SEMI-EQUATIONAL THEORIES

ARTEM CHERNIKOV AND ALEX MENNEN

Abstract. We introduce and study (weakly) semi-equational theories, generalizing equationality in stable theories (in the sense of Srour) to the NIP context. In particular, we establish a connection to distality via one-sided strong honest definitions; demonstrate that certain trees are semi-equational, while algebraically closed valued fields are not weakly semi-equational; and obtain a general criterion for weak semi-equationality of an expansion of a distal structure by a new predicate.

§1. Introduction. Equations and equational theories were introduced by Srour [38–40] in order to distinguish "positive" information in an arbitrary first order theory, i.e., to find a well-behaved class of "closed" sets among the definable sets, by analogy to the algebraic sets among the constructible ones in algebraically closed fields. We recall the definition:

DEFINITION 1.1. (1) A partitioned formula $\varphi(x, y)$, with x, y tuples of variables, is an *equation* (with respect to a first-order theory T) if there do not exist $\mathcal{M} \models T$ and tuples $(a_i, b_i : i \in \omega)$ in $\mathcal{M}$ such that $\mathcal{M} \models \varphi(a_i, b_j)$ for all $j < i$ and $\mathcal{M} \models \neg\varphi(a_i, b_i)$ for all i.

(2) A theory T is *equational* if every formula $\varphi(x, y)$, with x, y arbitrary finite tuples of variables, is equivalent in T to a Boolean combination of finitely many equations $\varphi_1(x, y), \dots, \varphi_n(x, y)$.

It is immediate from the definition that every equational theory is stable. Structural properties of equational theories in relation to forking and stability theory are studied in [19–22, 33]. Many natural stable theories are equational; [19] provided the first example of a stable non-equational theory. More recently it was demonstrated that the stable theory of non-abelian free groups is not equational [28, 34], and further examples are constructed in [25]. It is demonstrated in [24] that all theories of separably closed fields are equational (generalizing earlier work of Srour [37]). See also [29] for an accessible introduction to equationality.

We propose a generalization of equations and equational theories to the larger class of NIP theories (see Section 1.2 for a more detailed discussion):

Received October 26, 2022.

2020 *Mathematics Subject Classification.* Primary 03C45.

Key words and phrases. semi-equational theories, equational theories, NIP, distality, classification theory, expansions by a predicate.

This version of the article was significantly shortened for the journal publication, resulting in some details being omitted. For the full version of the article see [7].

0022-4812/00/0000-0000
DOI:10.1017/jsl.2023.28

DEFINITION 1.2. Let T be a first-order theory and $\mathbb{M} \models T$ a monster model of T.

(1) A partitioned formula $\varphi(x, y)$ is a *semi-equation* (in T) if there is no sequence $(a_i, b_i : i \in \omega)$ with $a_i \in \mathbb{M}^x, b_i \in \mathbb{M}^y$ such that for all $i, j \in \omega$, $\models \varphi(a_i, b_j) \iff i \neq j$.

(2) A (partitioned) formula $\varphi(x, y)$ is a *weak semi-equation* if there are no $b \in \mathbb{M}^y$ and an ($\emptyset$-)indiscernible sequence $(a_i : i \in \mathbb{Z})$ with $a_i \in \mathbb{M}^x$ such that the subsequence $(a_i : i \neq 0)$ is indiscernible over b, $\models \varphi(a_i, b)$ for all $i \neq 0$, but $\models \neg\varphi(a_0, b)$.

(3) A theory T is (*weakly*) *semi-equational* if every formula $\varphi(x, y) \in \mathcal{L}$, with x, y arbitrary finite tuples of variables, is a Boolean combination of finitely many (weak) semi-equations $\psi_1(x, y), \ldots, \psi_n(x, y) \in \mathcal{L}$.

Semi-equations are in particular weak semi-equations, every weakly semi-equational theory is NIP, and in a stable theory all three notions coincide (see Proposition 2.10). Some parts of the basic theory of equations naturally generalize to (weak) semi-equations, but there are also some new phenomena and complications appearing outside of stability. In particular, weak semi-equationality provides a simultaneous generalization of equationality and distality, bringing out some curious parallels between those two notions (see Section 4). In this paper we develop the basic theory of (weak) semi-equations, and investigate (weak) semi-equationality in some examples. We view this as a first step, and a large number of questions remain open and can be found throughout the paper.

In Section 1.2 we provide some equivalent characterizations of (weak) semi-equationality in terms of indiscernibles. We discuss closure of (weak) semi-equations under Boolean combinations (Proposition 2.3), reducts and expansions (Proposition 2.6). In Section 2.2 we discuss how (weak) semi-equationality relates to the more familiar notions: all weakly semi-equational theories are NIP, distal theories are weakly semi-equational, and in a stable theory a formula is an equation if and only if it is a (weak) semi-equation (Proposition 2.10). In Section 2.3 we introduce some quantitive parameters associated with semi-equations. This parameter is related to breadth (Definition 2.15) of the family defined by the instances of a formula, and we observe that a formula is a semi-equation if and only if the family of its instances has finite breadth (Proposition 2.16). The case when this parameter is minimal, i.e., 1-semi-equations, provides a generalization of weakly normal formulas characterizing 1-based stable theories (Proposition 2.19). Hence 1-semi-equationality can be viewed as a form of “linearity,” or “1-basedness” for NIP theories. We discuss its connections to a different form of “linearity” considered in [5], namely basic relations and almost linear Zarankiewicz bounds (see Proposition 2.23 and Remark 2.24), observing that $(2, 1)$-semi-equational theories do not define infinite fields.

In Section 3 we consider some examples of semi-equational theories. In Section 3.1 we show that an o-minimal expansion of a group is linear if and only if it is $(2, 1)$-semi-equational. It remains open if the field of reals is semi-equational (Problem 3.4). We demonstrate that arbitrary unary expansions of linear orders (Section 3.2) and many ordered abelian groups (Section 3.4) are 1-semi-equational. In Section 3.5 we demonstrate that the theory of infinitely branching dense trees is semi-equational (Theorem 3.12), but not 1-semi-equational (even after naming

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

parameters, see Theorem 3.13 and Corollary 3.14). Semi-equationality of arbitrary trees remains open (Problem 3.17). In Section 3.3 we observe that dense circular orders are not semi-equational, but become 1-semi-equational after naming a single constant (in contrast to equationality being preserved under naming and forgetting constants).

In Section 4 we consider the relation of weak semi-equationality and distality in more detail. We show that in an NIP theory, weak semi-equationality of a formula is equivalent to the existence of a *one-sided* strong honest definition for it (Theorem 4.8). This is a simultaneous generalization of the existence of strong honest definitions in distal theories from [10] and the isolation property for the positive part of φ-types for equations (replacing a conjunction of finitely many instances of φ by some formula θ; see Fact 4.2).

In Section 5.2 we show that many theories of NIP valued fields with an infinite stable residue field, e.g., ACVF, are not weakly semi-equational (see Theorem 5.1 and Remark 5.10). In Section 5.1 we provide a sufficient criterion for when a formula is not a Boolean combination of weak semi-equations (generalizing the criterion for equations from [28]). We then apply it to show that the partitioned formula $\psi(x_1, x_2; y_1, y_2) := v(x_1 - y_1) < v(x_2 - y_2)$ is not a Boolean combination of weak semi-equations via a detailed analysis of the behavior of indiscernible sequences. It remains open if the field $\mathbb{Q}_p$ is semi-equational (Problem 5.12).

In Section 6 we consider preservation of weak semi-equationality in expansions by naming a new predicate, partially adapting a result for NIP from [9]. Namely, we demonstrate in Theorem 6.7 that if $\mathcal{M} \models T$ is distal, A is a subset of $\mathcal{M}$ with a distal induced structure and the pair (M, A) is almost model complete (i.e., every formula in the pair is equivalent to a Boolean combination of formulas which only quantify existentially over the predicate; see Definition 6.6), then the pair $(\mathcal{M}, A)$ is weakly semi-equational. This implies in particular that dense pairs of *o*-minimal structures are weakly semi-equational (but not distal by [18]).

§2. Semi-equations and their basic properties. Let T be a complete theory in a language $\mathcal{L}$, and we work inside a sufficiently saturated and homogeneous monster model $\mathbb{M} \models T$. All sequences of elements are assumed to be small relative to the saturation of $\mathbb{M}$, and we write $x, y, \ldots$ to denote finite tuples of variables. Given two linear orders I, J, $I + J$ denotes the linear order given by their sum (i.e., $I < J$), and (0) denotes a linear order with a single element. We write $\mathbb{N} = \{0, 1, \ldots\}$ and for $k \in \mathbb{N}$, $[k] = \{1, \ldots, k\}$. Given a partitioned formula $\varphi(x, y)$, we let $\varphi^*(y, x) := \varphi(x, y)$.

2.1. Some basic properties of (weak) semi-equations.

Remark 2.1. By Ramsey and compactness we may equivalently replace ω by an arbitrary infinite linear order in Definition 1.2(1), and $\mathbb{Z}$ by $I_L + (0) + I_R$ with I_L, I_R arbitrary infinite linear orders in Definition 1.2(2).

By Ramsey, compactness, and taking automorphisms we also have:

Proposition 2.2. *A formula $\varphi(x, y)$ is a semi-equation if and only if there are no b, infinite linear orders I_L, I_R, and an indiscernible sequence $(a_i)_{i \in I_L + (0) + I_R}$ such that $\models \varphi(a_i, b)$ for $i \in I_L + I_R$, but $\not\models \varphi(a_0, b)$.*

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

Proposition 2.3. (1) *If $\varphi(x, y)$ is a semi-equation, then $\varphi(x, y)$ is a weak semi-equation. Hence every semi-equational theory is weakly semi-equational.*

(2) *Semi-equations are closed under conjunctions and exchanging the roles of the variables.*

(3) *Weak semi-equations are closed under conjunctions and disjunctions.*

Proof. (1) Clear from definitions using Proposition 2.2.

(2) Suppose $\varphi(x, y) \wedge \psi(x, y)$ is not a semi-equation. By Proposition 2.2, there are b and an indiscernible sequence $(a_i)_{i\in\mathbb{Z}}$ such that $\models \varphi(a_i, b) \wedge \psi(a_i, b) \iff i \neq 0$. Either $\not\models \varphi(a_0, b)$, in which case $\varphi(x, y)$ is not a semi-equation, or $\not\models \psi(a_0, b)$, in which case $\psi(x, y)$ is not a semi-equation. And $\varphi(x, y)$ is a semi-equation if and only if $\varphi^*(y, x) := \varphi(x, y)$ is a semi-equation by the symmetry of the definition.

(3) For conjunctions, the same as the proof of (2), but with the stipulation that $(a_i)_{i\neq 0}$ is b-indiscernible added. Now suppose $\varphi(x, y) \vee \psi(x, y)$ is not a weak semi-equation. Then there are b and an indiscernible sequence $(a_i)_{i\in\mathbb{Z}}$ such that $(a_i)_{i\neq 0}$ is b-indiscernible, and $\models \varphi(a_i, b) \vee \psi(a_i, b) \iff i \neq 0$. Either $\models \varphi(a_1, b)$ or $\models \psi(a_1, b)$, and then, by b-indiscernibility, either $\models \varphi(a_i, b)$ for all $i \neq 0$ or $\models \psi(a_i, b)$ for all $i \neq 0$. In the first case, $\varphi(x, y)$ is not a weak semi-equation, and in the second case, $\psi(x, y)$ is not a weak semi-equation. ⊣

Remark 2.4. (1) To see that neither property is closed under negation, note that $x = y$ is a semi-equation (hence also a weak semi-equation), but $x \neq y$ is not a weak semi-equation in the theory of infinite sets.

(2) To see that semi-equations need not be closed under disjunction, note that in a linear order, $x < y$ and $y < x$ are both semi-equations, but their disjunction is equivalent to $x \neq y$, which is not.

Problem 2.5. Are weak semi-equations closed under exchanging the roles of the variables, at least in NIP theories? Fact 6.4 can be viewed as establishing this for the definition of distality; however, the proof is not sufficiently local with respect to a formula witnessing failure of distality.

Proposition 2.6. *Assume we are given languages $\mathcal{L} \subseteq \mathcal{L}'$, a complete $\mathcal{L}$-theory T and an $\mathcal{L}'$-theory T' with $T \subseteq T'$, and a formula $\varphi(x, y) \in \mathcal{L}$.*

(1) *The formula $\varphi(x, y)$ is a semi-equation in T if and only if it is in T'.*

(2) *If $\varphi(x, y)$ is a weak semi-equation in T, then it is a weak semi-equation in T'.*

Proof. (1) Left to right is immediate from the definition (Proposition 2.2). For the converse, assume that in some model of T we can find an infinite sequence $(a_i, b_i)_{i\in I}$ such that for all $i, j \in I$, $\models \varphi(a_i, b_j) \iff i \neq j$. By completeness of T, we can find arbitrarily long finite sequences with the same property in *every* model of T, in particular in some model of T'. By compactness we can thus find an infinite sequence with the same property in a model of T', demonstrating that $\varphi(x, y)$ is not a semi-equation in T'.

(2) If $\varphi(x, y) \in \mathcal{L}$ is not a weak semi-equation in T', then (in a monster model of T', and hence of T) there are b and an $\mathcal{L}'$-indiscernible $(a_i)_{i\in I_L+(0)+I_R}$ such that $(a_i)_{i\in I_L+I_R}$ is $\mathcal{L}'$-indiscernible over b and $\models \varphi(a_i, b)$ for $i \in I_L + I_R$, but $\not\models \varphi(a_0, b)$, for infinite linear orders I_L, I_R. Then, in particular, $(a_i)_{i\in I_L+(0)+I_R}$ is $\mathcal{L}$-indiscernible,

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

and $(a_i)_{i\in I_L+I_R}$ is $\mathcal{L}$-indiscernible over b, so $\varphi(x,y)$ is a not a weak semi-equation in T. $\dashv$

Remark 2.7. The converse to Proposition 2.6(2) does not hold. Let $T' := \mathrm{DLO}$ be the theory of dense linear orders, and T its reduct to $\mathcal{L} := \{=\}$. Then the $\mathcal{L}$-formula $x \neq y$ is not a weak semi-equation in T by inspection, but it is a weak semi-equation in T' since it is equivalent to a disjunction of weak semi-equations $(x<y)\vee(x>y)$ (Proposition 2.3).

Problem 2.8. Is weak semi-equationality of a theory preserved under reducts? This appear to be open already for equationality (see [20, Question 3.10]), and fails for semi-equationality (see Section 3.3).

Problem 2.9. Is (weak) semi-equationality of theories invariant under bi-interpretability without parameters? Equivalently, if T is (weakly) semi-equational, does it follow that so is T^{eq}?

2.2. Relationship to equations and NIP. We provide some evidence that semi-equationality can be naturally viewed as a generalization of equationality (in the sense of Srour) in stable theories to the NIP context.

Proposition 2.10. (1) *Weak semi-equations are NIP formulas; hence, weakly semi-equational theories are NIP.*
(2) *Equations are semi-equations.*
(3) *A formula is an equation if and only if it is both stable and a semi-equation.*
(4) *In a stable theory, all weak semi-equations are equations. In particular, a stable theory is equational if and only if it is (weakly) semi-equational.*

Proof. (1) If $\varphi(x,y)$ is not NIP, then there are an indiscernible sequence $(a_i)_{i\in\mathbb{N}}$ and b such that $\models \varphi(a_i,b) \iff i$ is even. For any finite set of formulas $\Delta(x_1,\dots,x_n,y)$, by Ramsey's theorem, there is an infinite $I \subseteq 2\mathbb{N}$ on which the truth value of all formulas in $\Delta\left(a_{i_1},\dots,a_{i_n},b\right)$ is constant for all $i_1<\cdots<i_n\in I$. Thus, by letting $a'_0 := a_i$ for some sufficiently large odd i, we can find an indiscernible sequence $(a'_i)_{i\in I_L+(0)+I_R}$ (using $I_L \sqcup I_R = I$, and $a'_i = a_i$ for $i\in I$) for some infinite I_R and arbitrarily large finite I_L, such that $(a'_i)_{i\in I_L+I_R}$ is Δ-indiscernible over b. By compactness, it follows that $\varphi(x,y)$ is not a weak semi-equation.

(2) If $\varphi(x,y)$ is not a semi-equation, then there is a sequence $(a_i,b_i)_{i\in\mathbb{N}}$ such that $\models\varphi(a_i,b_j) \iff i\neq j$. In particular, $\models\varphi(a_i,b_j)$ for all $j<i$, and $\not\models\varphi(a_i,b_i)$, so this is a counterexample to the descending chain condition, and $\varphi(x,y)$ is not an equation.

(3) If $\varphi(x,y)$ is not an equation, then by Ramsey and compactness there is an indiscernible sequence $(a_i,b_i)_{i\in\mathbb{N}}$ such that $\models\varphi(a_i,b_j)$ for all $j<i$, and $\not\models \varphi(a_i,b_i)$. If $\varphi(a_i,b_j)$ holds for $i<j$ then $\varphi(x,y)$ is not a semi-equation. Otherwise, $\varphi(x,y)$ is not stable.

(4) If $\varphi(x,y)$ is not an equation, by Ramsey and compactness we can choose an indiscernible sequence $(a_i,b_i)_{i\in\mathbb{Z}}$ such that $\models\varphi(a_i,b_j)$ for all $j<i$, and $\not\models \varphi(a_i,b_i)$. The indiscernible sequence $(a_i,b_i)_{i\in\mathbb{Z}}$ is totally indiscernible by stability of T; hence, we have $\models\varphi(a_i,b_0) \iff i\neq 0$, and also $(a_i : i\neq 0)$ is indiscernible over b_0. This shows that $\varphi(x,y)$ is not a weak semi-equation. $\dashv$

2.3. Weakly normal formulas, (k, n)-semi-equations, and breadth.

DEFINITION 2.11 (see [32, Chapter 4, Definition 1.1]). A formula $\varphi(x, y)$ is *k-weakly normal* if for every $b_1, \dots, b_k \in \mathbb{M}^y$ such that $\models \exists x\, \varphi(x, b_1) \wedge \cdots \wedge \varphi(x, b_k)$, there are some $i \neq j \in [k]$ such that $\models \forall x\, \varphi(x, b_i) \leftrightarrow \varphi(x, b_j)$. It is *weakly normal* if it is k-weakly normal for some k (by compactness this is equivalent to: an infinite collection of pairwise distinct instances of $\varphi(x, y)$ must have empty intersection).

A formula $\varphi(x, y)$ is *normal* in the sense of [31] if and only if it is 2-weakly normal. Weakly normal formulas are special kinds of equations characterizing "linearity" of forking in stable theories (see [32, Chapter 4, Proposition 1.5, Remark 1.8.4, and Lemma 1.9]):

FACT 2.12. *A stable theory T is 1-based if and only if in T, every formula $\varphi(x, y) \in \mathcal{L}$, with x, y arbitrary finite tuples of variables, is equivalent to a Boolean combination of finitely many weakly normal formulas $\psi_1(x, y), \dots, \psi_n(x, y) \in \mathcal{L}$.*

We introduce some numeric parameters characterizing semi-equations, minimal values of which give rise to a generalization of weak normality.

