INDISCERNIBLES IN MONADICALLY NIP THEORIES

SAMUEL BRAUNFELD[1] AND MICHAEL C. LASKOWSKI[2]

ABSTRACT. We prove various results around indiscernibles in monadically NIP theories. First, we provide several characterizations of monadic NIP in terms of indiscernibles, mirroring previous characterizations in terms of the behavior of finite satisfiability. Second, we study (monadic) distality in hereditary classes and complete theories. Here, via finite combinatorics, we prove a result implying that every planar graph admits a distal expansion. Finally, we prove a result implying that no monadically NIP theory interprets an infinite group, and note an example of a (monadically) stable theory with no distal expansion that does not interpret an infinite group.

1. INTRODUCTION

In the setting of hereditary classes of relational structures, the standard model-theoretic condition of an NIP theory collapses to the much stronger *monadically NIP* theory, i.e. a theory that remains NIP under arbitrary expansions by unary predicates. Thus monadic NIP is conjectured to be an important dividing line for various combinatorial problems, such as the algorithmic tractability of first-order model checking [13]. The manipulation of indiscernible sequences is a prominent thread in results involving monadic NIP, and this paper consists of three independent sections concerning this. The first section shows that we may provide various characterizations of monadic NIP in terms of indiscernible sequences. The second section describes how monadic NIP interacts with distality, another model-theoretic condition characterized by the behavior of indiscernibles. One motivation for considering distality is that it also has combinatorial implications, such as the strong Erdős-Hajnal property (an improved form of Ramsey's theorem). Another motivation comes from the fact that in monadically NIP theories, every global invariant 1-type is either generically stable or distal, so one might hope to understand monadic NIP by understanding separately how it interacts with stability and with distality. It's interaction with stability leads to the relatively well-understood property of monadic stability, so we begin the study of its interaction with distality here. Last, a recurring theme in

model theory is understanding theories by the algebraic objects, particularly groups, that can appear in them, and in the third section we show that monadically NIP theories can encode almost no algebraic structure.

In the first section, we provide various characterizations of monadic NIP in terms of indiscernible sequences, paralleling the characterizations provided in [7] in terms of finite satisfiability.

Theorem 1.1 (Theorem 2.12). *Let T be a complete theory. Then the following are equivalent.*

(1) T is monadically NIP.
(2) T admits widening of indiscernibles.
(3) T is dp^+-minimal.
(4) T is dp-minimal and has endless indiscernible triviality.

There are seemingly good reasons to desire the characterizations in terms of indiscernibility: finite satisfiability requires working over an infinite model and so seems less suited to working with classes of finite structures, where monadic stability and monadic NIP have been intensively studied. Also, one of the issues encountered in [8] studying monadic NIP in the setting of existentially closed models is that finite satisfiability is not very well-behaved there, since a finitely satisfiable type cannot necessarily be extended to one over a larger base. However, the Erdős-Rado theorem still allows us to obtain and manipulate indiscernible sequences there.

In the second part of this section, we show that dp^+-minimality, as opposed to dp-minimality, ensures that a singleton can cut an indiscernible sequence in at most one place. This clarifies how models of a monadically NIP theory can be decomposed over an indiscernible sequence, which leads to a notion of independence paralleling finite satisfiability.

In the next section, we turn to distality in monadically NIP theories. We begin by characterizing monadic distality in terms of avoiding totally indiscernible sets of singletons (Proposition 3.2) and showing that monadic distality is equivalent to monadic NIP and distality (Corollary 3.3), mirroring a characterization of monadic stability. One of the main points of [8] is that universal theories are the right setting for monadic model-theoretic properties, so we then study (monadic) distality in universal theories and hereditary classes of finite structures. One of the main results here is that every graph class with bounded twin-width admits a distal expansion by a linear order.

Theorem 1.2 (Theorem 3.13). *Let $\mathcal{C}$ be a hereditary class of binary relational structures with bounded twin-width (for example, a planar graph class). Then $\mathcal{C}$ admits a distal expansion by a linear order, and thus so does every model of $Th(\mathcal{C})_\forall$.*

Theorem 3.13 produces distal expansions of infinite structures without having to analyze definability in indiscernible sequences in those structures, instead relying purely on finite combinatorics. As an example corollary, we obtain the following result about distal expansions of infinite graphs, which

on its face is not connected with monadic NIP or with hereditary classes of finite structures.

Corollary 1.3. *Let G be an infinite planar graph. Then G admits a distal expansion by a linear order.*

In the final section, we show that monadically NIP theories cannot interpret an infinite group, although we prove a more general result.

