A.3. *Suppose that $(X, <^X)$ and $(Y, <^Y)$ are both ω-orders, $X \subseteq Y$, and $<^X \sim <^Y|_X$. Then there is some ω-order $\lessdot^Y$ on Y such that $\lessdot^Y \sim <^Y$ and $\lessdot^Y|_X = <^X$.*

Proof. Let $X_0 \subseteq X$ be finite such that $<^X$ and $<^Y|_X$ agree on $X \setminus X_0$. We define an order on Y which agrees with $<^Y$ on $Y \setminus X_0$ and places X_0 in a way which agrees with $<^X$ on X. Formally, we prove the lemma by induction on $|X_0|$. If X_0 is empty there is nothing to do. Let $x \in X_0$, $Z = X \setminus \{x\}$, $W = Y \setminus \{x\}$, $<^Z = <^X|_Z$ and $<^W = <^Y|_W$. Note that $<^Z$ and $<^W$ are still ω-orders. By the induction hypothesis, there is some order $\lessdot^W$ on W such that $\lessdot^W \sim <^W$ and $<^Z \subseteq \lessdot^W$.

Let $F = \{y \in W \mid \exists z \in X\ (z <^X x \wedge y \trianglelefteq^W z)\}$; this is the cut on $(W, \lessdot^W)$ induced by the cut of x on $(Z, <^Z)$. Note that F is downwards closed in $\lessdot^W$ and that F is finite: let $x' \in X$ be such that $x <^X x'$. Then if $y \in F$ and z witnesses this, then $z <^X x <^X x'$ so that $z <^Z x'$ and hence $y \trianglelefteq^W z \lessdot^W x'$. But $(W, \lessdot^W)$ is an ω-order and hence $\{y \in W \mid y \lessdot^W x'\}$ is finite.

Let $\lessdot^Y$ extend $\lessdot^W$ and be such that for all $y \in Y$, $y \lessdot^Y x$ iff $y \in F$. To show that $\lessdot^Y$ is an ω-order, it is enough to show that $\{y \in Y \mid y <^Y x\}$ is finite, but this is precisely F. Since $F \cap X = \{z \in Z \mid z <^Z x\} = \{x' \in X \mid x' <^X x\}$, it follows that $\lessdot^Y \supseteq <^X$. Finally, if $\lessdot^W$ and $<^W$ agree on $W \setminus W_0$ where $W_0 \subseteq W$ is finite, then $<^Y$ and $\lessdot^Y$ agree on $Y \setminus (W_0 \cup \{x\})$ so that $<^Y \sim \lessdot^Y$ as required. $\square$

Lemma A.4. *Suppose that $(X_i)_{i<\omega}$ is an increasing sequence of countable sets, $(X_i, <^i)$ are ω-orders, and $<^{i+1}|_{X_i} \sim <^i$ for all $i < \omega$. Then there are ω-orders $\lessdot^i$ on X_i such that $\lessdot^0 = <^0$, and $\lessdot^i \sim <^i$, and $\lessdot^{i+1}|_{X_i} = \lessdot^i$ for all $i < \omega$.*

Proof. Inductively define $\lessdot^i$ as follows. Let $\lessdot^0 = <^0$. Suppose we defined $\lessdot^i$. Since $<^i \sim \lessdot^i$ and $<^{i+1}|_{X_i} \sim <^i$, it follows that $<^{i+1}|_{X_i} \sim \lessdot^i$. By Lemma A.3, there is some ω-order $\lessdot^{i+1}$ on X_{i+1} such that $\lessdot^{i+1} \sim <^{i+1}$ and $\lessdot^{i+1}|_{X_i} = \lessdot^i$, as required. $\square$

Proof of Theorem A.2. We define the orders $<^\alpha$ by induction on $\omega \leq \alpha < \omega_1$. Define $<^\omega$ as the canonical order on ω.

Suppose that $\alpha = \beta+1$ and $<^\beta$ has been defined. Let $<^\alpha = <^\beta \cup \{(\beta,\gamma) \mid \gamma < \beta\}$. In other words, we put β as the first element of $<^\alpha$ without changing anything else. Now, if $\gamma \le \beta$ then $<^\alpha|_\gamma = <^\beta|_\gamma \sim <^\gamma$ by induction.

Now suppose that $\alpha > \omega$ is a limit ordinal. Let $(\alpha_i)_{i<\omega}$ be an increasing sequence of ordinals, cofinal in α, where $\alpha_0 = \omega$. Apply Lemma A.4 to the sequence $(\alpha_i, <^{\alpha_i})$ (which we can by the induction hypothesis) to get ω-orders $\lessdot^{\alpha_i}$ on α_i such that $\lessdot^{\alpha_0} = <^{\alpha_0} = <^\omega$, and $\lessdot^{\alpha_i} \sim <^{\alpha_i}$, and $\lessdot^{\alpha_{i+1}} \supseteq \lessdot^{\alpha_i}$.