DEFINITION 2.13. For $k, n \in \mathbb{N}$, a formula $\varphi(x, y)$ is a *(k, n)-semi-equation* if, for every $b_1, \dots, b_k \in \mathbb{M}^y$, if $\models \exists x\ \varphi(x, b_1) \wedge \cdots \wedge \varphi(x, b_k)$, then for some pairwise distinct $i_1, \dots, i_n, j \in [k]$, $\models \forall x\ \left(\varphi\left(x, b_{i_1}\right) \wedge \cdots \wedge \varphi(x, b_{i_n})\right) \rightarrow \varphi(x, b_j)$. And $\varphi(x, y)$ is an *n-semi-equation* if it is a (k, n)-semi-equation for some k. A theory T is *n-semi-equational* (respectively, *(k, n)-semi-equational*) if every formula $\varphi(x, y) \in \mathcal{L}$, with x, y arbitrary finite tuples of variables, is equivalent in T to a Boolean combination of n-semi-equations (respectively, (k, n)-semi-equations) $\psi_1(x, y), \dots, \psi_n(x, y) \in \mathcal{L}$.

PROPOSITION 2.14. (1) *If $\varphi(x, y)$ is a (k, n)-semi-equation, then $n < k$, and $\varphi(x, y)$ is also an (ℓ, m)-semi-equation for any $\ell \geq k$ and $n \leq m < \ell$. If $\varphi(x, y)$ is an n-semi-equation, then it is also an m-semi-equation for every $m \geq n$.*

(2) *A formula is a semi-equation if and only if it is an n-semi-equation for some n, if and only if it is an $(n, n-1)$-semi-equation for some n.*

PROOF. (1) Clear from the definitions.

(2) If $\varphi(x, y)$ is not a semi-equation, let $(a_i, b_i)_{i \in \mathbb{N}}$ be such that $\models \varphi(a_i, b_j) \iff i \neq j$. Then for any (k, n) we have $\models \varphi(a_0, b_1) \wedge \cdots \wedge \varphi(a_0, b_k)$, but for any pairwise distinct $i_1, \dots, i_n, j \in [k]$, $\models \varphi\left(a_j, b_{i_1}\right) \wedge \cdots \wedge \varphi(a_j, b_{i_n}) \wedge \neg\varphi(a_j, b_j)$; hence, $\varphi(x, y)$ is not a (k, n)-semi-equation. Conversely, for any $k \in \mathbb{N}$, if $\varphi(x, y)$ is not a $(k, k-1)$-semi-equation, then there exist $b_1, \dots, b_k$ such that for each $j \in [k]$, there is a_j such that $\models \varphi(a_j, b_i)$ for $i \neq j$, but $\not\models \varphi(a_j, b_j)$. Hence if $\varphi(x, y)$ is not a $(k, k-1)$-semi-equation for any k, then by compactness $\varphi(x, y)$ is not a semi-equation. And if $\varphi(x, y)$ is not an n-semi-equation, then it is not an $(n+1, n)$-semi-equation by definition, so a formula that is not an n-semi-equation for any n is also not a $(k, k-1)$-semi-equation for any k. ⊣

We recall the notion of breadth from lattice theory.

DEFINITION 2.15 [2, Section 2.4]. Given a set X and $d \in \mathbb{N}_{\geq 1}$, a family of subsets $\mathcal{F} \subseteq \mathcal{P}(X)$ has *breadth d* if any nonempty intersection of finitely many sets in $\mathcal{F}$ is the intersection of at most d of them, and d is minimal with this property.

Proposition 2.16. *A formula $\varphi(x, y)$ is a $(k+1, k)$-semi-equation if and only if the family of sets $\mathcal{F}_\varphi := \{\varphi(\mathbb{M}, b) \mid b \in \mathbb{M}^y\}$ has breadth at most k. In particular, $\varphi(x, y)$ is a semi-equation if and only if the family of sets $\mathcal{F}_\varphi$ has finite breadth.*

Proof. The family of sets $\{\varphi(\mathbb{M}, b) \mid b \in \mathbb{M}^y\}$ has breadth at most k if and only if every finite consistent conjunction of instances of φ is implied by the conjunction of at most k of those instances. In particular this applies to consistent conjunctions of $(k+1)$ instances of φ, showing that if the breadth of $\mathcal{F}_\varphi$ is $\leq k$, then it is a $(k+1, k)$-semi-equation. Conversely, assume $\varphi(x, y)$ is a $(k+1, k)$-semi-equation. Given any consistent conjunction of $n > k$ instances of φ, any $(k+1)$ of them contain an instance implied by the other k instances. Removing this implied instance, we reduce to the case of a consistent conjunction of $(n-1)$ instances, and after $(n-k)$ steps to a conjunction of k instances of φ implying all the other ones. The "in particular" part is Proposition 2.14(2). $\dashv$

Example 2.17. Let T be an NIP theory expanding a group, and let a formula $\varphi(x, y)$ be such that for every $b \in \mathbb{M}^y$, $\varphi(\mathbb{M}, b)$ is a subgroup. Then, by Baldwin and Saxl [4], there exists $n \in \omega$ such that for all finite $B \subseteq \mathbb{M}^y$, there is $B_0 \subseteq B$ with $|B_0| \leq n$ such that $\bigcap_{b \in B_0} \varphi(\mathbb{M}, b) = \bigcap_{b \in B} \varphi(\mathbb{M}, b)$. So $\varphi(x, y)$ is a semi-equation by Proposition 2.16.

Remark 2.18. If $\varphi(x, y)$ is stable with infinitely many distinct instances $\varphi(\mathbb{M}, b), b \in \mathbb{M}^y$, then either $\varphi(x, y)$ is not a semi-equation, or $\neg\varphi(x, y)$ is not a semi-equation (combining [2, Proposition 2.20] and Proposition 2.16).

The following suggests that 1-semi-equationality can be viewed as a generalization of being 1-based from stable to the NIP context.

Proposition 2.19. *A formula $\varphi(x, y)$ is weakly normal if and only if it is stable and a* 1*-semi-equation. Hence a stable theory is* 1*-based if and only if it is* 1*-semi-equational.*

Proof. Clearly every k-weakly normal formula is a $(k, 1)$-semi-equation and is also an equation, hence stable. Conversely, suppose that $\varphi(x, y)$ is a $(k, 1)$-semi-equation and is stable, or just NSOP: there is some $\ell \in \omega$ such that there is no strictly increasing chain of sets of the form $\varphi(\mathbb{M}, b_0) \subsetneq \cdots \subsetneq \varphi(\mathbb{M}, b_\ell)$. We will show that then $\varphi(x, y)$ is k^ℓ-weakly normal. Let $(b_\eta)_{\eta \in [k]^\ell}$ be such that $\models \exists x \bigwedge_{\eta \in [k]^\ell} \varphi(x, b_\eta)$. For $\sigma \in [k]^{\leq \ell}$, we will show by induction on $m := \ell - |\sigma|$ that there are pairwise distinct $\eta_0, \ldots, \eta_m \in [k]^\ell$ extending σ (as sequences) such that $\varphi(\mathbb{M}, b_{\eta_0}) \subseteq \varphi(\mathbb{M}, b_{\eta_1}) \subseteq \cdots \subseteq \varphi(\mathbb{M}, b_{\eta_m})$. With $m = \ell$, so that $\sigma = \langle\rangle$ is the empty sequence, this implies by the choice of ℓ that there are $\eta \neq \eta' \in [k]^\ell$ such that $\varphi(\mathbb{M}, b_\eta) = \varphi(\mathbb{M}, b_{\eta'})$, as desired. The base case $(m = 0)$ is trivial, with $\eta_0 = \sigma$.

Assume the claim holds for m, and let $\sigma \in [k]^{\ell-(m+1)}$. For each $i \in [k]$, there exist pairwise distinct $\eta_{i,0}, \ldots, \eta_{i,m} \in [k]^\ell$ extending $\sigma^\frown i$ such that $\varphi(\mathbb{M}, b_{\eta_{i,0}}) \subseteq \cdots \subseteq \varphi(\mathbb{M}, b_{\eta_{i,m}})$. Among the sets $\{\varphi(\mathbb{M}, b_{\eta_{i,0}}) \mid i \in [k]\}$, one must be contained in another by $(k, 1)$-semi-equationality. Say $\varphi(\mathbb{M}, b_{\eta_{j,0}}) \subseteq \varphi(\mathbb{M}, b_{\eta_{i,0}})$ for some $i \neq j$. Then $\varphi(\mathbb{M}, b_{\eta_{j,0}}) \subseteq \varphi(\mathbb{M}, b_{\eta_{i,0}}) \subseteq \varphi(\mathbb{M}, b_{\eta_{i,1}}) \subseteq \cdots \subseteq \varphi(\mathbb{M}, b_{\eta_{i,m}})$, and

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

$\eta_{j,0}, \eta_{i,0}, \eta_{i,1}, \dots, \eta_{i,m}$ are pairwise distinct and extend σ, as desired. The "in particular" part follows by Fact 2.12. ⊣

Remark 2.20. It is well known that the family of weakly normal formulas is closed under conjunctions (but we could not find a direct reference). While semi-equations are closed under conjunctions by Proposition 2.3(2), this is not the case for the family of 1-semi-equations. Indeed, in a dense linear order, the formulas $x < y_1$ and $x > y_2$ are 1-semi-equations, but the formula $\varphi(x; y_1, y_2) := y_2 < x < y_1$ is not a 1-semi-equation since we can have any number of intervals with a non-empty intersection, so that none of them is contained in the other.

We observe a connection to another notion of "linearity" for NIP theories considered in [5], where various combinatorial results are proved for relations that are Boolean combinations of *basic* relations. The following is [5, Definition 2], in the case of binary relations (using the equivalence in [5, Proposition 2.8 and Remark 2.9]).

Definition 2.21. A binary relation $R \subseteq X \times Y$ is *basic* if there exist a linear order $(S, <)$ and functions $f : X \to S, g : Y \to S$ such that for any $a \in X, b \in Y$, $(a, b) \in R \iff f(a) < g(b)$.

Fact 2.22 [16, Claim 1 in the proof of Proposition 2.5]. *Let X be a set and $\mathcal{F} \subseteq \mathcal{P}(X)$ a family of subsets of X such that there are no $A, B \in \mathcal{F}$ satisfying $A \cap B \neq \emptyset, B \setminus A \neq \emptyset$ and $B \setminus A \neq \emptyset$ simultaneously. Then there exists a linear order $<$ on X so that every $A \in \mathcal{F}$ is a $<$-convex subset of X.*

Proposition 2.23.
(1) *Given a formula $\varphi(x, y) \in \mathcal{L}$, if the relation $R_\varphi := \{(a, b) \in \mathbb{M}^x \times \mathbb{M}^y : \models \varphi(a, b)\}$ is basic, then $\varphi(x, y)$ is a $(2, 1)$-semi-equation.*
(2) *If $\varphi(x, y)$ is $(2, 1)$-semi-equation, then $R_\varphi = R_1 \cap R_2$ for some (not necessarily definable) basic relations $R_1, R_2 \subseteq \mathbb{M}^x \times \mathbb{M}^y$.*

Proof. (1) Let $(S, <), f, g$ be as in Definition 2.21 for R_φ. Given any $b_1, b_2 \in \mathbb{M}^y$, the sets $\{x \in S : x < g(b_i)\}$ for $i \in \{1, 2\}$ are initial segments of S. Say $g(b_1) \leq g(b_2)$. Then for any $a \in \mathbb{M}^x$, $f(a) < g(b_1) \Rightarrow f(a) < g(b_2)$, so $\varphi(\mathbb{M}, b_1) \subseteq \varphi(\mathbb{M}, b_2)$, and the other case is symmetric.

(2) If $\varphi(x, y)$ is a $(2, 1)$-semi-equation, then the family $\mathcal{F}_\varphi$ of subsets of $\mathbb{M}^x$ satisfies the assumption in Fact 2.22. Hence there exists a (not necessarily definable) linear ordering $<'$ of $\mathbb{M}^x$ so that for every $b \in \mathbb{M}^y$, $\varphi(\mathbb{M}, b)$ is $<'$-convex. Let $(S, <)$ be the Dedekind completion of $(\mathbb{M}^x, <')$. Consider the functions $g_1, g_2 : \mathbb{M}^y \to S$ so that $g_1(b)$ is the infimum of $\varphi(\mathbb{M}, b)$ in S, and $g_2(b)$ is the supremum of $\varphi(\mathbb{M}, b)$ in S. Then $R_\varphi = \{(a, b) \in \mathbb{M}^x \times \mathbb{M}^y : g_1(b) \leq a\} \cap \{(a, b) \in \mathbb{M}^x \times \mathbb{M}^y : a \leq g_2(b)\}$, and both of these relations are basic (see [5, Remark 2.7]). ⊣

Remark 2.24. (1) In view of Proposition 2.23(2), if $\varphi(x, y)$ is a Boolean combination of $(2, 1)$-semi-equations, then by [5, Theorem 2.17 and Remark 2.20] the relation R_φ satisfies an almost linear Zarankiewicz bound. In particular, no infinite field can be defined in a $(2, 1)$-semi-equational theory (see [5, Corollary 5.11] or [42, Proposition 6.3] for a detailed explanation).

(2) If $\varphi(x, y)$ is a $(2, 1)$-semi-equation, then R_φ need not be basic. Indeed, the family of cosets of a subgroup is $(2, 1)$-semi-equational. If it was basic, then its complement is also basic, hence $(2, 1)$-semi-equational by the lemma above. But if

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

the index of the subgroup is ≥ 3, the family of complements of cosets is clearly not $(2, 1)$-semi-equational.

Problem 2.25. If $\varphi(x, y)$ is a $(k, 1)$-semi-equation for $k \geq 3$, is it still a Boolean combination of basic relations?

Problem 2.26. Show that no infinite field is definable in a 1-semi-equational theory.

Problem 2.27. Is every 1-semi-equational theory rosy? (Note that dense trees are not 1-semi-equational by Theorem 3.13.)

§3. Examples of semi-equational theories.

3.1. O-minimal structures. All *o*-minimal theories (and more generally, ordered dp-minimal theories) are distal (see [36]); hence, they are weakly semi-equational by Remark 4.4. Semi-equationality appears more subtle. An *o*-minimal structure $\mathcal{M}$ is *linear* if it has the *CF property* in the sense of [23], i.e., if every interpretable normal family of curves is of dimension at most 1. This is a weakening of local modularity of the pregeometry induced by the algebraic closure, and by the *o*-minimal trichotomy [30] it is equivalent to no infinite field being definable in $\mathcal{M}$. We will only need the following fact about linear *o*-minimal structures from [23]:

Fact 3.1. *Let* $T = \mathrm{Th}(\mathcal{M})$, *with* $\mathcal{M} = (M; <, +, \ldots)$ *a linear o-minimal expansion of a group. Let* $\mathcal{L} = (<, +, \ldots)$ *be the language of* T. *A partial endomorphism of* $\mathcal{M}$ *is a map* $f : (-c, c) \to M$, *for* c *an element of* M *or* ∞, *such that if* $a, b, a + b$ *are all in the domain, then* $f(a + b) = f(a) + f(b)$. *Let* $\mathcal{M}'$ *be the reduct of* $\mathcal{M}$ *to the language* $\mathcal{L}'$ *consisting of* $+, <$, *constant symbols naming* $\mathrm{acl}_{\mathcal{L}}(\emptyset)$, *and for each* $\mathcal{L}(\emptyset)$*-definable partial endomorphism* $f : (-c, c) \to M$ *with* $c \in \mathrm{acl}_{\mathcal{L}}(\emptyset)$ *or* $c = \infty$, *a unary function symbol interpreted as f on* $(-c, c)$ *and as* 0 *outside of the domain of f. Let* $T' := \mathrm{Th}_{\mathcal{L}'}(\mathcal{M}')$.

(1) [23, *Proposition* 4.2] *A subset of* M^n *is* $\emptyset$*-definable in* $\mathcal{M}$ *if and only if it is* $\emptyset$*-definable in* $\mathcal{M}'$.
(2) [23, *Corollary* 6.3] T' *admits quantifier elimination in the language* $\mathcal{L}'$.

Proposition 3.2. *Let* $T = \mathrm{Th}(\mathcal{M})$, *with* $\mathcal{M} = (M; <, \ldots)$ *an o-minimal structure.*

(1) *If* T *is an expansion of an ordered group and linear, then* T *is* $(2, 1)$*-semi-equational.*
(2) *Conversely, if T is* $(2, 1)$*-semi-equational, then T is linear.*

Proof. (1) Let $\mathcal{L} = (<, +, \ldots)$, $\mathcal{M}'$ and $\mathcal{L}'$ be as in Fact 3.1. By Fact 3.1(1) it suffices to show that $T' := \mathrm{Th}_{\mathcal{L}'}(\mathcal{M}')$ is $(2, 1)$-semi-equational. By Fact 3.1(2), it then suffices to show that every atomic $\mathcal{L}'$-formula $\varphi(x, y)$, with x, y arbitrary finite tuples of variables, is equivalent in T' to a Boolean combination of $(2, 1)$-semi-equations. By the proof of Theorem 4.3 in [1], every atomic $\mathcal{L}'$-formula $\varphi(x, y)$ is equivalent in T' to a Boolean combination of atomic formulas of the form $f(x) \square g(y) + c$, where $\square \in \{<, =, >\}$, $f : M^{|x|} \to M, g : M^{|y|} \to M$ are total multivariate $\mathcal{L}'(\emptyset)$-definable homomorphisms and $c \in \mathrm{dcl}_{\mathcal{L}'}(\emptyset)$. Every formula of this form clearly defines a basic relation on $M^{|x|} \times M^{|y|}$, hence is a $(2, 1)$-semi-equation by Proposition 2.23(1).

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

(2) By the o-minimal trichotomy theorem (see [30] and Remark 2 after the statement of Theorem 1.7 there), if $\mathcal{M}$ is not linear, then it defines an infinite field. But then Remark 2.24(1) implies that T is not $(2, 1)$-semi-equational. $\dashv$

PROBLEM 3.3. Is every o-minimal 1-semi-equational structure linear? A positive answer would follow from a positive answer to Problem 2.26.

PROBLEM 3.4. Which o-minimal theories are semi-equational? In particular, is $\mathrm{Th}(\mathbb{R}, +, \times)$ semi-equational?

3.2. Colored linear orders. Given a linearly ordered set $(S, <)$, a binary relation $R \subseteq S^2$ is *monotone* if $(x, y) \in R$, $x' \leq x$, and $y \leq y'$ implies $(x', y') \in R$.