Proposition 1.4 (Proposition 4.3)**.** *Let T be a theory with finite dp-rank and endless indiscernible triviality. Then T does not trace-define an infinite cancellative magma.*

This is part of a project set out in [26] to identify some notion of triviality and a corresponding Zilber dichotomy for NIP theories, and endless indiscernible triviality may be an appropriate such notion, perhaps in a restricted setting.

Returning to distality, the most commonly known reason for an NIP theory to not admit a distal expansion is if it interprets an infinite field of finite characteristic (e.g., see [1]). As a corollary of Proposition 1.4, we record the fact that there is a (monadically) stable theory with no distal expansion that does not interpret an infinite group (Corollary 4.10).

2. Indiscernibles and naming parameters

2.1. **Characterizations of monadic NIP via indiscernibles.** Before stating the relevant definitions, our goal in this subsection is to prove Theorem 1.1. The point of view resulting from this theorem is that we may decompose models of a monadically NIP theory into an indiscernible sequence, rather than into the M-f.s. sequences used in [7]. In particular, by iterating the fact that T admits widening of indiscernibles, it will follow that given any $M \prec \mathfrak{C}$ and Dedekind complete indiscernible sequence $\mathcal{I} = (\bar{a}_i : i \in I)$ in $\mathfrak{C}$, we may "widen" $\mathcal{I}$ to an indiscernible sequence $\mathcal{I}^+ = (\bar{a}_i \bar{b}_i : i \in I)$ such that $M \subset \bigcup \mathcal{I}^+$. This is essentially the point of view given on monadically NIP theories in [4], which also proves that monadically NIP theories admit widening of indiscernibles. Then dp^+-minimality is somewhat analogous to the f.s. dichotomy of [7], describing how single additional point interacts with a partial decomposition. Widening of indiscernibles corresponds to [7, Lemma 3.2] that any partial decomposition can be extended to contain any additional point, and similarly iterating gives the statement that a partial decomposition can always be extended to contain any model. The last characterization is analogous to [7, Lemma 3.3] that dependence is trivial and transitive. Dp-minimality may be more transparently viewed as a transitivity statement via the following characterization: if there are indiscernible sequences $\mathcal{I}, \mathcal{J}$ and a singleton b such that neither $\mathcal{I}$ nor $\mathcal{J}$ are indiscernible over b, then $\mathcal{I}$ and $\mathcal{J}$ are not mutually indiscernible.

We now define the notions appearing in Theorem 1.1. All indiscernible sequences will be assumed infinite. One can view dp-minimality as describing

how an indiscernible sequence $\mathcal{I} = (\bar{a}_i : i \in I)$ can be 'injured' by the addition of parameters.

Definition 2.1. A theory T is *dp-minimal* if for every indiscernible sequence $\mathcal{I} = (\bar{a}_i : i \in \mathbb{Q})$ and every singleton b, there is $i^* \in \mathbb{R} \cup \{ \pm\infty \}$ such that $\mathcal{I}_{<i^*}$ and $\mathcal{I}_{>i^*}$ are mutually indiscernible over b.

A theory T is *dp^+-minimal* if we may instead choose $i^* \in \mathbb{R} \cup \{ \pm\infty \}$ so that $\mathcal{I}_{<i^*}$ and $\mathcal{I}_{>i^*}$ are mutually indiscernible over $\bar{a}_{i^*}b$ if $i^* \in \mathbb{Q}$, and over b otherwise.

The definition of dp-minimality above is equivalent to the characterization given in [24, Theorem 4.18 $(iii)_2$] (with $\kappa = 2$ and p the partial type $\{ x = x \}$). The characterization works with an arbitrary infinite index set and allows one of $\mathcal{I}_{<i^*}$ and $\mathcal{I}_{>i^*}$ to be finite, in which case we only demand that the other is indiscernible over b. These are clearly equivalent on $\mathbb{Q}$-indexed sequences, so the characterization implies our definition. It suffices to show our definition implies the characterization on countable sequences, which can always be extended to $\mathbb{Q}$-indexed sequences, where our definition applies and produces a cut i^* that also works for the characterization.

The name dp^+-minimality is based on the notion of strongly$^+$-dependent theories introduced in [23] (see Conclusion 2.11 there) and of dp^+-rank mentioned in [24, Remark 4.19].