Let $<^* = \bigcup\{\lessdot^{\alpha_i} \mid i < \omega\}$. We define an ω-order $<^\alpha$ on α by, roughly speaking, concatenating the finite orders obtained by taking, for each $i < \omega$ in turn, those elements of α_i which are $\lessdot^{\alpha_i}$-less than i and have not yet been taken, ordered by $\lessdot^{\alpha_i}$. Formally: for $\beta < \alpha$, let $(-\infty,\beta)$ be $\{\gamma < \alpha \mid \gamma <^* \beta\}$. Define inductively sets $b_i \subseteq \alpha$ for $i < \omega$ as follows: $b_i = \alpha_i \cap (-\infty, i) \setminus \bigcup_{j<i} b_j$ (so $b_0 = \emptyset = (-\infty,0) \cap \alpha_0$, since $\lessdot^{\alpha_0} = <^\omega$ is the canonical order on ω). Note that $b_i \cap b_j = \emptyset$ for $i \ne j$ and that $\alpha = \bigcup_{i<\omega} b_i$ (if $\beta < \alpha$, then $\beta < \alpha_i$ for some $i < \omega$ and hence $\beta \lessdot^{\alpha_i} m$ for some $m < \omega$, so that $\beta \in (-\infty, m)$ and hence $\beta \in b_k$ for some $k \le \max\{i,m\}$). Finally, note that each b_i is finite since $(\alpha_i, \lessdot^{\alpha_i})$ has order type ω.

Order each b_i by $<^*$ and put the b_i's in order to define $<^\alpha$. More formally, let $<^{lex}$ be the lexicographical order on $\omega \times \alpha$ (taking the canonical order on ω and $<^*$ on α). For $\beta < \alpha$ let $i(\beta)$ be such that $\beta \in b_{i(\beta)}$, and for $\beta, \gamma < \alpha$ put $\gamma <^\alpha \beta$ iff $(i(\beta),\beta) <^{lex} (i(\gamma),\gamma)$.

We check that $<^\alpha$ is as required. The order type of $(\alpha, <^*)$ is ω since each b_i is finite, so that for any $\beta < \alpha$, $\{\gamma < \alpha \mid \gamma <^\alpha \beta\}$ is finite. Now suppose that $\beta < \alpha$. Then $\beta < \alpha_i$ for some $i < \omega$. Since $<^\beta \sim <^{\alpha_i}|_\beta$ and $<^{\alpha_i} \sim \lessdot^{\alpha_i}$, to show $<^\beta \sim <^\alpha|_\beta$ it suffices to check that $\lessdot^{\alpha_i} \sim <^\alpha|_{\alpha_i}$.

To show this we show that if $\gamma, \beta \in \alpha_i \setminus \bigcup_{j \le i} b_j$ then

$$(*) \qquad \gamma \lessdot^{\alpha_i} \beta \iff \gamma <^\alpha \beta.$$

Indeed, if $\gamma, \beta \in b_j$ for some $j < \omega$ then (*) follows from the fact that $<^\alpha$ equals $<^*$ on b_j, so extends $\lessdot^{\alpha_i}|_{b_j}$. Suppose that $\gamma \in b_j$, $\beta \in b_k$ and $j \ne k$, so without loss $i < j < k$. Then, since $\gamma, \beta \in \alpha_i$, necessarily $\gamma \in (-\infty, j)$ and $\beta \notin (-\infty, j)$. In this case, $\gamma \lessdot^{\alpha_i} j \lessdot^{\alpha_i} \beta$ (since this is true for $<^*$) and $\gamma <^\alpha \beta$ by definition. □

References

[BKS21] Martin Bays, Itay Kaplan, and Pierre Simon. Density of compressible types and some consequences, 2021.

[CS13] Artem Chernikov and Pierre Simon. Externally definable sets and dependent pairs. *Israel J. Math.*, 194(1):409–425, 2013.

[CS15] Artem Chernikov and Pierre Simon. Externally definable sets and dependent pairs II. *Trans. Amer. Math. Soc.*, 367(7):5217–5235, 2015.

[EK20] Shlomo Eshel and Itay Kaplan. On uniform definability of types over finite sets for NIP formulas. *Journal of Mathematical Logic*, 0(0):2150015, 2020.

[Jec03] Thomas Jech. *Set theory*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2003. The third millennium edition, revised and expanded.

[She09] Saharon Shelah. Dependent first order theories, continued. *Israel J. Math.*, 173:1–60, 2009.

[Sim15] Pierre Simon. *A guide to NIP theories*, volume 44 of *Lecture Notes in Logic*. Association for Symbolic Logic, Chicago, IL; Cambridge Scientific Publishers, Cambridge, 2015.

Institut für Mathematische Logik und Grundlagenforschung, Fachbereich Mathematik und Informatik, Universität Münster, Einsteinstrasse 62, 48149 Münster, Germany

Email address: mbays@sdf.org

Einstein Institute of Mathematics, Hebrew University of Jerusalem, 91904, Jerusalem Israel.

Email address: omer.bn@mail.huji.ac.il

Einstein Institute of Mathematics, Hebrew University of Jerusalem, 91904, Jerusalem Israel.

Email address: kaplan@math.huji.ac.il

Dept. of Mathematics, University of California, Berkeley, 970 Evans Hall #3840, Berkeley, CA 94720-3840 USA

Email address: simon@math.berkeley.edu