FACT 3.5. *Let* $\mathcal{M} = \left(M, <, (C_i)_{i \in I}, (R_j)_{j \in J}\right)$ *be a linear order expanded by arbitrary unary* (C_i) *and monotone binary* (R_j) *relations. Then* $\mathrm{Th}(\mathcal{M})$ *is* $(2, 1)$-*semi-equational.*

PROOF. Let $\mathcal{M}'$ be an expansion of $\mathcal{M}$ obtained by naming all $\mathcal{L}_{\mathcal{M}}(\emptyset)$-definable unary and monotone binary relations, then a subset of M^n is $\emptyset$-definable in $\mathcal{M}$ if and only if it is $\emptyset$-definable in $\mathcal{M}'$, so it suffices to show that $T' := \mathrm{Th}(\mathcal{M}')$ is $(2, 1)$-semi-equational. By [35, Proposition 4.1], T' eliminates quantifiers. Note that if $R(x, y)$ is monotone, then it is a $(2, 1)$-semi-equation (given any $b_1 \leq b_2 \in M$, for any $a \in M$ we have $\models R(a, b_1) \to R(a, b_2)$ by monotonicity; hence, $R(M, b_1) \subseteq R(M, b_2)$). And any unary relation $C_i(x)$ is trivially a $(2, 1)$-semi-equation; hence, T' is $(2, 1)$-semi-equational. $\dashv$

3.3. Cyclic orders. A cyclic order $\circlearrowright(x, y, z)$ (see, e.g., [6, Section 5] or [41]) is *dense* if its underlying set is infinite and for every distinct a, b, there is c such that $\circlearrowright(a, b, c)$, and d such that $\circlearrowright(d, b, a)$. The theory $T_{\circlearrowright}$ of dense cyclic orders is complete and has quantifier elimination (see, e.g., [7, Proposition 3.7]).

PROPOSITION 3.6. (1) *The theory* $T_{\circlearrowright}$ *is not semi-equational.*
(2) *The theory* $T_{\circlearrowright}$ *expanded with one constant symbol* c *is* $(2, 1)$-*semi-equational.*

PROOF. (1) We show that $\psi(x_1, x_2; y) := \circlearrowright(x_1, x_2; y)$ is not a Boolean combination of semi-equations. By quantifier elimination, the formulas $\circlearrowright(x_1, x_2; y)$ and $\circlearrowright(x_2, x_1; y)$ each isolate a complete 3-type (over $\emptyset$). Any Boolean combination of formulas that is equivalent to $\circlearrowright(x_1, x_2; y)$ must contain some formula $\varphi(x_1, x_2; y)$ that is implied by $\circlearrowright(x_1, x_2; y)$ and is inconsistent with $\circlearrowright(x_2, x_1; y)$, or vice versa. Assume the former. Let $(c_i)_{i \in \mathbb{Z}}$ be such that $\models \circlearrowright(c_k, c_i, c_j)$ for $i < j < k$. Let $a_{1,i} = c_{2i}$, $a_{2,i} = c_{2i+2}$, and $b_i = c_{2i+1}$. Then $\models \circlearrowright(a_{2,i}, a_{1,i}; b_j) \Leftrightarrow i = j$ and $\models \circlearrowright(a_{1,i}, a_{2,i}; b_j) \Leftrightarrow i \neq j$, so $\models \varphi(a_{1,i}, a_{2,i}; b_j) \Leftrightarrow i \neq j$, so $\varphi(x_1, x_2; y)$ is not a semi-equation. If instead, $\varphi(x_1, x_2; y)$ is implied by $\circlearrowright(x_2, x_1; y)$ and inconsistent with $\circlearrowright(x_1, x_2; y)$, we switch the roles of x_1 and x_2 to get the same result.

(2) Let $<$ be defined by $x < y \Leftrightarrow \circlearrowright(x, y, c)$. Then $<$ is a dense linear order on the complement of $\{c\}$, so $x < y$ is a $(2, 1)$-semi-equation. We have that $\circlearrowright(x, y, z)$ is equivalent to $x < y < z \vee y < z < x \vee z < x < y \vee (z = c \wedge x < y) \vee (y = c \wedge z < x) \vee (x = c \wedge y < z)$. Hence $\circlearrowright(x, y, z)$ is a Boolean combination of $(2, 1)$-semi-equations (with c as a parameter), under any partition of the variables.

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

By quantifier elimination, it follows that every formula is a Boolean combination of $(2, 1)$-semi-equations (using c as a parameter). $\dashv$

This example shows that a theory being semi-equational, or 1-semi-equational, is not preserved under forgetting constants (naming constants clearly preserves k-semi-equationality). This is in contrast to equationality [20, Proposition 3.5] and distality [36, Corollary 2.9], which are invariant under naming or forgetting constants. This is also an example of a distal, non-semi-equational theory.

PROBLEM 3.7. Is weak semi-equationality of theories preserved by forgetting constants?

3.4. Ordered abelian groups. We consider ordered abelian groups, as structures in the language $\mathcal{L}_{\mathrm{CH}}$ introduced in [13]. Given an ordered abelian group $(G, +, <)$ and prime p, for $a \in G \setminus pG$ we let $G_p(a)$ be the largest convex subgroup of G such that $a \notin G_p(a) + pG$, and for $a \in pG$ let $G_p(a) := \{0\}$. Let $\mathcal{S}_p := \{G_p(a) : a \in G\}$. Then the $\mathcal{L}_{\mathrm{CH}}$-structure $\bar{G}$ corresponding to G consists of the main sort $\mathcal{G}$ for G, an auxiliary sort $\mathcal{S}_p$ for each p, along with countably many further auxiliary sorts and relations between them. A relative quantifier elimination result is obtained for such structures in [13], to which we refer for the details (see also [3, Section 3.2] for a quick summary).

PROPOSITION 3.8. *Every ordered abelian group (either as a pure ordered abelian group, or the corresponding structure $\bar{G}$) with finite auxiliary sorts $\mathcal{S}_p$ for all prime p is 1-semi-equational (this includes Presburger arithmetic, and any ordered abelian group with a strongly dependent theory by* [12, 14, 15, 17]).

PROOF. Since every auxiliary sort is finite and linearly ordered by a (definable) relation in $\mathcal{L}_{\mathrm{CH}}$, all auxiliary sorts are contained in $\mathrm{dcl}(\emptyset)$. Hence we only need to verify that every formula $\varphi(x, y)$ with x, y tuples of the main sort $\mathcal{G}$ is a Boolean combination of 1-semi-equations in the expansion with every element of every auxiliary sort named by a new constant symbol (countably many in total). As explained in [3, Proposition 3.14], it then follows from the relative quantifier elimination that $\varphi(x, y)$ is equivalent to a Boolean combination of atomic formulas of the form $\pi_\alpha(f(x)) \diamond_\alpha \pi_\alpha(g(y)) + k_\alpha$, where $\diamond \in \{=, <, \equiv_m\}$, $k \in \mathbb{Z}$, α is an element of an auxiliary sort, f, g are $\mathbb{Z}$-linear functions on $\mathcal{G}$, G_α is a corresponding convex subgroup of G, $\pi_\alpha : G \to G/G_\alpha$ is the quotient map, 1_α is the minimal positive element of G/G_α if it is discrete or $0 \in G/G_\alpha$ otherwise, and $k_\alpha = k \cdot 1_\alpha$ in G/G_α, and for $g, h \in G/G_\alpha$ we have $g \equiv_m h$ if $g - h \in m\,(G/G_\alpha)$ (note that these relations on G are expressible in the pure language of ordered abelian groups).

It is straightforward from Definition 2.13 that if $\varphi(x, y)$ is a (k, n)-semi-equation and $f(x), g(y)$ are $\emptyset$-definable functions, then the formula $\psi(x, y) := \varphi(f(x), g(y))$ is also a (k, n)-semi-equation. Using this (in an expansion of $\bar{G}$ naming π_α, and the ordered group structure on G/G_α together with the constants for k_α), we only have to show that the relations $x = y, x < y, x \in y + m\,(G/G_\alpha)$ on G/G_α are $(2, 1)$-semi-equations, which is straightforward. $\dashv$

PROBLEM 3.9. Is every ordered abelian group 1-semi-equational, or at least (weakly) semi-equational?

3.5. Trees. In this section we use "$\wedge$" to denote "meet," and "&" to denote conjunction. By a *tree* we mean a meet-semilattice $(M, \wedge)$ with an associated partial order $\leq$ (defined by $x \leq y \iff x \wedge y = x$) so that all of its initial segments are linear orders. An *infinitely branching dense tree* is a tree whose initial segments are dense linear orders and such that for each element x, there are infinitely many elements any two of which have meet x.

FACT 3.10 (see, e.g., [7, Lemma 3.14] or [27, Section 1]). *The theory of infinitely branching dense trees is complete and eliminates quantifiers in the language* $\{\wedge\}$.

LEMMA 3.11. *In any tree $\mathcal{M} = (M, \wedge)$ with no additional structure, if every formula of the form $\varphi(x; y_1, y_2)$ with x, y_1, y_2 singletons is a Boolean combination of semi-equations, then every formula is a Boolean combination of semi-equations.*

PROOF. By [35, Corollary 4.6] (using that $x \leq y \iff x \wedge y = x$), in any tree $\mathcal{M} = (M, \wedge)$ we have: two tuples $\bar{a} = (a_i : i \in [n]), \bar{b} = (b_j : j \in [n]) \in M^n$ have the same type if and only if (a_i, a_j, a_k) and (b_i, b_j, b_k) have the same type for every $i, j, k \in [n]$. Hence for any $\bar{a}, \bar{b}$, $\operatorname{tp}\left(\bar{a}\bar{b}\right)$ is implied by the set of formulas satisfied by three-element subtuples of $\bar{a}\bar{b}$. So if every partitioned formula with three total free variables is a Boolean combination of semi-equations, then $\operatorname{tp}\left(\bar{a}\bar{b}\right)$ is implied by a Boolean combination of semi-equations. In view of this, it is sufficient to assume that every formula of the form $\varphi(x; y_1, y_2)$ is a Boolean combination of semi-equations, because then by symmetry, every formula of the form $\varphi(x_1, x_2; y)$ is as well, and every partitioned formula with one of the parts empty (i.e., $\varphi(; y_1, y_2, y_3)$ or $\varphi(x_1, x_2, x_3;))$ is automatically a semi-equation. ⊣

THEOREM 3.12. *The theory of infinitely branching dense trees is semi-equational.*

PROOF. Let $\mathcal{M} = (M, \wedge)$ be an infinitely branching dense tree. By Lemma 3.11, it is enough to check that every formula $\varphi(x; y_1, y_2)$ is a Boolean combination of semi-equations, and, by Fact 3.10, it is enough to check this for positive atomic formulas $\varphi(x; y_1, y_2)$. Using the fact that $\wedge$ is associative, commutative, and idempotent, each such formula is equivalent to a formula of the form $\bigwedge A = \bigwedge B$ for non-empty $A, B \subseteq \{x, y_1, y_2\}$. By a direct case analysis (see [7, Theorem 3.16] for the details) every such formula is either a tautology, or does not mention x, or an equality between two variables, or is equivalent to a Boolean combination of the following formulas (possibly replacing y_2 by y_1):

(1) $x = x \wedge y_1$, i.e., $x \leq y_1$—a semi-equation: given $(a_i, b_i)_{i \in \mathbb{Z}}$ such that $\models a_i \leq b_j \iff i \neq j$, $a_i \leq b_0$ for $i \neq 0$, so $(a_i)_{i \neq 0}$ forms a chain. This is not consistent with $a_1 \leq b_2, a_2 \leq b_1, a_1 \nleq b_1, a_2 \nleq b_2$.

(2) $x \wedge y_1 \wedge y_2 = y_1 \wedge y_2$, i.e., $x \geq y_1 \wedge y_2$—a semi-equation for the same reason.

(3) $x \wedge y_1 = x \wedge y_2$—a negated semi-equation: given $(a_i, b_i, b'_i)_{i \in \mathbb{Z}}$ such that $\models a_i \wedge b_j = a_i \wedge b'_j \iff i = j$, for every $i \neq 0$ we have: either $a_i \wedge b_0 > a_0 \wedge b_0$ or $a_i \wedge b'_0 > a_0 \wedge b_0$. By pigeonhole, there are $i_1 \neq i_2$ such that the same case holds for both. Without loss of generality, $a_1 \wedge b_0 > a_0 \wedge b_0$ and $a_2 \wedge b_0 > a_0 \wedge b_0$. But then $a_1 \wedge a_2 > a_0 \wedge b_0 = a_0 \wedge a_1$, so a_1 and a_2 meet strictly closer to each other than to a_0. But, since $a_1 \wedge b_1 \leq a_0 \wedge b_1$ and $a_1 \wedge b_1 \leq a_2 \wedge b_1$, it must also be true that $a_0 \wedge a_2 \geq a_1 \wedge b_1 = a_1 \wedge a_0$, so a_0 and a_2 meet at least as closely to each other as to a_1. These are inconsistent.

(4) $x \wedge y_1 = x \wedge y_1 \wedge y_2$ (i.e., $x \wedge y_1 \leq y_2$)—a negated semi-equation: given $(a_i, b_i, b'_i)_{i \in \mathbb{Z}}$ such that $\models a_i \wedge b_j \leq b'_j \iff i = j$, in particular $a_0 \wedge b_0 \leq b'_0$ and $a_i \wedge b_0 \not\leq b'_0$ for $i \neq 0$. Since the initial segment below b_0 is totally ordered, it follows that $a_0 \wedge b_0 < a_i \wedge b_0$ for $i \neq 0$. $a_1 \wedge a_2 \geq (a_1 \wedge b_0) \wedge (a_2 \wedge b_0) > a_0 \wedge b_0 = a_0 \wedge a_1$. That is, a_1 and a_2 meet strictly closer together with each other than with a_0. But, by switching the roles of the indices 0 and 2 in that argument, a_0 and a_1 must meet strictly closer together with each other than with a_2 as well, a contradiction. ⊣

THEOREM 3.13. *In an infinitely branching dense tree $\mathcal{M} = (M, \wedge)$, the formula $x < y$ is not a Boolean combination of* 1*-semi-equations* (*without parameters*).

PROOF. By quantifier elimination, there are four complete 2-types over $\emptyset$ axiomatized by $\{x = y,\ x > y,\ x < y,\ x \perp y\}$, where $\perp$ denotes incomparable elements. Thus, up to equivalence, there are only 16 formulas $\varphi(x, y)$ with x, y singletons without parameters. By a direct case analysis (see [7, Theorem 3.17] for the details) the only 1-semi-equations among them are $x \neq x$, $x = x$, $x = y$, $x > y$, $x \geq y$. None of them separate $x < y$ from $x \perp y$, so any Boolean combination of 1-semi-equations implied by $x < y$ must also be implied by $x \perp y$, so $x < y$ is not equivalent to a Boolean combination of 1-semi-equations. ⊣

COROLLARY 3.14. *In any expansion of an infinitely branching dense tree $\mathcal{M} = (M, \wedge)$ by naming constants, the formula $x < y$ is not a Boolean combination of* 1*-semi-equations.*

PROOF. Suppose $x < y$ is equivalent to a Boolean combination of 1-semi-equations with parameters $c = (c_1, \dots, c_n)$. Say $x < y \iff \Phi(\varphi_1(x, y, c), \dots, \varphi_k(x, y, c))$, where Φ is a Boolean formula in k variables, and $\varphi_1(x, y, c), \dots, \varphi_k(x, y, c)$ are 1-semi-equations. Let d be an element such that $d \perp \bigwedge_{i \leq n} c_i$. For each i, let $\psi_i(x, y)$ be the formula $\exists z \left(\mathrm{tp}(z) = \mathrm{tp}(c) \,\&\, \left(x \wedge y \perp \bigwedge_{i \leq n} z_i\right) \& \, \varphi_i(x, y, z)\right)$. As $\mathrm{tp}(c)$ is isolated by quantifier elimination, this is indeed a first-order formula. For $a, b > d$, if $\models \varphi_i(a, b, c)$, then $\models \psi_i(a, b)$. By quantifier elimination and [35, Lemma 4.4], the converse also holds. Thus, for $a, b > d$, $\models a < b \iff \models \Phi(\varphi_1(a, b, c), \dots, \varphi_k(a, b, c)) \iff \models \Phi(\psi_1(a, b), \dots, \psi_k(a, b))$. Since all singletons have the same type, it follows that this holds for all a, b. It thus remains to show that each $\psi_i(x, y)$ is a 1-semi-equation, contradicting Theorem 3.13. If this were not the case for some $i \leq k$, then there would be $(b_j)_{j \in \mathbb{N}}$ and a such that $\models \psi_i(a, b_j)$ for all $j \in \mathbb{N}$, but such that for every $j \neq \ell \in \mathbb{N}$, there is $a_{j,\ell}$ such that $\models \psi_i(a_{j,\ell}, b_j)$ but $\not\models \psi_i(a_{j,\ell}, b_\ell)$. But, again because all singletons have the same type, and every finite set of elements has a lower bound, it is consistent that furthermore all of these elements are above d. But then this would also provide a counterexample to $\varphi_i(x, y)$ being a 1-semi-equation. ⊣

REMARK 3.15. Since $x > y$ is a $(2, 1)$-semi-equation and $x < y$ is not, this shows that being an (n, k)-semi-equation for fixed n, k (or even being a Boolean combination of them) is not preserved under exchanging the roles of the variables (while being a semi-equation is preserved).

REMARK 3.16. Note also that every tree admits an expansion in which $x < y$ is a Boolean combination of $(2, 1)$-semi-equations. In a tree, let $\leq_{\mathrm{lex}}$ be a linear order

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

refining $\leq$ such that for a, b, b' such that $a \perp b$ and $b \wedge b' > b \wedge a$, $a \leq_{\text{lex}} b \iff a \leq_{\text{lex}} b'$. Then let $\leq_{\text{revlex}}$ be given by $x \leq_{\text{revlex}} y :\iff x \leq y \vee (x \perp y \& y \leq_{\text{lex}} x)$. Then $\leq_{\text{revlex}}$ satisfies the same conditions as $\leq_{\text{lex}}$ (so both are $(2,1)$-semi-equations as both are linear orders), and $x \leq y \iff x \leq_{\text{lex}} y \& x \leq_{\text{revlex}} y$.

Problem 3.17. Is every theory of trees semi-equational? Is every theory of trees (expanded by constants) not 1-semi-equational?

§4. Weak semi-equations and strong honest definitions. In this section we discuss how (weak) semi-equationality naturally generalizes both distality and equationality.

Definition 4.1. Given a formula $\varphi(x, y) \in \mathcal{L}$ and a type p, we denote by $p^+_\varphi := \{\varphi(x, b) : \varphi(x, b) \in p\}$ the positive φ-part of the type p.

Given small sets $A, B, C \subseteq \mathbb{M}$, let $A \mathop{\smile\!\!\!\!\!\!|}^{u}_{C} B$ denote that $\operatorname{tp}(A/BC)$ is finitely satisfiable in C. We recall the following characterization of equations from [24, Lemma 2.4], which in turn is a variant of [38, Theorem 2.5]. Note that Fact 4.2(3) is equivalent to [24, Lemma 2.4(3)] since in stable theories non-forking is symmetric and equivalent to finite satisfiability over models. Existence of k in Fact 4.2(2) is not stated explicitly in [24, Lemma 2.4(2)], but is immediate from the proof.

Fact 4.2. *Given a formula $\varphi(x, y)$ in a stable theory T, the following are equivalent*:

(1) *$\varphi(x, y)$ is an equation (equivalently, $\varphi^*(y, x) := \varphi(x, y)$ is an equation).*
(2) *There is some $k \in \mathbb{N}$ such that for any $a \in \mathbb{M}^x$ and small $B \subseteq \mathbb{M}^y$, there is a subset B_0 of B of size at most k such that $\operatorname{tp}^+_\varphi(a/B_0) \vdash \operatorname{tp}^+_\varphi(a/B)$.*

On the other hand, we recall one of the standard characterizations of distality (see, e.g., [3, Corollary 1.11]), which we use as a definition here:

Definition 4.3. A theory is *distal* if and only if every formula $\varphi(x, y)$ is *distal*, that is, for any I_L and I_R infinite linear orders, $b \in \mathbb{M}^y$ and indiscernible sequence $(a_i)_{i \in I_L + (0) + I_R}$ with $a_i \in \mathbb{M}^x$ such that $(a_i)_{i \in I_L + I_R}$ is indiscernible over b, $\models \varphi(a_0, b) \iff \models \varphi(a_i, b)$ for $i \in I_L + I_R$.