Definition 2.2. A theory T *admits widening of indiscernibles* if for every infinite indiscernible sequence $\mathcal{I} = (\bar{a}_i : i \in I)$ and every singleton $b \in \mathfrak{C}$, there is a one-point extension $(J, \leq) \supseteq (I, \leq)$, an extension $(\bar{a}_j : j \in J)$ of $\mathcal{I}$, an element $j^* \in J$ and a sequence $(b_j : j \in J)$ such that $b_{j^*} = b$ and $(\bar{a}_j b_j : j \in J)$ is indiscernible.

Definition 2.3. A theory T has *endless indiscernible triviality* if for every indiscernible sequence $\mathcal{I} = (\bar{a}_i : i \in I)$ without endpoints, and for every set B, if $\mathcal{I}$ is indiscernible over each $b \in B$ then $\mathcal{I}$ is indiscernible over B.

A theory has *indiscernible triviality* if the statement above holds for arbitrary infinite indiscernible sequences (possibly with endpoints).

We recall the f.s. dichotomy and some related notions from [7].

Definition 2.4. T has the *f.s. dichotomy* if, for all models M, all finite tuples $\bar{a}, \bar{b} \in \mathfrak{C}$, if $\mathrm{tp}(\bar{b}/M\bar{a})$ is finitely satisfied in M, then for any singleton c, either $\mathrm{tp}(\bar{b}/M\bar{a}c)$ or $\mathrm{tp}(\bar{b}c/M\bar{a})$ is finitely satisfied in M.

Definition 2.5. Let M be a model, let $C \supseteq M$, and let $(I, \leq)$ be any linearly ordered index set.

- Suppose $\langle A_i : i \in I \rangle$ is any sequence of sets, indexed by $(I, \leq)$. For $J \subseteq I$, A_J denotes $\bigcup_{j \in J} A_j$, and for $i^* \in I$, $A_{<i^*}$ denotes $\bigcup_{i < i^*} A_i$. $A_{\leq i^*}$ and $A_{>i^*}$ are defined analogously.
- For $C \supseteq M$, an *M-f.s. sequence over C*, is a sequence of sets $\langle A_i : i \in I \rangle$ such that $\mathrm{tp}(A_i/A_{<i}C)$ is finitely satisfied in M for every $i \in I$. When $C = M$ we simply say $\langle A_i : i \in I \rangle$ is an M-f.s. sequence.

Definition 2.6. We call $C \supseteq M$ *full* if, for every n, every $p \in S_n(M)$ is realized in C.

The relevance of fullness is the following stationarity result.

Lemma 2.7 ([7, Lemma 2.12]). *Suppose $C \supseteq M$ is full and $p \in S(C)$ is finitely satisfied in M. Then for any set $D \supseteq C$, there is a unique $q \in S(D)$ extending p that remains finitely satisfied over M.*

Definition 2.8. Fix any full $C \supset M$ and $q(\bar{x}) \in S(C)$ finitely satisfied in M. For any linear order $(I, \leq)$ (even finite) the *I-Morley sequence* $q^{(I)}$ is the unique type in the variables $(\bar{x}_i : i \in I)$ satisfying $q = \mathrm{tp}(\bar{x}_i/D)$ and $\mathrm{tp}(\bar{x}_i/D\bar{x}_{<i})$ is finitely satisfiable in M for every $i \in I$.

Lemma 2.9. *Suppose $C \supseteq M$ is full and $r(\bar{x}, \bar{y}) \in S(C)$ is finitely satisfied in M, and let $q(\bar{x})$ be the restriction of $r(\bar{x}, \bar{y})$ to $\bar{x}$. For any linear order $(I, \leq)$ and any $(\bar{a}_i : i \in I)$ realizing $q^{(I)}$, there are $(\bar{b}_i : i \in I)$ such that $(\bar{a}_i\bar{b}_i : i \in I)$ realizes $r^{(I)}$.*

Proof. Choose any $(\bar{a}_i'\bar{b}_i' : i \in I)$ realizing $r^{(I)}$. Then $(\bar{a}_i' : i \in I)$ also realizes $q^{(I)}$, so there is an automorphism $\sigma \in Aut(\mathfrak{C})$ fixing D pointwise with $\sigma(\bar{a}_i') = \bar{a}_i$. Then $(\bar{a}_i\sigma(\bar{b}_i') : i \in I)$ also satisfies $r^{(I)}$. $\qquad\square$

The following lemma is essentially [4, Theorem 4.14], although the f.s. dichotomy allows us to give a simple proof.