There is a straightforward relationship between weak semi-equationality and distality:

Remark 4.4. A formula $\varphi(x, y)$ is distal if and only if both $\varphi(x, y)$ and $\neg\varphi(x, y)$ are weak semi-equations. In particular, every distal theory is weakly semi-equational.

Problem 4.5. Is there an NIP theory without a (weakly) semi-equational expansion? We note that while the theory ACF_p for $p > 0$ is known not to have a distal expansion [11], it is equational, and hence semi-equational.

An NIP theory is distal if and only if every formula admits a *strong honest definition*:

Fact 4.6 [10, Theorem 21]. *A theory T is distal if and only if for every formula $\varphi(x, y)$ there is a formula $\theta(x; y_1, \ldots, y_k)$, called a strong honest definition for $\varphi(x, y)$, such that for any finite set $C \subseteq \mathbb{M}^y$ ($|C| \geq 2$) and $a \in \mathbb{M}^x$, there is $b \in C^k$ such that $\models \theta(a; b)$ and $\theta(x; b) \vdash \operatorname{tp}_\varphi(a/C)$.*

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

We now show that in an NIP theory, weak semi-equationality is equivalent to the existence of *one-sided* strong honest definitions, which is also a generalization of Fact 4.2 (replacing a conjunction of finitely many instances of φ by some formula θ). We will need the following (p,k)-theorem of Matoušek from combinatorics:

FACT 4.7 [26]. *Let $\mathcal{F}$ be a family of subsets of some set X. Assume that the VC co-dimension of $\mathcal{F}$ is bounded by k. Then for every $p \geq k \in \mathbb{N}$, there is $N \in \mathbb{N}$ such that: for every finite subfamily $\mathcal{G} \subseteq \mathcal{F}$, if $\mathcal{G}$ has the (p,k)-property, meaning that among any p subsets of $\mathcal{G}$ some k intersect, then there is a subset of X of size N intersecting all sets in $\mathcal{G}$.*

THEOREM 4.8. *Let T be NIP, and let $\varphi(x,y)$ be a formula. The following are equivalent:*

(1) *The formula $\varphi^*(y,x) := \varphi(x,y)$ is a weak semi-equation.*
(2) *For every small $B \subseteq \mathbb{M}^y$ and $a \in \mathbb{M}^x$ with $\models \varphi(a,b)$ for all $b \in B$ there are $\theta(x;y_1,\dots,y_k), c \in (\mathbb{M}^y)^k$ such that $c \mathop{\smile\!\!\!\!\!\mid}^u_B a$, $\models \theta(a,c)$ and $\theta(x,c) \vdash \mathrm{tp}^+_\varphi(a/B)$.*
(3) *There is some formula $\theta(x;y_1,\dots,y_k)$ and number N such that for any finite set $B \subseteq \mathbb{M}^y$ with $|B| \geq 2$ and $a \in \mathbb{M}^x$, there is some $B_0 \subseteq B$ with $|B_0| \leq N$ such that $\mathrm{tp}^+_\theta(a/B_0) \vdash \mathrm{tp}^+_\varphi(a/B)$.*

PROOF. (1) implies (2). We follow closely the proof of [10, Proposition 19]. Assume that a,B are such that $\models \varphi(a,b)$ for all $b \in B$. Let $\mathcal{M} \preceq \mathbb{M}$ contain a,B, and let $(\mathcal{M}',B') \succ (\mathcal{M},B)$ be a $\kappa := |M|^+$-saturated elementary extension (with B named by a new predicate). We may assume $\mathcal{M}' \prec \mathbb{M}$ is a small submodel. Take $p(x) := \mathrm{tp}(a/B')$.

CLAIM 4.9. *Assume that $q(y) \in S_y(B')$ is a type finitely satisfiable in B. Then $p(x) \cup q(y) \vdash \varphi(x,y)$.*

PROOF. Let $\hat{q} \in S_y(\mathbb{M})$ be an arbitrary global type extending q and finitely satisfiable in B, and form the Morley product $\hat{q}^{(\omega)}(y_1,y_2,\dots) := \bigotimes_{i\in\mathbb{N}} q(y_i) \in S_{(y_1,y_2,\dots)}(\mathbb{M})$, also finitely satisfiable in B. For any set $C \subseteq \mathbb{M}$, we let $q|_C := \hat{q} \restriction_C$ (respectively, $q^{(\omega)}|_C := \hat{q}^{(\omega)} \restriction_C$) be the restriction of $\hat{q}$ (respectively, of $\hat{q}^{(\omega)}$) to formulas with parameters in C. As T is NIP, by [10, Lemma 5] there is some D with $B \subseteq D \subseteq B'$, $|D| < \kappa$ such that for any two realizations $I, I' \subseteq B'$ of $q^{(\omega)}|_D$ we have $aI \equiv_D aI'$. Fix some $I \models q^{(\omega)}|_D$ in B' (exists by saturation of $(\mathcal{M}',B')$ and finite satisfiability of $q^{(\omega)}|_{\mathbb{M}}$ in B) and $J \models q^{(\omega)}|_{\mathbb{M}}$ (in some larger monster model $\mathbb{M}' \succ \mathbb{M}$).

We claim that $I+J$ is indiscernible over aB. Indeed, as $q^{(\omega)}|_{\mathbb{M}}$ is finitely satisfiable in B, by compactness and saturation of $(\mathcal{M}',B')$ there is some $J' \models q^{(\omega)}|_{aDI}$ in B'. If $I+J$ is not aB-indiscernible, then $I'+J'$ is not aB-indiscernible for some finite subsequence I' of I. As by construction both $I'+J'$ and J' realize $q^{(\omega)}|_D$ in B', it follows by the choice of D that J' is not indiscernible over aB—contradicting the choice of J'.

Now let $b^* \in \mathbb{M}$ be any realization of q, then the sequence $I + (b^*) + J$ is Morley in $q|_{\mathbb{M}}$ over B, hence indiscernible (even over B). And $I+J$ is indiscernible over a (even over aB) by the previous paragraph. Note also that $\models \varphi(a,b)$ for every $b \in B'$ (by assumption we had $\models \varphi(a,b)$ for all $b \in B$, but $a \in \mathcal{M}$ and $(\mathcal{M}',B') \succ$

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

$(\mathcal{M}, B)$). Hence $\models \varphi(a,b)$ for every $b \in I + J$. And since $\varphi^*(y,x)$ is a weak semi-equation, this implies $\models \varphi(a,b^*)$. That is, for any $a \models p$ and $b^* \models q$, we have $\models \varphi(a,b^*)$, as wanted. $\dashv$

Now let S' be the set of types over B' finitely satisfiable in B, then S' is a closed subset of $S_y(B')$. By the claim, for every $q \in S'$ we have $p(x) \cup q(y) \vdash \varphi(x,y)$; hence, by compactness $\theta_q(x) \cup \psi_q(y) \vdash \varphi(x,y)$ for some formulas $\theta_q(X) \in p, \psi_q(y) \in q$. As $\{\psi_q(y) : q \in S'\}$ is a covering of the closed set S', it has a finite sub-covering $\{\psi_{q_k} : k \in K\}$. Let $\theta(x) := \bigwedge_{k \in K} \theta_{q_k}(x) \in p(x)$. As in particular $\mathrm{tp}(b/B) \in S'$ for every $b \in B$, we thus have $\theta(x) \in \mathcal{L}(B')$ (and $B' \mathop{\smile\!\!\!\!\!\!|}^{u}_{B} a$), $\models \theta(a)$ and $\theta(x) \vdash \mathrm{tp}^+_\varphi(a/B)$.

(2) implies (3). Let a, B be given. We either have that $\models \neg\varphi(a,b)$ holds for all $b \in B$, in which case $\mathrm{tp}^+_\theta(a/B_0) = \mathrm{tp}^+_\varphi(a/B) = \emptyset$, and $\emptyset \vdash \emptyset$ trivially. Or we replace B by $\{b \in B : \models \varphi(a,b)\}$, and follow the proof of (1) implies (2) in [10, Theorem 21].

We provide the details. By (2), given small $B \subseteq \mathbb{M}^y$ and $a \in \mathbb{M}^x$ such that $\models \varphi(a,b)$ for all $b \in B$, there exist $\theta(x; y_1, \dots, y_\ell)$ and $c \in (\mathbb{M}^y)^\ell$ such that $c \mathop{\smile\!\!\!\!\!\!|}^{u}_{B} a$, $\models \theta(a,c)$, and $\theta(x,c) \vdash \mathrm{tp}^+_\varphi(a/B)$. Then given any finite $B_0 \subseteq B$, there is $d \in B^\ell$ such that $\mathrm{tp}_\varphi(d/aB_0) = \mathrm{tp}_\varphi(c/aB_0)$, so in particular $\models \theta(a,d)$ and $\theta(x,d) \vdash \mathrm{tp}^+_\varphi(a/B_0)$.

Now fix an arbitrary function $f : \mathcal{L} \to \mathbb{N}$ and let $n_\theta := f(\theta(x; y_1, \dots, y_\ell))$ for every partitioned formula $\theta \in \mathcal{L}$ with x the same as before and ℓ arbitrary. Let T_f be a theory in the language $\mathcal{L} \cup \{P(x), a\}$ with P a new unary predicate and a a new constant symbol, so that T_f expands T with the following axioms: $\forall x \in P\, \varphi(a,x)$ and, for every formula $\theta(x; y_1, \dots, y_\ell) \in \mathcal{L}$, an axiom $\exists b_1, \dots, b_{n_\theta} \in P\, \forall c \in P^\ell\, (\neg\theta(a,c)) \vee \exists x \left(\theta(x,c) \wedge \bigvee_{i \le n_\theta} \neg\varphi(a,b_i)\right)$. By the previous paragraph, the theory T_f is inconsistent. By compactness, there is a finite inconsistent subset of T_f only requiring finitely many of these formulas $\theta_1, \dots, \theta_k$.

Thus there are finitely many formulas $\theta_1(x; y_1, \dots, y_{\ell_1}), \dots, \theta_k(x; y_1, \dots, y_{\ell_k}) \in \mathcal{L}$ such that: given $B \subseteq \mathbb{M}^y$ and $a \in \mathbb{M}^x$ such that $\models \varphi(a,b)$ for all $b \in B$, there is $i \le k$ such that for all $B_0 \subseteq B$ with $|B_0| \le n_{\theta_i}$, there is $c \in B^{\ell_i}$ such that $\models \theta_i(a;c)$ and $\theta_i(x;c) \vdash \mathrm{tp}^+_\varphi(a/B_0)$.

For each formula $\theta(x; y_1, \dots, y_\ell) \in \mathcal{L}$, let $\rho_\theta(x, y; z) := \theta(x;z) \wedge \forall w\, \theta(w;z) \to \varphi(w,y)$, and $n_\theta := \mathrm{VC}(\rho_\theta) + 1$ in the above argument (where VC is the VC-dimension), and let $\theta_1, \dots, \theta_k$ be as given by the previous paragraph for this choice of the n_θ's. Then for an arbitrary $a \in \mathbb{M}^x$ and finite $B \subseteq \mathbb{M}^y$ such that $\models \varphi(a;b)$ for $b \in B$, there is $i_{a,B} \le k$ such that: for all $B_0 \subseteq B$ with $|B_0| \le n_{\theta_{i_{a,B}}}$, there is $c \in B^{\ell_{i_{a,B}}}$ with $\models \theta_{i_{a,B}}(a;c)$ and $\theta_{i_{a,B}}(x;c) \vdash \mathrm{tp}^+_\varphi(a/B_0)$. For $b \in B$, consider the set

$$F^b_{a,B} := \left\{c \in B^{\ell_{i_{a,B}}} : \models \theta_{i_{a,B}}(a;c),\ \theta_{i_{a,B}}(x;c) \vdash \varphi(x;b)\right\}.$$

Note that $F^b_{a,B} = \left\{c \in B^{\ell_{i_{a,B}}} : \models \rho_{\theta_{i_{a,B}}}(a,b;c)\right\}$, and let $\mathcal{F}_{a,B} := \left\{F^b_{a,B} : b \in B\right\}$. By Fact 4.7 applied to $\mathcal{F}_{a,B}$, with $p = k = n_{\theta_{i_{a,B}}}$, there is $N_{i_{a,B}}$ (depending on $i_{a,B}$ but not otherwise depending on a, B) such that if every $n_{\theta_{i_{a,B}}}$ sets from $\mathcal{F}_{a,B}$ intersect, then there is $B_0 \subseteq B^{\ell_{i_{a,B}}}$ with $|B_0| \le N_{i_{a,B}}$ intersecting all sets from $\mathcal{F}_{a,B}$. Furthermore,

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

by choice of $i_{a,B}$, the condition that every $n_{\theta_{i_{a,B}}}$ sets from $\mathcal{F}_{a,B}$ intersect holds. And there are only k many possible values of $i_{a,B}$, so we let $N := \max_{1\leq i\leq k} N_i$.

We thus found $N \in \mathbb{N}$ such that: for all $a \in \mathbb{M}^x$ and finite $B \subseteq \mathbb{M}^y$ with $\models \varphi(a;b)$ for all $b \in B$, there is $i_{a,B} \leq k$ and $B_1 \subseteq B^{\ell_{i_{a,B}}}$ with $|B_1| \leq N$ intersecting all sets from $\mathcal{F}_{a,B}$, meaning that for every $b \in B$ there is $c \in B_1$ such that $\models \theta_{i_{a,B}}(a;c)$ and $\theta_{i_{a,B}}(x;c) \vdash \varphi(x;b)$. That is, $\mathrm{tp}^+_{\theta_{i_{a,B}}}(a/B_1) \vdash \mathrm{tp}^+_{\varphi}(a/B)$.

Finally, let $\theta(x;y_1,\dots,y_\ell)$ be a formula that can code for any $\theta_i\left(x;y_1,\dots,y_{\ell_i}\right)$ when parameters range over a set with at least two elements. For all $a \in \mathbb{M}^x$ and finite $B \subseteq \mathbb{M}^y$ with $|B| \geq 2$, for which $\models \varphi(a;b)$ for all $b \in B$, there is $B_0 \subseteq B$ with $2 \leq |B_0| \leq \ell N + 2$ (consisting of the coordinates of B_1 from the previous paragraph, and two points for coding) such that $\mathrm{tp}^+_{\theta}(a/B_0) \vdash \mathrm{tp}^+_{\varphi}(a/B)$, as desired.

(3) implies (1). This follows almost verbatim from the proof of (2) implies (1) in [10, Theorem 21]. Let $I + d + J$ be an indiscernible sequence in $\mathbb{M}^y$, with I and J infinite, and $I + J$ indiscernible over $a \in \mathbb{M}^x$, and suppose $\models \varphi(a,b)$ for $b \in I + J$. Let $I_1 \subset I$ with $|I_1| = N + 1$. Then there is some $I_0 \subseteq I_1$ such that $|I_0| \leq N$ and $\mathrm{tp}^+_{\theta}(a/I_0) \vdash \mathrm{tp}^+_{\varphi}(a/I_1)$. Let $b \in I_1 \setminus I_0$. By indiscernibility of $I + d + J$, there is some $\sigma \in \mathrm{Aut}(\mathbb{M})$ such that $\sigma(I_1) \subset I + d + J$ and $\sigma(b) = d$. We have $\sigma(I_0) \subseteq I + J$, so by a-indiscernibility of $I + J$, $\models \theta(a, \sigma(c))$ for every $c \in I_0^k$ for which $\models \theta(a,c)$, and hence $a \models \sigma\left(\mathrm{tp}^+_{\varphi}(a/I_1)\right)$. And $\varphi(x,b) \in \mathrm{tp}^+_{\varphi}(a/I_1)$, so $\varphi(x,d) \in \sigma\left(\mathrm{tp}^+_{\varphi}(a/I_1)\right)$, and hence $\models \varphi(a,d)$. ⊣

Problem 4.10. Can the assumption that T is NIP be omitted? (Note that the proof of (3) implies (1) does not use it.)

Proposition 2.16 immediately implies an analog of Fact 4.2 for semi-equations, telling us that $\varphi(x;y)$ is a one-sided strong honest definition for itself:

Corollary 4.11. *A formula $\varphi(x,y)$ (equivalently, $\varphi^*(y,x)$) is a semi-equation if and only if there is some $k \in \mathbb{N}$ such that: for every finite $B \subset \mathbb{M}^y$ and $a \in \mathbb{M}^x$ there is some $B_0 \subseteq B$ with $|B_0| \leq k$ such that* $\mathrm{tp}^+_{\varphi}(a/B_0) \vdash \mathrm{tp}^+_{\varphi}(a/B)$.

§5. Non weakly semi-equational valued fields. In this section we demonstrate that many valued fields are not weakly semi-equational. By an ac-*valued field field* we mean a three-sorted structure $(K, k, \Gamma, v, \mathrm{ac})$ in the Denef-Pas language, where K is a field, $v : K \to \Gamma$ is a valuation, with (ordered) value group Γ and residue field k, and $\mathrm{ac} : K \to k$ is the angular component map. As usual, $\mathcal{O} = \mathcal{O}_v$ denotes the valuation ring of v, and for $x \in \mathcal{O}$, $\bar{x}$ denotes the residue of x in k. The following is the main theorem of the section:

Theorem 5.1. *Let K be an ac-valued field for which the residue field k contains a non-constant totally indiscernible sequence (for instance, if k is infinite and stable), and which eliminates quantifiers of the main field sort (for example, a Henselian ac-valued field of equicharacteristic* 0 *with an algebraically closed residue field). Then K is not weakly semi-equational.*

Before presenting its proof, we need to develop some auxiliary results. First we provide a general sufficient criterion for when a formula is not a Boolean combination of weak semi-equations in Section 5.1. Then we discuss valuational

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

independence in Section 5.2. In Section 5.3 we describe a particular configuration of elements in a valued field indented to satisfy this sufficient criterion with respect to the formula $\psi(x_1, x_2; y_1, y_2) := v(x_1 - y_1) < v(x_2 - y_2)$, and reduce demonstrating that it has all of the required properties to Claims 5.7 and 5.8 which express a certain amount of indiscernibility of our configuration. We also explain how both claims can be proved by induction on the complexity of the formula and reduce to several essential cases that have to be considered; and show in Claim 5.9 valuational independence of some elements of our configuration which will be helpful in the proof of the claims. We then prove Claim 5.7 in Section 5.4 and Claim 5.8 in Section 5.5, concluding the proof of Theorem 5.1. Finally, in Section 5.6 we discuss some further applications of Theorem 5.1 and examples.

5.1. Boolean combinations of weak semi-equations. We provide a sufficient criterion for when a formula is not a Boolean combination of weak semi-equations (analogous to a criterion for equations from [28]).

LEMMA 5.2. *If $\varphi(x, y)$ and $\psi(x, y)$ are weak semi-equations, then there are no $b \in \mathbb{M}^y$ and array $(a_{i,j})_{i,j\in\mathbb{Z}}$ with $a_{i,j} \in \mathbb{M}^x$ such that:*

- *Every row (i.e., $(a_{i,j} : j \in \mathbb{Z})$ for a fixed $i \in \mathbb{Z}$) and every column (i.e., $(a_{i,j} : i \in \mathbb{Z})$ for a fixed $j \in \mathbb{Z}$) is indiscernible (over $\emptyset$).*
- *Rows and columns without their 0-indexed elements (i.e., $\left(a_{i,j}\right)_{j\neq 0}$ for fixed i, and $\left(a_{i,j}\right)_{i\neq 0}$ for fixed j) are b-indiscernible.*
- $\models \varphi\left(a_{i,j}, b\right) \wedge \neg\psi\left(a_{i,j}, b\right) \iff i = 0 \vee j \neq 0.$

PROOF. Assume there exist an array $(a_{i,j} : i, j \in \mathbb{Z})$ and b with these properties. For any fixed $i \neq 0$, we have $\models \varphi(a_{i,j}, b)$ for all $j \neq 0$, $(a_{i,j})_{j\in\mathbb{Z}}$ is indiscernible, and $(a_{i,j})_{j\neq 0}$ is b-indiscernible, so, by weak semi-equationality of φ, $\models \varphi(a_{i,0}, b)$. But $\not\models \varphi(a_{i,0}, b) \wedge \neg\psi(a_{i,0}, b)$, so $\models \psi(a_{i,0}, b)$. Now the sequence $(a_{i,0})_{i\in\mathbb{Z}}$ is indiscernible, $(a_{i,0})_{i\neq 0}$ is b-indiscernible, and $\models \psi(a_{i,0}, b)$ for all $i \neq 0$, so, by weak semi-equationality of ψ, $\models \psi(a_{0,0}, b)$—contradicting $\models \varphi(a_{0,0}, b) \wedge \neg\psi(a_{0,0}, b)$. ⊣

LEMMA 5.3. *If $\varphi(x, y)$ is a Boolean combination of weak semi-equations, then there are no $b \in \mathbb{M}^y$ and array $(a_{i,j})_{i,j\in\mathbb{Z}}$ with $a_{i,j} \in \mathbb{M}^x$ such that:*

- *Rows and columns of $\left(a_{i,j}\right)_{i,j\in\mathbb{Z}}$ are indiscernible.*
- *Rows and columns without their 0-indexed elements (i.e., $\left(a_{i,j}\right)_{j\neq 0}$ for fixed i, and $\left(a_{i,j}\right)_{i\neq 0}$ for fixed j) are b-indiscernible.*
- $\models \varphi\left(a_{i,j}, b\right) \iff i = 0 \vee j \neq 0.$
- *All $a_{i,j}$ with $i = 0$ or $j \neq 0$ have the same type over b.*

PROOF. Any conjunction of finitely many weak semi-equations and negations of weak semi-equations is of the form $\psi(x, y) \wedge \neg\theta(x, y)$ for some weak semi-equations $\psi(x, y)$ and $\theta(x, y)$, because weak semi-equations are closed under conjunction and under disjunction (Proposition 2.3(3)), so negations of weak semi-equations are also closed under conjunction. Thus any Boolean combination of weak semi-equations is equivalent, via its disjunctive normal form, to $\bigvee_{k\in I}(\psi_k(x, y) \wedge \neg\theta_k(x, y))$ for some finite index set I and weak semi-equations $\psi_k(x, y)$ and $\theta_k(x, y)$ for $k \in I$. Given b and $(a_{i,j})_{i.j\in\mathbb{Z}}$ as above, since

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

$i = 0 \vee j \neq 0 \iff \models \varphi(a_{i,j}, b) \iff \models \bigvee_{k\in I} (\psi_k(a_{i,j}, b) \wedge \neg\theta_k(a_{i,j}, b))$, and all $a_{i,j}$ with $i = 0$ or $j \neq 0$ have the same type over b, there is some k such that $\models \psi_k(a_{i,j}, b) \wedge \neg\theta_k(a_{i,j}, b) \iff i = 0 \vee j \neq 0$, contradicting Lemma 5.2. ⊣

5.2. Valuational independence.

DEFINITION 5.4. Let K be a field with valuation v.

(1) We say that $a_1, \dots, a_n \in K$ are *valuationally independent* if, for every polynomial $f(x_1, \dots, x_n) = \sum_i c_i x_1^{\alpha_{1,i}} \dots x_n^{\alpha_{n,i}}$ (where i runs over some finite index set, $c_i, \alpha_{1,i}, \dots, \alpha_{n,i} \in \mathbb{Z}$, and $(\alpha_{1,i}, \dots, \alpha_{n,i}) \neq (\alpha_{1,j}, \dots, \alpha_{n,j})$ for $i \neq j$) we have

$$v(f(a_1, \dots, a_n)) = \min_i v\left(c_i a_1^{\alpha_{1,i}} \dots a_n^{\alpha_{n,i}}\right).$$

That is, if the valuation of every polynomial applied to $a_1, \dots, a_n$ is the minimum of the valuations of its monomials (including their coefficients).

(2) An infinite set is *valuationally independent* if every finite subset is.

EXAMPLE 5.5. (1) A set of elements with valuation 0 is valuationally independent if and only if their residues are algebraically independent.

(2) In a valued field of pure characteristic, every set of elements whose valuations are $\mathbb{Z}$-linearly independent is valuationally independent. In mixed characteristic $(0, p)$, every set of elements whose valuations, together with $v(p)$, are $\mathbb{Z}$-linearly independent, is valuationally independent. In an ac-valued field, this is the only way for a set of elements with angular component 1 to be valuationally independent.

5.3. Reducing the proof of Theorem 5.1 to two claims. We will show that the partitioned formula $\psi(x_1, x_2; y_1, y_2) := v(x_1 - y_1) < v(x_2 - y_2)$ is not a Boolean combination of weak semi-equations.

Without loss of generality we may assume that K is a monster model. By Lemma 5.3, it suffices to find b, b' and $(a_i)_{i\in\mathbb{Z}}, \left(a'_j\right)_{j\in\mathbb{Z}}$ in K such that the sequences $(a_i)_{i\in\mathbb{Z}}$ and $\left(a'_j\right)_{j\in\mathbb{Z}}$ are mutually indiscernible (so that rows and columns of the array $\left(a_i a'_j\right)_{i,j\in\mathbb{Z}}$ are indiscernible), $(a_i)_{i\neq 0}$ is indiscernible over $bb'\left(a'_j\right)_{j\in\mathbb{Z}}$, $\left(a'_j\right)_{j\neq 0}$ is indiscernible over $bb'(a_i)_{i\in\mathbb{Z}}$ (so that the rows and the columns of the array $\left(a_i a'_j\right)_{i,j\in\mathbb{Z}}$ with their 0-indexed elements removed are indiscernible over bb'), $\models v(a_i - b) < v\left(a'_j - b'\right) \iff i \neq 0 \vee j = 0$, and all pairs $\left(a_i, a'_j\right)$ with $i \neq 0$ or $j = 0$ have the same type over bb'.

To find these elements, first let $0 < \gamma_0 < \gamma_1 < \gamma_2 < \gamma_3 < \gamma_4 < \gamma_5 < \gamma_6 \in \Gamma$ be an increasing *indiscernible* sequence of positive elements of the value group (exists by Ramsey and saturation).

CLAIM 5.6. *The elements $\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5$ are $\mathbb{Z}$-linearly independent (in Γ viewed as a $\mathbb{Z}$-module). If K has mixed characteristic $(0, p)$, then $v(p), \gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5$ are $\mathbb{Z}$-linearly independent.*

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

PROOF. If $n_0\gamma_0 + \cdots + n_5\gamma_5 = 0$ with $n_0, \ldots, n_5 \in \mathbb{Z}$ not all 0, let $i \leq 5$ be maximal such that $n_i \neq 0$. Now $n_0\gamma_0 + \cdots + n_i\gamma_i = 0$, and $n_i \neq 0$. By indiscernibility of the sequence $(\gamma_1, \ldots, \gamma_6)$, $n_0\gamma_0 + \cdots + n_{i-1}\gamma_{i-1} + n_i\gamma_6 = 0$, but then $n_i(\gamma_i - \gamma_6) = 0$, contradicting that $n_i \neq 0$, $\gamma_i \neq \gamma_6$, and Γ is ordered and thus torsion-free. In mixed characteristic, the same argument can be repeated starting from $n_0\gamma_0 + \cdots + n_5\gamma_5 = mv(p)$ with $n_0, \ldots, n_5, m \in \mathbb{Z}$. ⊣

Next let $a_\infty, a'_\infty \in K$ be such that $v(a_\infty) = \gamma_0$, $v(a'_\infty) = \gamma_1$, and $\operatorname{ac}(a_\infty) = \operatorname{ac}(a'_\infty) = 1$. Let $(\tilde{a}_i)_{i\in\mathbb{Z}} + \left(\tilde{b}\right)$ and $\left(\tilde{a}'_j\right)_{j\in\mathbb{Z}} + \left(\tilde{b}'\right)$ be arbitrary mutually totally indiscernible sequences in the residue field k. Such sequences exist by assumption on k and saturation, e.g., splitting a totally indiscernible sequence into two disjoint subsequences. We define $a_i := a_\infty + \alpha \operatorname{lift}(\tilde{a}_i)$ and $a'_j := a'_\infty + \beta \operatorname{lift}\left(\tilde{a}'_j\right)$ for $i, j \in \mathbb{Z}$, for some $\alpha, \beta \in K$ with $v(\alpha) = \gamma_2$, $v(\beta) = \gamma_3$, and $\operatorname{ac}(\alpha) = \operatorname{ac}(\beta) = 1$. Here $\operatorname{lift}(x)$ is some arbitrary element of $\mathcal{O}$ such that $\overline{\operatorname{lift}(x)} = x$. Let b, b' be such that $v(a_0 - b) = \gamma_4$, $v\left(a'_0 - b'\right) = \gamma_5$, $\operatorname{ac}(a_0 - b) = \tilde{b} - \tilde{a}_0$, and $\operatorname{ac}\left(a'_0 - b'\right) = \tilde{b}' - \tilde{a}'_0$. All of these elements are fixed for the rest of the section.

It is clear that $\models v(a_i - b) < v\left(a'_j - b'\right) \iff i \neq 0 \vee j = 0$, because $v(a_0 - b) = \gamma_4$, $v(a_i - b) = \gamma_2$ for $i \neq 0$, $v\left(a'_0 - b'\right) = \gamma_5$, and $v\left(a'_j - b'\right) = \gamma_3$ for $j \neq 0$.

We will prove the following two claims. Given a sequence $(x_i)_{i\in I}$ and $J \subseteq I$, we will write x_J to denote the subsequence $(x_i : i \in J)$.

CLAIM 5.7. (1) *Let $\varphi\left(x; z; w; b', a'_J\right)$ be a formula with parameters b' and a'_J for some $J \subseteq \mathbb{Z}$, tuples of variables x of sort K, z of sort k, and w of sort Γ_∞. Let I_1, I_2 be tuples of distinct indices from $\mathbb{Z}$, with $|I_1| = |I_2| = |x|$. Let $\sigma \in \operatorname{Aut}(k)$ be such that $\sigma\left(\tilde{a}_{I_1}\right) = \tilde{a}_{I_2}$ (preserving the ordering of the tuples), $\sigma\left(\tilde{a}'_J\right) = \tilde{a}'_J$, and $\sigma\left(\tilde{b}'\right) = \tilde{b}'$. Then for any tuples $c \in k^z$, $d \in \Gamma_\infty^w$ we have $\models \varphi\left(a_{I_1}; c; d; b'; a'_J\right) \iff \models \varphi\left(a_{I_2}; \sigma(c); d; b'; a'_J\right)$.*

(2) *Likewise, let $\varphi(y; z; w; b, a_I)$ be a formula with parameters b and a_I for some $I \subseteq \mathbb{Z}$, tuples of variables y of sort K, z of sort k, and w of sort Γ_∞. Let J_1, J_2 be tuples of distinct indices from $\mathbb{Z}$, with $|J_1| = |J_2| = |y|$, and let $\sigma \in \operatorname{Aut}(k)$ be such that $\sigma\left(\tilde{a}'_{J_1}\right) = \tilde{a}'_{J_2}$, $\sigma(\tilde{a}_I) = \tilde{a}_I$, and $\sigma\left(\tilde{b}\right) = \tilde{b}$. Then for any tuples $c \in k^z$, $d \in \Gamma_\infty^w$ we have*

$$\models \varphi\left(a'_{J_1}; c; d; b; a_I\right) \iff \models \varphi\left(a'_{J_2}; \sigma(c); d; b; a_I\right).$$

CLAIM 5.8. *Let $\varphi(x; y; z; w; b; b')$ be a formula with parameters b, b', where x and y are single variables of sort K, and z and w are tuples of variables of sort k and Γ_∞, respectively. Let $\sigma_i \in \operatorname{Aut}(k)$ be such that $\sigma_i(\tilde{a}_i) = \tilde{b}$, $\sigma_i(\tilde{a}_0) = \tilde{a}_0$, $\sigma_i\left(\tilde{a}'_0\right) = \tilde{a}'_0$, and $\sigma_i\left(\tilde{b}'\right) = \tilde{b}'$. Let $\sigma'_j \in \operatorname{Aut}(k)$ be such that $\sigma'_j\left(\tilde{b}'\right) = \tilde{a}'_j$, $\sigma'_j\left(\tilde{a}'_0\right) = \tilde{a}'_0$, $\sigma'_j(\tilde{a}_i) = \tilde{a}_i$, and $\sigma'_j\left(\tilde{b}\right) = \tilde{b}$. Let $\pi \in \operatorname{Aut}(\Gamma_\infty)$ be such that $\pi(\gamma_2) = \gamma_4$, $\pi(\gamma_0) = \gamma_0$, $\pi(\gamma_1) = \gamma_1$, and $\pi(\gamma_5) = \gamma_5$, and let $\tau \in \operatorname{Aut}(\Gamma_\infty)$ be such that $\tau(\gamma_5) = \gamma_3$, $\tau(\gamma_0) = \gamma_0$, $\tau(\gamma_1) = \gamma_1$, and $\tau(\gamma_2) = \gamma_2$. Then, for $i, j \neq 0$, $c \in k^z$,* *and* $d \in \Gamma_\infty^w$, $\models \varphi\left(a_0; a_0'; \sigma_i(c); \pi(d); b; b'\right) \iff \models \varphi\left(a_i; a_0'; c; d; b; b'\right) \iff \models \varphi\left(a_i; a_j'; \sigma_j'(c); \tau(d); b; b'\right)$.

Assuming these two claims, from the $|z| = |w| = 0$ case of Claim 5.7, we get that $(a_i)_{i\in\mathbb{Z}}$ is totally indiscernible over $b'\left(a_j'\right)_{j\in\mathbb{Z}}$, and $\left(a_j'\right)_{j\in\mathbb{Z}}$ is totally indiscernible over $b\,(a_i)_{i\in\mathbb{Z}}$. In particular $(a_i)_{i\in\mathbb{Z}}$ and $\left(a_j'\right)_{j\in\mathbb{Z}}$ are mutually totally indiscernible.

In describing $(a_i)_{i\in\mathbb{Z}}, \left(a_j'\right)_{j\in\mathbb{Z}}, b, b'$, we have made exactly the same assumptions about a_0 as about b, and the same assumptions about a_0' as about b', in the sense that if we replace a_0 with b or replace a_0' with b', the resulting elements $(a_i)_{i\in\mathbb{Z}}, \left(a_j'\right)_{j\in\mathbb{Z}}, b, b'$ could have come from the same construction. Thus, as Claim 5.7 implies that $(a_i)_{i\neq 0}$ is totally indiscernible over $a_0 b'\left(a_j'\right)_{j\in\mathbb{Z}}$, and $\left(a_j'\right)_{j\neq 0}$ is totally indiscernible over $a_0' b\,(a_i)_{i\in\mathbb{Z}}$, it must also be the case that $(a_i)_{i\neq 0}$ is totally indiscernible over $bb'\left(a_j'\right)_{j\in\mathbb{Z}}$, and $\left(a_j'\right)_{j\neq 0}$ is totally indiscernible over $b'b\,(a_i)_{i\in\mathbb{Z}}$.

From the $|z| = |w| = 0$ case of Claim 5.8, we get that

$$\operatorname{tp}\left(a_i, a_j'/b, b'\right) = \operatorname{tp}\left(a_i, a_0'/b, b'\right) = \operatorname{tp}\left(a_0, a_0'/b, b'\right)$$

for $i, j \neq 0$; hence, all $\left(a_i, a_j'\right)$ with $i \neq 0$ or $j = 0$ have the same type over bb'.

Thus these two claims establish the conditions needed for Lemma 5.3 to imply that $v(x_1 - y_1) < v(x_2 - y_2)$ is not a Boolean combination of weak semi-equations.

Both claims will be proved by induction on the parse tree of the formula φ (without parameters). There are five cases that must be considered:

Case 1. The formula φ is of the form $t_1 \leq t_2$, where t_1, t_2 are terms of sort Γ_∞. Such terms are $\mathbb{N}$-linear combinations of variables of sort Γ_∞ and valuations of polynomials in variables of sort K; i.e., of the form $\boldsymbol{n} \cdot x + \boldsymbol{m} \cdot v(f(y))$, where $x = \left(x_1, \ldots, x_{\ell_1}\right)$ is a tuple of variables of sort Γ_∞, y is a tuple of variables of sort K, f is a tuple of polynomials $\left(f_1(y), \ldots, f_{\ell_2}(y)\right)$, $\boldsymbol{n} = \left(n_1, \ldots, n_{\ell_1}\right) \in \mathbb{N}^{|x|}$, $\boldsymbol{m} = \left(m_1, \ldots, m_{\ell_2}\right) \in \mathbb{N}^{|f|}$, $v\left(f(y)\right)$ is an abbreviation for the tuple $\left(v\left(f_1(y)\right), \ldots, v\left(f_\ell(y)\right)\right)$, and "$\cdot$" is the dot product.

Case 2. φ is of the form $t_1 =_k t_2$, where t_1, t_2 are terms of sort k. Terms of sort k are polynomials applied to variables of sort k and angular components of terms of sort K; i.e., of the form $f(x, \operatorname{ac}(g(y)))$, where f is a polynomial, $g = (g_1, \ldots, g_\ell)$ is a tuple of polynomials, x is a tuple of variables of sort k, y is a tuple of variables of sort K, and $\operatorname{ac}(g(y))$ is an abbreviation for the tuple $\left(\operatorname{ac}(g_1(y)), \ldots, \operatorname{ac}(g_\ell(y))\right)$. Since $t_1 =_k t_2$ if and only if $t_1 - t_2 =_k 0$, every formula of this form is equivalent to a formula of the form $f(x, \operatorname{ac}(g(y))) =_k 0$.

Case 3. φ is a Boolean combination of formulas for which the claim holds.

Case 4. φ is of the form $\exists u\, \psi$, with u a variable of sort k, and the claim holds for ψ.

Case 5. φ is of the form $\exists u\, \psi$, with u a variable of sort Γ_∞, and the claim holds for ψ.