Lemma 2.10. *If T has the f.s. dichotomy then for every infinite indiscernible $(\bar{a}_i : i \in I)$ where $(I, \leq)$ is Dedekind complete and for every singleton $b \in \mathfrak{C}$, there are $(b_i : i \in I)$ and some $i^* \in I$ such that $b_{i^*} = b$ and $(\bar{a}_i b_i : i \in I)$ is indiscernible.*

Proof. Choose M and a full $D \supseteq M$ such that $(\bar{a}_i : i \in I)$ is an M-f.s. sequence over D, by [7, Lemma 2.20].

Claim. There is some $i^* \in I$ such that $(\bar{a}_i : i < i^*)^\frown(\bar{a}_{i^*}b)^\frown(\bar{a}_i : i > i^*)$ is an M-f.s. sequence over D.

Proof of Claim. This is immediate from the proof of Lemma [7, Lemma 3.2], where the following is shown. Let I_0 be the maximal initial segment of I such that $\mathrm{tp}(b/D\bar{a}_{I_0})$ is finitely satisfiable in M, and let $I = I_0^\frown I_1$. If I_1 has a minimal element i_1^*, it is shown that we can take $i^* = i_1^*$. Otherwise, it is shown that $(\bar{a}_i : i \in I_0)^\frown(b)^\frown(\bar{a}_i : i \in I_1)$ is an M-f.s. sequence over D. In this case, I_0 must have a maximal element i_0^* by Dedekind completeness, and $(\bar{a}_i : i \in I_0\backslash i_0^*)^\frown(\bar{a}_{i_0^*}b)^\frown(\bar{a}_i : i \in I_1)$ is also an M-f.s. sequence over D. $\qquad\Diamond$

Let $r(\bar{x}, y) := \mathrm{tp}(\bar{a}_{i^*}b/D)$ and let $q(\bar{x}) := \mathrm{tp}(\bar{a}_i/D)$ for some (equivalently, for every) $i \in I$. Then, by Lemma 2.9 there are $\{b_i : i \in I\}$ such that $(\bar{a}_ib_i : i \in I)$ realizes $r^{(I)}$. By [7, Lemma 2.22], $(\bar{a}_ib_i : i \in I)$ is indiscernible over D. Finally, choose an automorphism of $\mathfrak{C}$ fixing $A_I D$ with $\sigma(b_{i^*}) = b$. Then $(\bar{a}_i\sigma(b_i) : i \in D)$ is as desired. $\qquad\square$

Before the main theorem, we need state the following obstruction to monadic NIP.

Definition 2.11. A *tuple-coding configuration* is a formula $\phi(\bar{x}, \bar{y}, z)$, a pair of mutually indiscernible sequences $(\bar{a}_i : i \in \mathbb{Q})$ and $(\bar{b}_j : j \in \mathbb{Q})$, and a set of singletons $\{ c_{i,j} \mid i \in \mathbb{Q}, j \in \mathbb{Q} \}$, such that $\models \phi(\bar{a}_i, \bar{b}_j, c_{k,\ell})$ if and only if $(i,j) = (k,\ell)$.

Theorem 2.12. *Let T be a complete theory. Then the following are equivalent.*

(1) T is monadically NIP.
(2) T admits widening of indiscernibles.
(3) T is dp^+-minimal.
(4) T is dp-minimal and has endless indiscernible triviality.

Proof. $(1) \Rightarrow (2)$ Let $\mathcal{I} = (\bar{a}_i : i \in I)$ be an infinite indiscernible sequence. Let $(I_2, \leq)$ be the Dedekind completion of $(I, \leq)$. By Compactness, choose $(\bar{a}_i : i \in I_2)$ extending $(\bar{a}_i : i \in I)$ that is also indiscernible. By Lemma 2.10 there is $i^* \in I_2$ and $(b_j : j \in I_2)$ such that $b_{i^*} = b$ and $(\bar{a}_i b_i : i \in I_2)$ is indiscernible. Finally, let $(J, \leq) \subseteq (I_2, \leq)$ be the subordering with universe $I \cup \{i^*\}$. Then $(\bar{a}_j b_j : j \in J)$ is as desired.

$(2) \Rightarrow (3)$ Suppose T admits widening of indiscernibles, $\mathcal{I} = (\bar{a}_i : i \in \mathbb{Q})$ is an infinite indiscernible sequence, and b is a singleton. Let $\mathcal{I}^+ = (\bar{a}_j b_j : j \in J)$ with $J \subset \mathbb{R} \cup \{ \pm\infty \}$ be the sequence produced by widening $\mathcal{I}$, with $b_{j^*} = b$. Then $\mathcal{I}^+_{<j^*}$ and $\mathcal{I}^+_{>j^*}$ are mutually indiscernible over $\bar{a}_{j^*} b$.

$(3) \Rightarrow (1)$ Suppose T is not monadically NIP, so by [7] T admits a tuple-coding configuration $\mathcal{I} = (\bar{a}_i : i \in \mathbb{Q})$, $\mathcal{J} = (\bar{b}_j : j \in \mathbb{Q})$, $\{ c_{i,j} \mid i, j \in \mathbb{Q} \}$, $\phi(\bar{x}, \bar{y}, z)$. Since $\mathcal{I}$ and $\mathcal{J}$ are mutually indiscernible, $\mathcal{I}^* = (\bar{a}_i \bar{b}_i : i \in \mathbb{Q})$ is also indiscernible. Then any $c_{i,j}$ with $i \neq j$ witnesses the failure of dp^+-minimality.

$(1) \Leftrightarrow (4)$ This is contained in the corrigendum to [7]. $\square$

We note one application in a special case. Call a theory T *binary* if, the type of any n-tuple of elements from any $M \models T$ is determined by the union of the types of its 2-element subsequences. Clearly, if T is binary, then it has indiscernible triviality.

Corollary 2.13. *If T is binary, then T is monadically NIP if and only if every completion of T is dp-minimal. In this case, T has indiscernible triviality (as opposed to just endless indiscernible triviality).*

2.2. **Uniqueness of the cut point.** In a dp-minimal theory, a singleton can "exceptionally relate" to several points of an indiscernible sequence, although removing any one of these points destroys the relation. For example, consider an infinite linearly independent set in a vector space and the sum of some finite subset. However, in a dp^+-minimal theory, we will show that singleton can "exceptionally relate" to only one cut in an endless indiscernible sequence. This implies if an indiscernible $\mathcal{I}$ is not indiscernible over b, then the choice

of i^* in Definition 2.1 is essentially unique, and is truly unique if the index sequence is also dense and Dedekind complete.

One outcome of this is that an endless, dense, Dedekind complete indiscernible sequence in a monadically NIP theory gives rise to a well-defined notion of independence (still depending on the sequence). Namely, given a model $M \subset \mathfrak{C}$ and such an indiscernible sequence $\mathcal{I}$, then we may "widen" $\mathcal{I}$ to $\mathcal{I}^+$ covering M. There will be some points that $\mathcal{I}$ remains indiscernible over, which would be inserted into every tuple of $\mathcal{I}^+$, which we call orthogonal points. Then two non-orthogonal points may be considered dependent over $\mathcal{I}$ if they are placed into the same tuple of $\mathcal{I}^+$ and independent otherwise, and orthogonal points are considered independent from everything. From the results of this section, it follows that at every step there is a unique tuple where each non-orthogonal point can be inserted, and furthermore this tuple does not depend on the order in which points are inserted.

The main result concerning indiscernible sequences is Proposition 2.21. The proof of this makes use of the parallel result Proposition 2.18 concerning M-f.s. sequences (Definition 2.5). We begin with some general lemmas on indiscernibles without any assumption of monadic NIP.

Lemma 2.14. *Suppose $(I, \leq)$ is endless and $I = I_0 \widehat{\ } I_1 \widehat{\ } I_2$ is any partition of I into three non-empty convex sets. If both $(\bar{a}_i : i \in I_0)$ and $(\bar{a}_i : i \in I_1 \cup I_2)$ are mutually indiscernible, and $(\bar{a}_i : i \in I_0 \cup I_1)$ and $(\bar{a}_i : i \in I_2)$ are mutually indiscernible, then $(\bar{a}_i : i \in I)$ is indiscernible.*

Proof. Given an increasing tuple $\bar{a}_{i_1} \ldots \bar{a}_{i_n}$, we may shift an individual $\bar{a}_{i_j}$ from $\mathcal{I}_1$ to $\mathcal{I}_0$ while preserving the type by using the second mutual indiscernibility assumption, and similarly may shift an individual $\bar{a}_{i_j}$ from $\mathcal{I}_2$ to $\mathcal{I}_1$ while preserving the type by using the first mutual indiscernibility assumption. Repeating this process, we eventually shift all individual subtuples into $\mathcal{I}_0$, which is assumed to be indiscernible. $\square$

Lemma 2.15. *Let $\mathcal{I} = (\bar{a}_i : i \in I)$ be a sequence, and suppose there is a cut $I = I_0 \widehat{\ } I_1$ such that I_0 and I_1 are mutually indiscernible. If there is $i^* \in I$ such that $\mathcal{I} \setminus \{ \bar{a}_{i^*} \}$ is indiscernible, then $\mathcal{I}$ is indiscernible.*