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

There are four more cases for how φ could be constructed, but they follow from the previous five cases: φ is of the form $t_1 =_\Gamma t_2$, where t_1, t_2 are terms of sort Γ_∞—this is equivalent to $t_1 \leq t_2 \wedge t_2 \leq t_1$, and is thus redundant with Cases 1 and 3; φ is of the form $t_1 =_K t_2$, where t_1, t_2 are terms of sort K—this is equivalent to $v(t_1 - t_2) = v(0)$, and is thus redundant with Cases 1 and 3; φ is of the form $\forall u\, \psi$, where u is a variable of sort k or Γ_∞—this is redundant with Cases 3–5; φ is of the form $\exists u\, \psi$, or $\forall u\, \psi$, where u is a variable of sort K—this case can be neglected by quantifier elimination, since we can always pick a formula equivalent to φ which has no quantifiers of sort K.

The following auxiliary result will be used in the proof of the claims.

Claim 5.9. *The elements* $a_\infty, a'_\infty, (\alpha\,\mathrm{lift}(\tilde{a}_i))_{i\in\mathbb{Z}}, \left(\beta\,\mathrm{lift}\left(\tilde{a}'_j\right)\right)_{j\in\mathbb{Z}}, b - a_0,$ $b' - a'_0$ *are valuationally independent.*

Proof. Define a valuation $v^* : \mathbb{Z}[u, v, x, y, z, w] \to \Gamma_\infty$ (with $|u| = |v| = |z| = |w| = 1$, $|x|, |y|$ arbitrary), by, for monomials (which in case of mixed characteristic is taken to include its coefficient),

$$v^*\left(n \cdot u^{r_\infty} x_1^{r_1} \dots x_{|x|}^{r_{|x|}} v^{s_\infty} y_1^{s_1} \dots y_{|y|}^{s_{|y|}} z^{t_1} w^{t_2}\right)$$
$$:= v(n) + r_\infty\gamma_0 + s_\infty\gamma_1 + \left(r_1 + \dots + r_{|x|}\right)\gamma_2 + \left(s_1 + \dots + s_{|y|}\right)\gamma_3 + t_1\gamma_4 + t_2\gamma_5,$$

and the valuation of a polynomial is the minimum of the valuations of its monomials. That way, for any $I, J \subseteq \mathbb{Z}$ with $|I| = |x|$ and $|J| = |y|$ we have

$$v^*(f(u, v, x, y, z, w)) = v\left(f\left(a_\infty, a'_\infty, \alpha \cdot \mathrm{lift}(\tilde{a}_I), \beta \cdot \mathrm{lift}(\tilde{a}'_J), b - a_0, b' - a'_0\right)\right)$$

when f is a monomial (where $\alpha \cdot \mathrm{lift}(\tilde{a}_I) := \left(\alpha\,\mathrm{lift}(\tilde{a}_i)\right)_{i\in I}$), and we need to prove that this holds for all polynomials f. Given a polynomial $f(u, v, x, y, z, w)$,

$$v^*(f) = v(n) + m_0\gamma_0 + m_1\gamma_1 + m_2\gamma_2 + m_3\gamma_3 + m_4\gamma_4 + m_5\gamma_5$$

for some $n, m_0, m_1, m_2, m_3, m_4, m_5 \in \mathbb{N}$ (with $v(n), m_0, m_1, m_2, m_3, m_4, m_5$ unique by Claim 5.6). Let $\tilde{f}(u, v, x, y, z, w)$ be the sum of monomials in f of the same valuation as f, so that every monomial appearing in $\tilde{f}(u, v, x, y, z, w)$ has degree m_0 in u, degree m_1 in v, total degree m_2 in x, total degree m_3 in y, degree m_4 in z, degree m_5 in w, and has leading coefficient with valuation $v(n)$, and $v^*\left(f - \tilde{f}\right) > v^*(f)$. Thus

$$\frac{\tilde{f}(u, v, x, y, z, w)}{n \cdot u^{m_0} v^{m_1} z^{m_4} w^{m_5}}$$

is a non-zero polynomial in x, y, all coefficients having valuation 0, so it reduces under the residue map to a nonzero polynomial in x, y. Since the set of elements in the tuples $\tilde{a}_I, \tilde{a}'_J$ is algebraically independent (they come from an infinite indiscernible sequence), it follows that

$$\frac{\tilde{f}\left(u, v, \tilde{a}_I, \tilde{a}'_J, z, w\right)}{n \cdot u^{m_0} v^{m_1} z^{m_4} w^{m_5}} \neq 0,$$

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

and thus a lift of it,

$$\frac{\tilde{f}\left(u, v, \operatorname{lift}\left(\tilde{a}_I\right), \operatorname{lift}\left(\tilde{a}'_J\right), z, w\right)}{n \cdot u^{m_0} v^{m_1} z^{m_4} w^{m_5}},$$

has valuation 0. Thus

$$v\left(\tilde{f}\left(a_\infty, a'_\infty, \operatorname{lift}\left(\tilde{a}_I\right), \operatorname{lift}\left(\tilde{a}'_J\right), b - a_0, b' - a'_0\right)\right) = v(n) + m_0\gamma_0 + m_1\gamma_1 + m_4\gamma_4 + m_5\gamma_5$$

and, by homogeneity of $\tilde{f}$,

$$\begin{aligned} & v\left(\tilde{f}\left(a_\infty, a'_\infty, \alpha \cdot \operatorname{lift}\left(\tilde{a}_I\right), \beta \cdot \operatorname{lift}\left(\tilde{a}'_J\right), b - a_0, b' - a'_0\right)\right) \\ &= v(n) + m_0\gamma_0 + m_1\gamma_1 + m_2\gamma_2 + m_3\gamma_3 + m_4\gamma_4 + m_5\gamma_5 = v^*(f). \end{aligned}$$

We have

$$v\left(\left(f - \tilde{f}\right)\left(a_\infty, a'_\infty, \alpha \cdot \operatorname{lift}\left(\tilde{a}_I\right), \beta \cdot \operatorname{lift}\left(\tilde{a}'_J\right), b - a_0, b' - a'_0\right)\right) \geq v^*\left(f - \tilde{f}\right) > v^*(f)$$

(the first inequality holds by the ultrametric property, combined with the fact that it holds for monomials), so it follows that

$$v\left(f\left(a_\infty, a'_\infty, \alpha \cdot \operatorname{lift}\left(\tilde{a}_I\right), \beta \cdot \operatorname{lift}\left(\tilde{a}'_J\right), b - a_0, b' - a'_0\right)\right) = v^*(f). \qquad \dashv$$

We are ready to prove the two claims.

5.4. Proof of Claim 5.7. Let $\varphi\left(x; z; w; b'; a'_J\right)$ with $x = \left(x_1, \ldots, x_{|x|}\right)$ and I_1, I_2, σ, c, d be as in the statement of the claim, and we analyze the five cases described above. We will assume without loss of generality that $j_1 = 0$, where $J = \left(j_1, \ldots, j_{|J|}\right)$ (since if 0 appears somewhere else in J, J may be re-ordered, and if 0 does not appear in J, it may be added). The proof for the part regarding a formula $\varphi(y; z; w; b; a_I)$ is identical, switching the roles of $(a_i)_{i \in \mathbb{Z}}$ and $\left(a'_j\right)_{j \in \mathbb{Z}}$, replacing b' with b', and replacing γ_5 with γ_4.

Case 1. $\varphi\left(x; z; w; b'; a'_J\right)$ is of the form $\boldsymbol{n}_1 \cdot w + \boldsymbol{m}_1 \cdot v\left(g\left(x, b', a'_J\right)\right) \leq \boldsymbol{n}_2 \cdot w + \boldsymbol{m}_2 \cdot v\left(h\left(x, b', a'_J\right)\right)$.

It is enough to show that for any polynomial $f(x, q, y)$ (with $|x| = |I_1|, |y| = |J|, |q| = 1$), we have $v\left(f\left(a_{I_1}, b', a'_J\right)\right) = v\left(f\left(a_{I_2}, b', a'_J\right)\right)$, because then

$$\begin{gathered} \boldsymbol{m}_1 \cdot v\left(g\left(a_{I_1}, b', a'_J\right)\right) = \boldsymbol{m}_1 \cdot v\left(g\left(a_{I_2}, b', a'_J\right)\right) \text{ and} \\ \boldsymbol{m}_2 \cdot v\left(h\left(a_{I_1}, b', a'_J\right)\right) = \boldsymbol{m}_2 \cdot v\left(h\left(a_{I_2}, b', a'_J\right)\right), \text{ so} \\ \models \boldsymbol{n}_1 \cdot d + \boldsymbol{m}_1 \cdot v\left(g\left(a_{I_1}, b', a'_J\right)\right) \leq \boldsymbol{n}_2 \cdot d + \boldsymbol{m}_2 \cdot v\left(h\left(a_{I_1}, b', a'_J\right)\right) \iff \\ \models \boldsymbol{n}_1 \cdot d + \boldsymbol{m}_1 \cdot v\left(g\left(a_{I_2}, b', a'_J\right)\right) \leq \boldsymbol{n}_2 \cdot d + \boldsymbol{m}_2 \cdot v\left(h\left(a_{I_2}, b', a'_J\right)\right). \end{gathered}$$

Given a polynomial $f(x, q, y)$, let

$$f^*(u, v, x, y, q) := f\left(x_1 + u, \ldots, x_{|x|} + u, q + y_1 + v, y_1 + v, \ldots, y_{|y|} + v\right),$$

with $|u| = |v| = |q| = 1$, $|x| = |I_1|$, $|y| = |J|$, so that

$$f^*\left(a_\infty, a'_\infty, \alpha \cdot \operatorname{lift}\left(\tilde{a}_{I_i}\right), \beta \cdot \operatorname{lift}\left(\tilde{a}'_J\right), b' - a'_0\right) = f\left(a_{I_i}, b', a'_J\right)$$

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

for $i \in \{1,2\}$ (using that $a_i = a_\infty + \alpha \cdot \operatorname{lift}(\tilde{a}_i)$ and $a'_j = a'_\infty + \beta \cdot \operatorname{lift}\left(\tilde{a}'_j\right)$ and $j_1 = 0$).

$$\text{Since } v\left(n \cdot a_\infty^{r_\infty} \left(\alpha \operatorname{lift}\left(\tilde{a}_{i_1}\right)\right)^{r_1} \dots \left(\alpha \operatorname{lift}\left(\tilde{a}_{i_{|x|}}\right)\right)^{r_{|x|}} \left(a'_\infty\right)^{s_\infty} \cdot \right.$$
$$\left. \cdot \left(\beta \operatorname{lift}\left(\tilde{a}'_{j_1}\right)\right)^{s_1} \dots \left(\beta \operatorname{lift}\left(\tilde{a}'_{j_{|y|}}\right)\right)^{s_{|y|}} \left(b' - a'_0\right)^t \right)$$
$$= v(n) + r_\infty \gamma_0 + s_\infty \gamma_1 + \left(r_1 + \dots + r_{|x|}\right)\gamma_2 + \left(s_1 + \dots + s_{|y|}\right)\gamma_3 + t\gamma_5,$$

regardless of $i_1, \dots, i_{|x|}$, if we let

$$n \cdot u^{r_\infty} v^{s_\infty} x_1^{r_1} \dots x_{|x|}^{r_{|x|}} y_1^{s_1} \dots y_{|y|}^{s_{|y|}} q^t$$

be a monomial in $f^*(u, v, x, y, q)$ minimizing

$$v(n) + r_\infty \gamma_0 + s_\infty \gamma_1 + \left(r_1 + \dots + r_{|x|}\right)\gamma_2 + \left(s_1 + \dots + s_{|y|}\right)\gamma_3 + t\gamma_5,$$

then by Claim 5.9,

$$\begin{aligned} v\left(f\left(a_{I_i}, b', a'_J\right)\right) &= v\left(f^*\left(a_\infty, a'_\infty, \alpha \cdot \operatorname{lift}\left(\tilde{a}_{I_i}\right), \beta \cdot \operatorname{lift}\left(\tilde{a}'_J\right), b' - a'_0\right)\right) \\ &= v(n) + r_\infty \gamma_0 + s_\infty \gamma_1 + \left(r_1 + \dots + r_{|x|}\right)\gamma_2 + \left(s_1 + \dots + s_{|y|}\right)\gamma_3 + t\gamma_5 \end{aligned}$$

for $i \in \{1,2\}$.

Case 2. $\varphi\left(x; z; w; b'; a'_J\right)$ is of the form $f\left(z, \operatorname{ac}\left(g\left(x, b, a'_J\right)\right)\right) =_k 0$.

It is enough to show that $f(\sigma(c), \operatorname{ac}(g(a_{I_2}, b', a'_J))) = \sigma(f(c, \operatorname{ac}(g(a_{I_1}, b', a'_J))))$, for which it is in turn enough to show that $\operatorname{ac}(g(a_{I_2}, b', a'_J)) = \sigma(\operatorname{ac}(g(a_{I_1}, b', a'_J)))$. Since $a_i = a_\infty + \alpha \cdot \operatorname{lift}(\tilde{a}_i)$ and $a'_j = a'_\infty + \beta \cdot \operatorname{lift}\left(\tilde{a}'_j\right)$, there is a polynomial $h(u, v, x, y, q)$ (with $|u| = |v| = |q| = 1$, $|x| = |I_1|$, $|y| = |J|$) such that

$$h\left(a_\infty, a'_\infty, \alpha \cdot \operatorname{lift}\left(\tilde{a}_{I_i}\right), \beta \cdot \operatorname{lift}\left(\tilde{a}'_J\right), b' - a'_0\right) = g\left(a_{I_i}, b', a'_J\right)$$

for $i \in \{1,2\}$. As in the proof of Case 1, there are $n, m_0, m_1, m_2, m_3, m_5 \in \mathbb{N}$ such that

$$\begin{aligned} &v\left(h\left(a_\infty, a'_\infty, \alpha \cdot \operatorname{lift}\left(\tilde{a}_{I_i}\right), \beta \cdot \operatorname{lift}\left(\tilde{a}'_J\right), b' - a'_0\right)\right) \\ &= v(n) + m_0\gamma_0 + m_1\gamma_1 + m_2\gamma_2 + m_3\gamma_3 + m_5\gamma_5 \end{aligned}$$

for $i \in \{1,2\}$. Let $\tilde{h}(u, v, x, y, q)$ be the sum of monomials in h with degree m_0 in u, degree m_1 in v, total degree m_2 in x, total degree m_3 in y, degree m_5 in q, and whose coefficient has valuation $v(n)$. That way

$$\begin{aligned} &v\left(\tilde{h}\left(a_\infty, a'_\infty, \alpha \cdot \operatorname{lift}\left(\tilde{a}_{I_i}\right), \beta \cdot \operatorname{lift}\left(\tilde{a}'_J\right), b' - a'_0\right)\right) \\ &= v(n) + m_0\gamma_0 + m_1\gamma_1 + m_2\gamma_2 + m_3\gamma_3 + m_5\gamma_5, \text{ and} \\ &v\left(\left(h - \tilde{h}\right)\left(a_\infty, a'_\infty, \alpha \cdot \operatorname{lift}\left(\tilde{a}_{I_i}\right), \beta \cdot \operatorname{lift}\left(\tilde{a}'_J\right), b' - a'_0\right)\right) \\ &> v(n) + m_0\gamma_0 + m_1\gamma_1 + m_2\gamma_2 + m_3\gamma_3 + m_5\gamma_5. \end{aligned}$$

Then $h^*(x,y) := \frac{\tilde{h}(u,v,x,y,q)}{n\cdot u^{m_0}v^{m_1}q^{m_5}}$ is a non-zero polynomial in x, y, all coefficients having valuation 0, so it reduces under the residue map to a nonzero polynomial in x, y. Since $\tilde{a}_{I_i}, \tilde{a}'_J$ are algebraically independent (by indiscernibility), it follows that $\overline{h^*}\left(\tilde{a}_{I_i}, \tilde{a}'_J\right) \neq 0$, so $h^*\left(\operatorname{lift}\left(\tilde{a}_{I_i}\right), \operatorname{lift}\left(\tilde{a}'_J\right)\right)$ has valuation 0, and hence its angular component is its residue, $\overline{h^*}\left(\tilde{a}_{I_i}, \tilde{a}'_J\right)$. We have

$$\overline{h^*}\left(\tilde{a}_{I_2}, \tilde{a}'_J\right) = \overline{h^*}\left(\sigma\left(\tilde{a}_{I_1}\right), \sigma\left(\tilde{a}'_J\right)\right) = \sigma\left(\overline{h^*}\left(\tilde{a}_{I_1}, \tilde{a}'_J\right)\right), \text{ thus}$$

$$\begin{aligned}
&\operatorname{ac}\left(\tilde{h}\left(a_\infty, a'_\infty, \alpha\cdot\operatorname{lift}\left(\tilde{a}_{I_2}\right), \beta\cdot\operatorname{lift}\left(\tilde{a}'_J\right), b'-a'_0\right)\right)\\
&= \operatorname{ac}\left(n\cdot a_\infty^{m_0}\left(a'_\infty\right)^{m_1}\alpha^{m_2}\beta^{m_3}\left(b'-a'_0\right)^{m_5}\right)\operatorname{ac}\left(\frac{\tilde{h}\left(a_\infty, a'_\infty, \operatorname{lift}\left(\tilde{a}_{I_2}\right), \operatorname{lift}\left(\tilde{a}'_J\right), b'-a'_0\right)}{n\cdot a_\infty^{m_0}\left(a'_\infty\right)^{m_1}\left(b'-a'_0\right)^{m_5}}\right)\\
&= \operatorname{ac}(n)\left(\tilde{a}'_0-\tilde{b}'\right)^{m_5}\overline{h^*}\left(\tilde{a}_{I_2}, \tilde{a}'_J\right) = \operatorname{ac}(n)\left(\tilde{a}'_0-\tilde{b}'\right)^{m_5}\sigma\left(\overline{h^*}\left(\tilde{a}_{I_1}, \tilde{a}'_J\right)\right)\\
&= \sigma\left(\operatorname{ac}(n)\left(\tilde{a}'_0-\tilde{b}'\right)^{m_5}\overline{h^*}\left(\tilde{a}_{I_1}, \tilde{a}'_J\right)\right)\\
&= \sigma\Bigg(\operatorname{ac}\left(n\cdot a_\infty^{m_0}\left(a'_\infty\right)^{m_1}\alpha^{m_2}\beta^{m_3}\left(b'-a'_0\right)^{m_5}\right)\cdot\\
&\qquad\cdot\operatorname{ac}\left(\frac{\tilde{h}\left(a_\infty, a'_\infty, \operatorname{lift}\left(\tilde{a}_{I_1}\right), \operatorname{lift}\left(\tilde{a}'_J\right), b'-a'_0\right)}{n\cdot a_\infty^{m_0}\left(a'_\infty\right)^{m_1}\left(b'-a'_0\right)^{m_5}}\right)\Bigg)\\
&= \sigma\left(\operatorname{ac}\left(\tilde{h}\left(a_\infty, a'_\infty, \alpha\cdot\operatorname{lift}\left(\tilde{a}_{I_1}\right), \beta\cdot\operatorname{lift}\left(\tilde{a}'_J\right), b'-a'_0\right)\right)\right).
\end{aligned}$$

$$\begin{aligned}
\text{Since } & v\left(\left(h-\tilde{h}\right)\left(a_\infty, a'_\infty, \alpha\cdot\operatorname{lift}\left(\tilde{a}_{I_i}\right), \beta\cdot\operatorname{lift}\left(\tilde{a}'_J\right), b'-a'_0\right)\right)\\
&> v\left(h\left(a_\infty, a'_\infty, \alpha\cdot\operatorname{lift}\left(\tilde{a}_{I_i}\right), \beta\cdot\operatorname{lift}\left(\tilde{a}'_J\right), b'-a'_0\right)\right),\\
\text{we have } & \operatorname{ac}\left(g\left(a_{I_i}, b', a'_J\right)\right) = \operatorname{ac}\left(h\left(a_\infty, a'_\infty, \alpha\cdot\operatorname{lift}\left(\tilde{a}_{I_i}\right), \beta\cdot\operatorname{lift}\left(\tilde{a}'_J\right), b'-a'_0\right)\right)\\
&= \operatorname{ac}\left(\tilde{h}\left(a_\infty, a'_\infty, \alpha\cdot\operatorname{lift}\left(\tilde{a}_{I_i}\right), \beta\cdot\operatorname{lift}\left(\tilde{a}'_J\right), b'-a'_0\right)\right); \text{ hence}
\end{aligned}$$

$$\operatorname{ac}\left(g\left(a_{I_2}, b', a'_J\right)\right) = \operatorname{ac}\left(\tilde{h}\left(a_\infty, a'_\infty, \alpha\cdot\operatorname{lift}\left(\tilde{a}_{I_2}\right), \beta\cdot\operatorname{lift}\left(\tilde{a}'_J\right), b-a'_0\right)\right)$$
$$= \sigma\left(\operatorname{ac}\left(\tilde{h}\left(a_\infty, a'_\infty, \alpha\cdot\operatorname{lift}\left(\tilde{a}_{I_1}\right), \beta\cdot\operatorname{lift}\left(\tilde{a}'_J\right), b'-a'_0\right)\right)\right) = \sigma\left(\operatorname{ac}\left(g\left(a_{I_1}, b', a'_J\right)\right)\right).$$

Case 3. Clear.

Case 4. $\varphi\left(x; z; w; b'; a'_J\right)$ is of the form $\exists u\, \psi\left(x; z, u; w; b'; a'_J\right)$, where u is a variable of sort k, and the claim holds for ψ. If $\models \varphi\left(a_{I_1}; c; d; b'; a'_J\right)$, then $\models \psi\left(a_{I_1}; c, e; d; b'; a'_J\right)$ for some $e \in k$. Then we have $\models \psi\left(a_{I_2}; \sigma(c), \sigma(e); d; b'; a'_J\right)$, so $\models \varphi\left(a_{I_2}; \sigma(c); d; b'; a'_J\right)$.

Case 5. $\varphi\left(x; z; w; b'; a'_J\right)$ is of the form $\exists u\, \psi\left(x; z; w, u; b'; a'_J\right)$, where u is a variable of sort Γ_∞, and the claim holds for ψ. If $\models \varphi\left(a_{I_1}; c; d; b'; a'_J\right)$, then $\models \psi\left(a_{I_1}; c; d, e; b'; a'_J\right)$ for some $e \in \Gamma_\infty$. Then we have $\models \psi\left(a_{I_2}; \sigma(c); d, e; b'; a'_J\right)$, so $\models \varphi\left(a_{I_2}; \sigma(c); d; b'; a'_J\right)$. This concludes the proof of Claim 5.7.

5.5. Proof of Claim 5.8. Let $\varphi(x; y; z; w; b; b'), \sigma_i, \sigma'_j, \pi, \tau$ be as in the claim.

Case 1. $\varphi(x; y; z; w; b; b')$ is of the form $\boldsymbol{n}_1\cdot w + \boldsymbol{m}_1\cdot v(g(x,y,b,b')) \leq \boldsymbol{n}_2\cdot w + \boldsymbol{m}_2\cdot v(h(x,y,b,b'))$. It is enough to show that for any polynomial

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

$f(x, y, u, u')$,

$$\pi^{-1}\left(v\left(f\left(a_0, a_0', b, b'\right)\right)\right) = v\left(f\left(a_i, a_0', b, b'\right)\right) = \tau^{-1}\left(v\left(f\left(a_i, a_j', b, b'\right)\right)\right)$$

for $i, j \neq 0$, because then

$$\boldsymbol{n}_1 \cdot \pi(d) + \boldsymbol{m}_1 \cdot v\left(g\left(a_0, a_0', b, b'\right)\right) = \pi\left(\boldsymbol{n}_1 \cdot d + \boldsymbol{m}_1 \cdot v\left(g\left(a_i, a_0', b, b'\right)\right)\right) \text{ and}$$
$$\boldsymbol{n}_1 \cdot \tau(d) + \boldsymbol{m}_1 \cdot v\left(g\left(a_i, a_j', b, b'\right)\right) = \tau\left(\boldsymbol{n}_1 \cdot d + \boldsymbol{m}_1 \cdot v\left(g\left(a_i, a_0', b, b'\right)\right)\right)$$

for $i, j \neq 0$, and likewise for $\boldsymbol{n}_2, \boldsymbol{m}_2, h$, and, as π and τ preserve order, this implies

$$\models \varphi\left(a_0; a_0'; \sigma_i(c); \pi(d); b; b'\right) \iff \models \varphi\left(a_i; a_0'; c; d; b; b'\right)$$
$$\iff \models \varphi\left(a_i; a_j'; \sigma_j'(c); \tau(d); b; b'\right).$$

To show this, let $f^*(x, y, u, v) := f(x+u, y+v, u, v)$. By Claim 5.6 and the choice of these elements, for $i, j \in \mathbb{Z}$, the valuations of $a_i - b, a_j' - b', b$, and b' are $\mathbb{Z}$-linearly independent (together with $v(p)$ if the characteristic is mixed), and hence these are valuationally independent. Let $nx^{e_1}y^{e_2}u^{e_3}v^{e_4}$ be the monomial in $f^*(x, y, u, v)$ minimizing $v(n) + e_1\gamma_2 + e_2\gamma_5 + e_3\gamma_0 + e_4\gamma_1$, so that by valuational independence,

$$v\left(f^*\left(a_i - b, a_0' - b', b, b'\right)\right) = v(n) + e_1\gamma_2 + e_2\gamma_5 + e_3\gamma_0 + e_4\gamma_1.$$

This monomial is unique by linear independence (Claim 5.6). Since π and τ preserve order, this monomial also minimizes

$$\pi\left(v(n) + e_1\gamma_2 + e_2\gamma_5 + e_3\gamma_0 + e_4\gamma_1\right) = v(n) + e_1\gamma_4 + e_2\gamma_5 + e_3\gamma_0 + e_4\gamma_1$$
$$= v\left(f^*\left(a_0 - b, a_0' - b', b, b'\right)\right), \text{ and}$$
$$\tau\left(v(n) + e_1\gamma_2 + e_2\gamma_5 + e_3\gamma_0 + e_4\gamma_1\right) = v(n) + e_1\gamma_2 + e_2\gamma_3 + e_3\gamma_0 + e_4\gamma_1$$
$$= v\left(f^*\left(a_i - b, a_j' - b', b, b'\right)\right) \text{ for } i, j \neq 0.$$

Case 2. $\varphi(x; y; z; w; b; b')$ is of the form $f\left(z, \operatorname{ac}\left(g\left(x, y, b, b'\right)\right)\right) =_k 0$.
It is enough to show that

$$\sigma_i\left(f\left(c, \operatorname{ac}\left(g\left(a_i, a_0', b, b'\right)\right)\right)\right) = f\left(\sigma_i(c), \operatorname{ac}\left(g\left(a_0, a_0', b, b'\right)\right)\right) \text{ and}$$
$$\sigma_j'\left(f\left(c, \operatorname{ac}\left(g\left(a_i, a_0', b, b'\right)\right)\right)\right) = f\left(\sigma_j'(c), \operatorname{ac}\left(g\left(a_i, a_j', b, b'\right)\right)\right),$$

for which it is in turn enough to show that

$$\sigma_i\left(\operatorname{ac}\left(g\left(a_i, a_0', b, b'\right)\right)\right) = \operatorname{ac}\left(g\left(a_0, a_0', b, b'\right)\right) \text{ and}$$
$$\sigma_j'\left(\operatorname{ac}\left(g\left(a_i, a_0', b, b'\right)\right)\right) = \operatorname{ac}\left(g\left(a_i, a_j', b, b'\right)\right).$$

Let $h(x, y, u, v) := g(x+u, y+v, u, v)$. Let $nx^{m_1}y^{m_2}u^{m_3}v^{m_4}$ be the (unique, by Claim 5.6) monomial in $h(x, y, u, v)$ minimizing $v(n) + m_1\gamma_2 + m_2\gamma_5 + m_3\gamma_0 + m_4\gamma_1$, so that by valuational independence, $v\left(h\left(a_i - b, a_0' - b', b, b'\right)\right) = v(n) + m_1\gamma_2 + m_2\gamma_5 + m_3\gamma_0 + m_4\gamma_1$. Since π and τ preserve order, this monomial also minimizes $v(n) + m_1\gamma_4 + m_2\gamma_5 + m_3\gamma_0 + m_4\gamma_1$ and $v(n) + m_1\gamma_2 + m_2\gamma_3 + m_3\gamma_0 + m_4\gamma_1$.

$$\text{For } i \neq 0,\ \ \mathrm{ac}\left(n\left(a_i-b\right)^{m_1}\left(a_0'-b'\right)^{m_2}b^{m_3}\left(b'\right)^{m_4}\right)$$
$$=\mathrm{ac}\left(n\right)\left(\mathrm{ac}\left(\alpha\right)\left(\tilde{a}_i-\tilde{a}_0\right)\right)^{m_1}\left(\tilde{b}'-\tilde{a}_0'\right)^{m_2}\mathrm{ac}\left(a_\infty\right)^{m_3}\mathrm{ac}\left(a_\infty'\right)^{m_4}$$
$$=\mathrm{ac}\left(n\right)\left(\tilde{a}_i-\tilde{a}_0\right)^{m_1}\left(\tilde{b}'-\tilde{a}_0'\right)^{m_2}.$$
$$\text{Similarly,}\ \ \mathrm{ac}\left(n\left(a_0-b\right)^{m_1}\left(a_0'-b'\right)^{m_2}b^{m_3}\left(b'\right)^{m_4}\right)$$
$$=\mathrm{ac}\left(n\right)\left(\tilde{b}-\tilde{a}_0\right)^{m_1}\left(\tilde{b}'-\tilde{a}_0'\right)^{m_2}\mathrm{ac}\left(a_\infty\right)^{m_3}\mathrm{ac}\left(a_\infty'\right)^{m_4}$$
$$=\mathrm{ac}\left(n\right)\left(\tilde{b}-\tilde{a}_0\right)^{m_1}\left(\tilde{b}'-\tilde{a}_0'\right)^{m_2}=\sigma_i\left(\mathrm{ac}\left(n\right)\left(\tilde{a}_i-\tilde{a}_0\right)^{m_1}\left(\tilde{b}'-\tilde{a}_0'\right)^{m_2}\right).$$
$$\text{And for } i,j\neq 0 \text{ we have}\ \ \mathrm{ac}\left(n\left(a_i-b\right)^{m_1}\left(a_j'-b'\right)^{m_2}b^{m_3}\left(b'\right)^{m_4}\right)$$
$$=\mathrm{ac}\left(n\right)\left(\mathrm{ac}\left(\alpha\right)\left(\tilde{a}_i-\tilde{a}_0\right)\right)^{m_1}\left(\mathrm{ac}\left(\beta\right)\left(\tilde{a}_j'-\tilde{a}_0'\right)\right)^{m_2}\mathrm{ac}\left(a_\infty\right)^{m_3}\mathrm{ac}\left(a_\infty'\right)^{m_4}$$
$$=\mathrm{ac}\left(n\right)\left(\tilde{a}_i-\tilde{a}_0\right)^{m_1}\left(\tilde{a}_j'-\tilde{a}_0'\right)^{m_2}=\sigma_j'\left(\mathrm{ac}\left(n\right)\left(\tilde{a}_i-\tilde{a}_0\right)^{m_1}\left(\tilde{b}'-\tilde{a}_0'\right)^{m_2}\right).$$
$$\text{Since}\ \ v\left(h\left(a_i-b,a_0'-b',b,b'\right)-n\left(a_i-b\right)^{m_1}\left(a_0'-b'\right)^{m_2}b^{m_3}\left(b'\right)^{m_4}\right)$$
$$>v\left(h\left(a_i-b,a_0'-b',b,b'\right)\right),\ \text{we have}$$
$$\mathrm{ac}\left(h\left(a_i-b,a_0'-b',b,b'\right)\right)=\mathrm{ac}\left(n\left(a_i-b\right)^{m_1}\left(a_0'-b'\right)^{m_2}b^{m_3}\left(b'\right)^{m_4}\right)$$
$$=\mathrm{ac}\left(n\right)\left(\tilde{a}_i-\tilde{a}_0\right)^{m_1}\left(\tilde{b}'-\tilde{a}_0'\right)^{m_2}.$$
$$\text{Likewise,}\ \ v\left(h\left(a_0-b,a_0'-b',b,b'\right)-n\left(a_0-b\right)^{m_1}\left(a_0'-b'\right)^{m_2}b^{m_3}\left(b'\right)^{m_4}\right)$$
$$>v\left(h\left(a_0-b,a_0'-b',b,b'\right)\right),\ \text{so}$$
$$\mathrm{ac}\left(h\left(a_0-b,a_0'-b',b,b'\right)\right)=\mathrm{ac}\left(n\left(a_0-b\right)^{m_1}\left(a_0'-b'\right)^{m_2}b^{m_3}\left(b'\right)^{m_4}\right)$$
$$=\sigma_i\left(\mathrm{ac}\left(n\right)\left(\tilde{a}_i-\tilde{a}_0\right)^{m_1}\left(\tilde{b}'-\tilde{a}_0'\right)^{m_2}\right).$$
$$\text{And}\ \ v\left(h\left(a_i-b,a_j'-b',b,b'\right)-n\left(a_i-b\right)^{m_1}\left(a_j'-b'\right)^{m_2}b^{m_3}\left(b'\right)^{m_4}\right)$$
$$>v\left(h\left(a_i-b,a_j'-b',b,b'\right)\right),\ \text{so}$$
$$\mathrm{ac}\left(h\left(a_i-b,a_j'-b',b,b'\right)\right)=\mathrm{ac}\left(n\left(a_i-b\right)^{m_1}\left(a_j'-b'\right)^{m_2}b^{m_3}\left(b'\right)^{m_4}\right)$$
$$=\sigma_j'\left(\mathrm{ac}\left(n\right)\left(\tilde{a}_i-\tilde{a}_0\right)^{m_1}\left(\tilde{b}'-\tilde{a}_0'\right)^{m_2}\right).$$

Since $g(a_i,a_0',b,b')=h(a_i-b,a_0'-b',b,b')$, $g(a_0,a_0',b,b')=h(a_0-b,a_0'-b',b,b')$, and $g\left(a_i,a_j',b,b'\right)=h\left(a_i-b,a_j'-b',b,b'\right)$, this is what we wanted to show.

Case 3. Clear.

Case 4. $\varphi\left(x;y;z;w;b;b'\right)$ is of the form $\exists u\,\psi\left(x;y;z,u;w;b;b'\right)$, where u is a variable of sort k, and the claim holds for ψ.

For $i,j\neq 0$, if $\models\varphi\left(a_i;a_0';c;d;b;b'\right)$, then $\models\psi\left(a_i;a_0';c,e;d;b;b'\right)$ for some $e\in k$. Then $\models\psi\left(a_0;a_0';\sigma_i\left(c\right),\sigma_i\left(e\right);\pi\left(d\right);b;b'\right)$ and $\models\psi(a_i;a_j';\sigma_j'(c),\sigma_j'(e);\tau(d);b;b')$, so $\models\varphi\left(a_0;a_0';\sigma_i\left(c\right);\pi\left(d\right);b;b'\right)$ and $\models\varphi\left(a_i;a_j';\sigma_j'\left(c\right);\tau\left(d\right);b;b'\right)$. Note that each of these implications is reversible.

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

Case 5. $\varphi(x;y;z;w;b;b')$ is of the form $\exists u\, \psi(x;y;z;w,u;b;b')$, where u is a variable of sort Γ_∞, and the claim holds for ψ. For $i,j \neq 0$, if $\models \varphi\left(a_i;a_0';c;d;b;b'\right)$, then $\models \psi\left(a_i;a_0';c;d,e;b;b'\right)$ for some $e \in \Gamma_\infty$. Then $\models \psi\left(a_0;a_0';\sigma_i(c);\pi(d),\pi(e);b;b'\right)$ and $\models \psi\left(a_i;a_j';\sigma_j'(c);\tau(d),\tau(e);b;b'\right)$; hence, $\models \varphi\left(a_0;a_0';\sigma_i(c);\pi(d);b;b'\right)$ and $\models \varphi\left(a_i;a_j';\sigma_j'(c);\tau(d);b;b'\right)$. Since π and τ are bijective, each of these implications is reversible. This concludes the proof of Claim 5.8, and hence of Theorem 5.1.

5.6. Some further applications of Theorem 5.1 and examples.

REMARK 5.10. Our proof of Theorem 5.1 also applies to any reduct of an ac-valued field K whose residue field has a non-constant totally indiscernible sequence to a language $\mathcal{L} \subseteq \mathcal{L}_{\text{Denef-Pas}}$ such that $\mathcal{L}$ contains the relation $v(x_1 - y_1) < v(x_2 - y_2)$, and every $\mathcal{L}$-formula is equivalent to a Boolean combination of $\mathcal{L}_{\text{Denef-Pas}}$-formulas with no quantifiers of the main sort. This gives us further examples of NIP theories that are not weakly semi-equational, such as:

(1) a Henselian valued field of equicharacteristic 0 whose residue field is algebraically closed;
(2) an algebraically closed valued field (of any characteristic);
(3) the reduct of either of the above to a valued vector space or valued abelian group;
(4) a generic abstract ultrametric space: a two-sorted structure $(\mathcal{M}, \Gamma_\infty)$, with a linear order $\leq$ on Γ_∞ that is dense with maximal element $\infty \in \Gamma_\infty$ and no minimal element, and a function $v : \mathcal{M}^2 \to \Gamma_\infty$, such that $v(x,y) = \infty \iff x = y$, $v(x,y) = v(y,x)$, and $v(x,z) \geq \max(v(x,y), v(y,z))$, and such that for every $\gamma \in \Gamma$ and $a \in \mathcal{M}$, there are $(b_i)_{i \in \mathbb{N}}$ in $\mathcal{M}$ such that $v(a,b_i) = v(b_i,b_j) = \gamma$ for $i,j \in \mathbb{N}$.

EXAMPLE 5.11. Let K be a valued field (viewed as a structure in the language of rings with a predicate for the valuation ring $\mathcal{O}$), $d \in \omega$ and let $\mathcal{F}$ be the family of all convex subsets of K^d in the sense of Monna (equivalently, the family of all translates of $\mathcal{O}$-submodules of K^d). Then $\mathcal{F}$ is a definable family, and a formula defining it is a semi-equation by [8, Theorem 4.3] and Proposition 2.16.