Proof. Given an increasing tuple $\bar{a}_{i_1} \ldots \bar{a}_{i_n}$, if some $\bar{a}_{i_j} = \bar{a}_{i^*}$, we may use the mutual indiscernibility to shift $\bar{a}_{i^*}$ to another tuple on the same side of the cut, possibly also shifting other $\bar{a}_{i_k}$ if this is needed to make space. $\square$

Lemma 2.16. *Let $\mathcal{I} = (\bar{a}_i : i \in I)$ be an endless indiscernible sequence, and let b be a singleton. If any of the following hold, then $\mathcal{I}$ is indiscernible over b.*

(1) *There are two distinct cuts $I = I_0 \widehat{\ } I_1$ and $I = J_0 \widehat{\ } J_1$ such that $(\bar{a}_i : i \in I_0)$ and $(\bar{a}_i : i \in I_1)$ are mutually indiscernible over b, as are $(\bar{a}_i : i \in J_0)$ and $(\bar{a}_i : i \in J_1)$.*

(2) *There is a cut $I = I_0 \widehat{\ } I_1$ such that $(\bar{a}_i : i \in I_0)$ and $(\bar{a}_i : i \in I_1)$ are mutually indiscernible over b, and an element $i^* \in I$ that is*

neither the maximal element of I_0 nor the minimal element of I_1 with $(\bar{a}_i : i < i^*)$ and $(\bar{a}_i : i > i^*)$ are mutually indiscernible over b.

(3) There are two distinct element $i^*, i' \in I$ with $(\bar{a}_i : i < i^*)$ and $(\bar{a}_i : i > i^*)$ mutually indiscernible over $a_{i^*}b$ and $(\bar{a}_i : i < i')$ and $(\bar{a}_i : i > i')$ mutually indiscernible over b, and there is no cut $I = I_0^\frown I_1$ such that $(\bar{a}_i : i \in I_0)$ and $(\bar{a}_i : i \in I_1)$ are mutually indiscernible over b and i^*, i' are the maximal element of I_0 and the minimal element of I_1.

Proof. (1) Suppose $J_0 \subset I_0$. Apply Lemma 2.14 to $J_0^\frown I_0 \backslash J_0^\frown I_1$.

(2) Suppose there is such a cut $I = I_0^\frown I_1$ and an $i^* \in I$ such that $(\bar{a}_i : i < i^*)$ and $(\bar{a}_i : i > i^*)$ are mutually indiscernible over b. Possibly reversing the order, we may suppose $i^* \in I_0$. Let $I \backslash \{i^*\} = J_0^\frown J_1$ with $J_0 = \{i \in I \mid i < i^*\}$ and $J_1 = \{i \in I_0 \mid i > i^*\}$. If i^* is not the maximal element of I_0, then this cut is distinct from $(I_0 \backslash \{i^*\})^\frown I_1$. So by Case (1), $\mathcal{I} \backslash \{\bar{a}_{i^*}\}$ is indiscernible over b. So by Lemma 2.15, $\mathcal{I}$ is indiscernible over b.

(3) Suppose there are such i^* and i' with $i' < i^*$ (the case $i' > i^*$ is symmetric). Applying Case (2) to $\mathcal{I} \backslash \{\bar{a}_{i^*}\}$, we have that i' is the immediate predecessor of i^*. Now consider the cut $I = I_0^\frown I_1$ with $I_0 = (i \leq i')$ and $I_1 = (i \geq i^*)$. Routine manipulation of indiscernibles shows $\mathcal{I}_{I_0}$ and $\mathcal{I}_{I_1}$ are mutually indiscernible over b. $\square$

Lemma 2.17. *Suppose $(I, \leq)$ is endless, I_0 is a proper initial segment, and I_1 is a proper final segment (but we are not assuming $I = I_0 \cup I_1$). If $(\bar{a}_i : i \in I)$ is indiscernible and M and a full $D \supset M$ are chosen so that $(\bar{a}_i : i \in I_0)^\frown (Y)^\frown (\bar{a}_i : i \in I_1)$ is an M-f.s. sequence over D and indiscernible over D, then $(\bar{a}_i : i \in I_0)$ and $(\bar{a}_i : i \in I_1)$ are mutually indiscernible over DY.*

Proof. This is immediate from [7, Proposition 2.17] stating that an M-f.s. sequence over a full $D \supset M$ is an order-congruence over D ([7, Definition 2.16]). $\square$

We now begin working in monadically NIP theories.