PROBLEM 5.12. Is the field $\mathbb{Q}_p$ semi-equational? It is weakly semi-equational by distality.

§6. Weak semi-equationality in expansions by a predicate.

6.1. Context. We recall the setting and some results from [9] (as usual, below x, y, z denote arbitrary finite tuples of variables). We start with a theory T in a language $\mathcal{L}$, and let $\mathcal{L}_{\mathrm{P}} := \mathcal{L} \cup \{\mathrm{P}(x)\}$, where P is a new unary predicate. Let $T_{\mathrm{P}} := \mathrm{Th}_{\mathcal{L}_{\mathrm{P}}}(M, A)$, where A is some subset of M (interpreted as P). We fix some monster model $(M', A') \succ (M, A)$ of T_{P}. An $\mathcal{L}_{\mathrm{P}}$-formula $\psi(x)$ is *bounded* if it is of the form $Q_0 y_0 \in \mathrm{P} \ldots Q_n y_n \in \mathrm{P}\, \varphi(x,y)$, where $Q_i \in \{\exists, \forall\}$ and $\varphi(x,y) \in \mathcal{L}$. We denote the set of all bounded $\mathcal{L}_{\mathrm{P}}$-formulas by $\mathcal{L}_{\mathrm{P}}^{\mathrm{bdd}}$ and say that the theory T_{P} is *bounded* if every $\mathcal{L}_{\mathrm{P}}$ formula is equivalent modulo T_{P} to a bounded one. Finally, for

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

$\mathcal{L} \subseteq \mathcal{L}' \subseteq \mathcal{L}_{\mathrm{P}}(M)$ we denote by $A_{\mathrm{ind}(\mathcal{L}')}$ the $\mathcal{L}'(\emptyset)$-induced structure on A, i.e., the structure $\left(A; \left(R_{\varphi}(x)\right)_{\varphi(x) \in \mathcal{L}'}\right)$ with $R_{\varphi} := \left\{a \in A^{|x|} : (M, A) \models \varphi(a)\right\}$.

REMARK 6.1. (1) The structures $A_{\mathrm{ind}\left(\mathcal{L}_{\mathrm{P}}^{\mathrm{bdd}}\right)}$ and $A_{\mathrm{ind}(\mathcal{L})}$ have the same definable subsets of A^n, for all $n \in \omega$. Indeed, given $\psi(x) = Q_0 y_0 \in \mathrm{P} \ldots Q_n y_n \in \mathrm{P}\,\varphi(x, y)$ with $\varphi(x, y) \in \mathcal{L}_{\mathrm{P}}^{\mathrm{bdd}}$. Then $R_{\psi}(A)$ can be defined in $A_{\mathrm{ind}(\mathcal{L})}$ by $Q_0 y_0 \ldots Q_n y_n R_{\varphi}(x, y)$.

(2) If T_{P} is bounded, then clearly $A_{\mathrm{ind}(\mathcal{L}_{\mathrm{P}})}$ and $A_{\mathrm{ind}\left(\mathcal{L}_{\mathrm{P}}^{\mathrm{bdd}}\right)}$ have the same definable subsets of A^n, for all n.

FACT 6.2. (1) [9, *Corollary* 2.5] *Assume that T is NIP, $A_{\mathrm{ind}(\mathcal{L})}$ is NIP, and T_{P} is bounded. Then T_{P} is NIP.*

(2) [9, *Corollary* 2.6] *In particular, if T is NIP, $A \preceq M$, and T_{P} is bounded, then T_{P} is NIP.*

Some results on preservation of equationality under naming a set by a predicate are obtained in [24]. As pointed out in [18], the exact analog with distality in place of NIP is false:

FACT 6.3 ([18, Theorem 5.1] and the examples after it). *The theory of dense pairs of o-minimal structures expanding a group is not distal (even though it is bounded and the induced structure on the submodel is distal). Their proof shows that the formula $\varphi(x, y) = \neg\exists u \in \mathrm{P}\,(x = u + y)$ is not a weak semi-equation in the theory of dense pairs.*

In this section we show that at least weak semi-equationality of T_{P} can be salvaged. We will need the following properties of indiscernible sequences and definable sets with distal induced structure.

FACT 6.4 [3, Proposition 1.17]. *Let T be NIP, and let D be an $\emptyset$-definable set. Assume that D_{ind} is distal. Let $(c_i : i \in \mathbb{Q})$ be an indiscernible sequence of tuples in $\mathbb{M}$ and let a tuple b from D be given. Assume that $(c_i : i \in \mathbb{Q} \setminus \{0\})$ is indiscernible over b, then $(c_i : i \in \mathbb{Q})$ is indiscernible over b as well.*

LEMMA 6.5. *Assume T is NIP and D is an $\emptyset$-definable set with D_{ind} distal. Let $(a_i : i \in \mathbb{Q})$ be an $\emptyset$-indiscernible sequence, b such that $(a_i : i \in \mathbb{Q} \setminus \{0\})$ is b-indiscernible, and $c \in D$ arbitrary. Then we can find a sequence $(c_i : i \in \mathbb{Q})$ such that:*

- *$a_i c_i \equiv_b a_1 c$ for all $i \in \mathbb{Q} \setminus \{0\}$,*
- *$(a_i c_i : i \in \mathbb{Q})$ is $\emptyset$-indiscernible, and*
- *$(a_i c_i : i \in \mathbb{Q} \setminus \{0\})$ is b-indiscernible.*

PROOF. By b-indiscernibility of $(a_i : i \in \mathbb{Q} \setminus \{0\})$, Ramsey, compactness, and taking automorphisms we can find a sequence $(c_i : i \in \mathbb{Q} \setminus \{0\})$ in D such that $(a_i c_i : i \in \mathbb{Q} \setminus \{0\})$ is b-indiscernible and $a_i c_i \equiv_b a_1 c$ for all $i \neq 0$. It remains to find a $c_0 \in D$ such that $(a_i c_i : i \in \mathbb{Q})$ is $\emptyset$-indiscernible. Let $I \subseteq \mathbb{Q} \setminus \{0\}$ be an arbitrary finite set and let $\bar{a}_0 := (a_i : i \in I)$. Let $\varepsilon > 0$ in $\mathbb{Q}$ be such that $I \subseteq \mathbb{Q} \setminus (-\varepsilon, \varepsilon)$. For each $i \in \mathbb{Q}$, let $a'_i := (a_i, \bar{a}_0)$ and consider the sequence $(a'_i : i \in (-\varepsilon, \varepsilon))$. It is $\emptyset$-indiscernible since the sequence $(a_i : i \in \mathbb{Q})$ is, and moreover $(a'_i : i \in (-\varepsilon, \varepsilon) \setminus \{0\})$ is indiscernible over $(c_i : i \in I) \subseteq D$ (since the sequence of pairs $(a_i c_i : i \in \mathbb{Q} \setminus \{0\})$

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

is indiscernible). Then by Fact 6.4 we have that $(a'_i : i \in (-\varepsilon, \varepsilon))$ is indiscernible over $(c_i : i \in I)$. In particular, there exists an automorphism σ sending $a'_{\frac{\varepsilon}{2}}$ to a'_0 and fixing $(c_i : i \in I)$, hence sending $a_{\frac{\varepsilon}{2}}$ to a_0 and fixing $(a_i c_i : i \in I)$. As by assumption $(a_i c_i : i \in I, i <- \varepsilon) + \left(a_{\frac{\varepsilon}{2}} c_{\frac{\varepsilon}{2}}\right) + (a_i c_i : i \in I, i > \varepsilon)$ is indiscernible, applying σ we have that there is $\tilde{c}_0 := \sigma\left(c_{\frac{\varepsilon}{2}}\right) \in D$ such that $(a_i c_i : i \in I, i <- \varepsilon) + (a_0 \tilde{c}_0) + (a_i c_i : i \in I, i > \varepsilon)$ is indiscernible. As I was arbitrary, we can then find c_0 as wanted by compactness. ⊣

Definition 6.6. A theory T_{P} is *almost model complete* if, modulo T_{P}, every $\mathcal{L}_{\mathrm{P}}$-formula $\psi(x)$ is equivalent to a Boolean combination of formulas of the form $\exists y_0 \in \mathrm{P} \dots \exists y_{n-1} \in \mathrm{P}\, \varphi(x, y)$, where $\varphi(x, y)$ is an $\mathcal{L}$-formula.

Theorem 6.7. *Assume that T is distal, $A_{\mathrm{ind}(\mathcal{L})}$ is distal, and T_{P} is almost model complete. Then T_{P} is weakly semi-equational.*

Proof. We know by Fact 6.2 that T_{P} is NIP. As T_{P} is almost model complete, so in particular bounded, by Lemma 6.1(1) and (2) the structures $A_{\mathrm{ind}(\mathcal{L}_{\mathrm{P}})}$ and $A_{\mathrm{ind}(\mathcal{L})}$ have the same definable subsets of A^n, for all n. Hence the full structure induced on P in T_{P} is distal, so Lemma 6.5 can be applied to T_{P} with $D := \mathrm{P}$.

Let (M', A') be a sufficiently saturated elementary extension of $(M, A) \models T_{\mathrm{P}}$. As T_{P} is almost model complete by assumption, it is enough to show that every formula in $\mathcal{L}_{\mathrm{P}}$ of the form

$$\varphi(x, y) = \exists z_0 \in \mathrm{P} \dots \exists z_{n-1} \in \mathrm{P}\, \psi(x, y, z),$$

where $\psi(x, x, z) \in \mathcal{L}$, is a weak semi-equation in T_{P}.

To check Definition 1.2, assume (using Remark 2.1) that we are given an $\mathcal{L}_{\mathrm{P}}$-indiscernible sequence of finite tuples $(a_i : i \in \mathbb{Q})$ and a finite tuple b, both in M', such that the sequence $(a_i : i \in \mathbb{Q} \setminus \{0\})$ is $\mathcal{L}_P(b)$-indiscernible and $\models \varphi(a_i, b)$ for all $i \neq 0$. In particular, there is some tuple c in P such that $\models \psi(a_1, b, c)$ holds. By Lemma 6.5 applied in T_{P} with $D := \mathrm{P}$, it follows that there is a sequence $(c_i : i \in \mathbb{Q})$ with $c_i \in \mathrm{P}$ such that $(a_i c_i : i \in \mathbb{Q})$ is $\mathcal{L}_{\mathrm{P}}$-indiscernible, $(a_i c_i : i \neq 0)$ is $\mathcal{L}_{\mathrm{P}}(b)$-indiscernible and $a_i c_i \equiv_b^{\mathcal{L}_{\mathrm{P}}} a_1 c$ for $i \neq 0$. In particular $\models \psi(a_i, b, c_i)$ for $i \neq 0$. But $\psi'(x, z; y) := \psi(x, y, z) \in \mathcal{L}$ is a semi-equation in T as T is distal, hence $\models \psi(a_0, b, c_0)$, and so $\models \varphi(a_0, b)$ holds—as wanted. ⊣

Corollary 6.8. *Dense pairs of o-minimal structures, as well as the other examples discussed after* [18, *Theorem* 5.1], *are weakly semi-equational.*

Problem 6.9. (1) In Theorem 6.7, can we relax the assumption to "T and $A_{\mathrm{ind}(\mathcal{L})}$ are weakly semi-equational?"

(2) Is there an analog of Theorem 6.7 for semi-equationality? Even a general result for equationality seems to be missing (the argument in [24] for *belles paires* of algebraically closed fields is specific to algebraically closed fields).

Acknowledgements. We thank the referee for many suggestions on improving the paper. We are grateful to Matthias Aschenbrenner, Gabe Conant, Allen Gehret, Anand Pillay, Sergei Starchenko, Erik Walsberg, and Martin Ziegler for some helpful comments and conversations.

Funding. Both authors were partially supported by the NSF CAREER grant DMS-1651321, and Chernikov was additionally supported by a Simons fellowship.

REFERENCES

[1] A. Anderson, *Combinatorial bounds in distal structures*, preprint, 2021, arXiv:2104.07769.

[2] M. Aschenbrenner, A. Dolich, D. Haskell, D. Macpherson, and S. Starchenko, *Vapnik–Chervonenkis density in some theories without the independence property, I*. ***Transactions of the American Mathematical Society***, vol. 368 (2016), no. 8, pp. 5889–5949.

[3] M. Aschenbrenner, A. Chernikov, A. Gehret, and M. Ziegler, *Distality in valued fields and related structures*. ***Transactions of the American Mathematical Society***, vol. 375 (2022), no. 7, pp. 4641–4710.

[4] J. T. Baldwin and J. Saxl, *Logical stability in group theory*. ***Journal of the Australian Mathematical Society***, vol. 21 (1976), no. 3, pp. 267–276.

[5] A. Basit, A. Chernikov, S. Starchenko, T. Tao, and C.-M. Tran, *Zarankiewicz's problem for semilinear hypergraphs*. ***Forum of Mathematics, Sigma***, vol. 9 (2021), p. e59.

[6] A. Chernikov and I. Kaplan, *Forking and dividing in* NTP_2 *theories*, this Journal, vol. 77 (2012), no. 1, 1–20.

[7] A. Chernikov and A. Mennen, *Semi-equational theories.* Extended version of this paper, preprint, 2022, arXiv:2204.13790v1.

[8] ———, *Combinatorial properties of non-archimedean convex sets*. ***Pacific Journal of Mathematics***, 2023, to appear. Preprint available from arXiv:2109.04591.

[9] A. Chernikov and P. Simon, *Externally definable sets and dependent pairs*. ***Israel Journal of Mathematics***, vol. 194 (2013), no. 1, pp. 409–425.

[10] ———, *Externally definable sets and dependent pairs II*. ***Transactions of the American Mathematical Society***, vol. 367 (2015), no. 7, pp. 5217–5235.

[11] A. Chernikov and S. Starchenko, *Regularity lemma for distal structures*. ***Journal of the European Mathematical Society***, vol. 20 (2018), no. 10, pp. 2437–2466.

[12] A. Chernikov, I. Kaplan, and P. Simon, *Groups and fields with* NTP_2. ***Proceedings of the American Mathematical Society***, vol. 143 (2015), no. 1, pp. 395–406.

[13] R. Cluckers and I. Halupczok, *Quantifier elimination in ordered abelian groups*. ***Confluentes Mathematici***, vol. 3 (2011), no. 4, pp. 587–615.

[14] A. Dolich and J. Goodrick, *A characterization of strongly dependent ordered abelian groups*. ***Revista Colombiana de Matemáticas***, vol. 52 (2018), no. 2, pp. 139–159.

[15] R. Farré, *Strong ordered abelian groups and dp-rank*, preprint, 2017, arXiv:1706.05471.

[16] V. Guingona and M. C. Laskowski, *On VC-minimal theories and variants*. ***Archive for Mathematical Logic***, vol. 52 (2013), no. 7, pp. 743–758.

[17] Y. Halevi and A. Hasson, *Strongly dependent ordered abelian groups and Henselian fields*. ***Israel Journal of Mathematics***, vol. 232 (2019), no. 2, pp. 719–758.

[18] P. Hieronymi and T. Nell, *Distal and non-distal pairs*, this Journal, vol. 82 (2017), no. 1, pp. 375–383.

[19] E. Hrushovski and G. Srour, *On stable non-equational theories*, 1989, unpublished manuscript.

[20] M. Junker, *A note on equational theories*, this Journal, vol. 65 (2000), no. 4, pp. 1705–1712.

[21] M. Junker and I. Kraus, *Theories with equational forking*, this Journal, vol. 67 (2002), no. 1, pp. 326–340.

[22] M. Junker and D. Lascar, *The indiscernible topology: A mock Zariski topology*. ***Journal of Mathematical Logic***, vol. 1 (2001), no. 01, pp. 99–124.

[23] J. Loveys and Y. Peterzil, *Linear o-minimal structures*. ***Israel Journal of Mathematics***, vol. 81 (1993), no. 1, pp. 1–30.

[24] A. Martin-Pizarro and M. Ziegler, *Equational theories of fields*, this Journal, vol. 85 (2020), no. 2, pp. 828–851.

[25] ———, *Trois couleurs: A new non-equational theory*. ***Fundamenta Mathematicae***, vol. 254 (2021), pp. 313–333.

https://doi.org/10.1017/jsl.2023.28 Published online by Cambridge University Press

[26] J. Matousek, *Bounded VC-dimension implies a fractional Helly theorem*. ***Discrete & Computational Geometry***, vol. 31 (2004), no. 2, pp. 251–255.

[27] R. Mennuni, *Weakly binary expansions of dense meet-trees*. ***Mathematical Logic Quarterly***, vol. 68 (2022), no. 1, pp. 32–47.

[28] I. Müller and R. Sklinos, *Nonequational stable groups*. ***Proceedings of the American Mathematical Society***, 2023, to appear. Preprint available from arXiv:1703.04169.

[29] A. O'Hara, ***An introduction to equations and equational theories***. Master's thesis, University of Waterloo, 2011.

[30] Y. Peterzil and S. Starchenko, *A trichotomy theorem for o-minimal structures*. ***Proceedings of the London Mathematical Society***, vol. 77 (1998), no. 3, pp. 481–523.

[31] A. Pillay, *Countable models of stable theories*. ***Proceedings of the American Mathematical Society***, vol. 89 (1983), no. 4, pp. 666–672.

[32] ———, ***Geometric Stability Theory***, Oxford Logic Guides, vol. 32, Oxford University Press, Oxford, 1996.

[33] A. Pillay and G. Srour, *Closed sets and chain conditions in stable theories*, this Journal, vol. 49 (1984), no. 4, pp. 1350–1362.

[34] Z Sela, *Free and hyperbolic groups are not equational*, preprint, 2012, arXiv:1204.5075.

[35] P. Simon, *On dp-minimal ordered structures*, this Journal, vol. 76 (2011), no. 2, pp. 448–460.

[36] ———, *Distal and non-distal NIP theories*. ***Annals of Pure and Applied Logic***, vol. 164 (2013), no. 3, pp. 294–318.

[37] G. Srour, *The independence relation in separably closed fields*, this Journal, vol. 51 (1986), no. 3, pp. 715–725.

[38] ———, *The notion of independence in categories of algebraic structures, Part I: Basic properties*. ***Annals of Pure and Applied Logic***, vol. 38 (1988), no. 2, pp. 185–213.

[39] ———, *The notion of independence in categories of algebraic structures, Part II: S-minimal extensions*. ***Annals of Pure and Applied Logic***, vol. 39 (1988), no. 1, pp. 55–73.

[40] ———, *The notion of independence in categories of algebraic structures, Part III: Equational classes*. ***Annals of Pure and Applied Logic***, vol. 47 (1990), no. 3, pp. 269–294.

[41] M. C. Tran and E. Walsberg, *A family of dp-minimal expansions of* $(\mathbb{Z}, +)$, preprint, 2017, arXiv:1711.04390.

[42] E. Walsberg, *Notes on trace equivalence*, preprint, 2021, arXiv:2101.12194.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA, LOS ANGELES
LOS ANGELES, CA 90095, USA
E-mail: chernikov@math.ucla.edu
E-mail: alexmennen@math.ucla.edu