Proposition 2.18 (Cut Lemma, f.s. version)**.** *Let T be monadically NIP. Suppose $(I, \leq)$ is endless, $(\bar{a}_i : i \in I)$ is indiscernible, and $b \in \mathfrak{C}$ is any singleton for which $(\bar{a}_i : i \in I)$ is not indiscernible over b. Let $M \subseteq D$ be such that M is a model, D/M is full, and $(\bar{a}_i : i \in I)$ is an M-f.s. sequence over D. Let I_0 be the maximal initial segment of I such that $\mathrm{tp}(b/A_{I_0}D)$ is finitely satisfied in M, and let $I = I_0^\frown I_1$. Let i_0^* denote the maximal element of I_0 if it exists, and i_1^* denote the minimal element of I_1 if it exists.*

(1) *If i_1^* does not exist (and possibly if it does), then*

$$(\bar{a}_i : i \in I_0)^\frown (b)^\frown (\bar{a}_i : i \in I_1)$$

is an M-f.s. sequence over D

(2) If i_1^* exists, then

$$(\bar{a}_i : i < i_1^*)^\frown (\bar{a}_{i_1^*} b)^\frown (\bar{a}_i : i > i_1^*)$$

is an M-f.s. sequence over D.

(3) If i_0^* exists, then

$$(\bar{a}_i : i < i_0^*)^\frown (\bar{a}_{i_0^*} b)^\frown (\bar{a}_i : i > i_0^*)$$

is an M-f.s. sequence over D if and only if

$$(\bar{a}_i : i \in I_0)^\frown (b)^\frown (\bar{a}_i : i \in I_1)$$

is an M-f.s. sequence over D.

Furthermore, these are the only three possibilities for inserting b so that the result is an M-f.s. sequence over D.

Proof. The proof of Clauses (1) and (2) is explicitly contained in the proof of [7, Lemma 3.2]. The "if" direction of Clause (3) is immediate from Condensation ([7, Definition 2.5]).

For the "only if" direction of Clause (3), suppose b can be inserted at i_0^*. Then we must show that $(\bar{a}_i : i \in I_0)^\frown (b)^\frown (\bar{a}_i : i \in I_1)$ is an M-f.s. sequence over D. If i_1^* does not exist, this follows from Clause (2). If i_1^* does exist, then Clause (1) gives that $(\bar{a}_i : i < i_1^*)^\frown (\bar{a}_{i_1^*} b)^\frown (\bar{a}_i : i > i_1^*)$ is an M-f.s. sequence over D, and together with our assumption that $(\bar{a}_i : i < i_0^*)^\frown (\bar{a}_{i_0^*} b)^\frown (\bar{a}_i : i > i_0^*)$ is an M-f.s. sequence over D, we are finished.

We now prove the Furthermore clause. We recall Lemma 2.17, which we will use repeatedly without mention. Suppose b can be inserted at some index i^*, which possibly represents a new cut. By the definition of I_0, we cannot have i^* above any element of I_1. If i_1^* exists and $i^* \in I$ with $i^* \neq i_1^*$, then by Lemma 2.16(3) there is a cut $I = J_0^\frown J_1$ such that i^* is the maximal element of J_0 and i_1^* is the minimal element of J_1, so $I_0 = J_0$.

We cannot have that i_1^* does not exist and $i^* \notin I$ defines a cut other than $I_0^\frown I_1$, since this would contradict Lemma 2.16(1). The case where i_1^* does not exist and $i^* \in I$ with $i^* \neq i_0^*$ is ruled out by Lemma 2.16(2), which also rules out case where i_1^* exists and $i^* \notin I$ defines a cut other than $I_0^\frown I_1$. $\square$

The following Corollary further clarifies the situation.

Corollary 2.19. *Let I, M, D, $(\bar{a}_i : i \in I)$ and $b \in \mathfrak{C}$ be as in Lemma 2.18. Suppose $(\bar{a}_i : i \in I)$ is not indiscernible over b and let $I = I_0^\frown I_1$ be the cut where I_0 is the maximal initial segment of I satisfying $\mathrm{tp}(b/A_{I_0}D)$ is finitely satisfied in M. Then:*

(1) *The only possible $i^* \in I$ such that $(\bar{a}_i : i < i^*)^\frown (\bar{a}_{i^*} b)^\frown (\bar{a}_i : i > i^*)$ is an M-f.s. sequence over D are the maximal element of I_0 or the minimal element of I_1. In particular, there are at most two such $i^* \in I$.*

(2) *If $(I, \leq)$ is dense, then there is at most one such $i^* \in I$.*

(3) *If $(I, \leq)$ is Dedekind complete, then there is at least one such $i^* \in I$.*

Proof. (1) Immediate from the statement of Lemma 2.18.

(2) If $(I, \leq)$ is dense, then it is not possible for I_0 to have a maximum and I_1 has a minimum, so we are done by (1).

(3) If $(I, \leq)$ is Dedekind complete, then either i_0^* or i_1^* from Lemma 2.18 must exist. If i_1^* exists, then we are finished by Clause (2) of Lemma 2.18. If i_1^* does not exist, then we are finished by Clauses (1) and (3) of Lemma 2.18. $\qquad\square$

Remark 2.20. Easy examples involving $T = DLO$ show that the remaining possibilities can indeed happen.

- If $a_q = q$ for all $q \in \mathbb{Q}$, then for $b = \pi$ and every element of M is larger than every rational, then $(a_q : q < \pi)^\frown(\pi)^\frown(a_q : q > \pi)$ is an M-f.s. sequence, but there is no $q \in \mathbb{Q}$ such that $(a_i : i < q)^\frown(a_q b)^\frown(a_i : i > q)$ is an M-f.s. sequence.
- If $a_i = i$ for all $i \in \mathbb{Z}$, $b = \pi$, every element of M is larger than every rational, then $(a_i : i < \pi)^\frown(\pi)^\frown(a_i : i > \pi)$ is an M-f.s. sequence, but also both $(a_i : i < 3)^\frown(3^\frown b)^\frown(a_i : i \geq 4)$ and $(a_i : i \leq 3)^\frown(b^\frown 4)^\frown(a_i : i > 4)$ are M-f.s. sequences.

Proposition 2.21 (Cut Lemma, indiscernible version). *Suppose T is monadically NIP and $(I, \leq)$ is endless. Suppose $(\bar{a}_i : i \in I)$ is indiscernible and $b \in \mathfrak{C}$ is a singleton such that $(\bar{a}_i : i \in I)$ is not indiscernible over b. Then one of the following holds.*

(1) *There is a cut $I = I_0^\frown I_1$ such that $(\bar{a}_i : i \in I_0)$ and $(\bar{a}_i : i \in I_1)$ are mutually indiscernible over b.*

   Furthermore in this case, the cut is unique. And $i^* \in I$ is such that $(\bar{a}_i : i < i^*)$ and $(\bar{a}_i : i > i^*)$ are mutually indiscernible over b if and only if i^* is either the maximal element of I_0 or the minimal element of I_1.

(2) *There is a unique $i^* \in I$ such that $(\bar{a}_i : i < i^*)$ and $(\bar{a}_i : i > i^*)$ are mutually indiscernible over b.*

   Furthermore in this case, they are mutually indiscernible over $a_{i^*} b$.

Proof. By [7, Lemma 2.20], choose M and a full $D \supset M$ so that $(\bar{a}_i : i \in I)$ is an M-f.s. sequence over D and indiscernible over D. We then apply the f.s. Cut Lemma. If there is a cut $I = I_0^\frown I_1$ such that $(\bar{a}_i : i \in I_0)^\frown(b)^\frown(\bar{a}_i : i \in I_1)$ form an M-f.s. sequence over D, then $(\bar{a}_i : i \in I_0)$ and $(\bar{a}_i : i \in I_1)$ are mutually indiscernible over b by Lemma 2.17. The uniqueness of the cut follows from Lemma 2.16(1). If i^* is either the maximal element of I_0 or the minimal element of I_1, it is immediate that $(\bar{a}_i : i < i^*)$ and $(\bar{a}_i : i > i^*)$ are mutually indiscernible over $\bar{a}_{i^*} b$. The "only if" part about i^* follows from Lemma 2.16(2).

If there is no cut $I = I_0^\frown I_1$ such that $(\bar{a}_i : i \in I_0)^\frown(b)^\frown(\bar{a}_i : i \in I_1)$ form an M-f.s. sequence over D, then there is $i^* \in I$ such that $(\bar{a}_i : i < i^*)^\frown(\bar{a}_{i^*} b)^\frown(\bar{a}_i : i > i^*)$ is an M-f.s. sequence over D, and so by Lemma 2.17, $(\bar{a}_i : i < i^*)$

and $(\bar{a}_i : i > i^*)$ are mutually indiscernible over $a_{i^*}b$. The uniqueness of i^* follows from Lemma 2.16(3). □

3. DISTALITY IN HEREDITARY CLASSES AND MONADIC DISTALITY

This section contains some collapse results regarding (monadic) distality. It is not immediate that distality of a hereditary class implies monadic NIP, but it implies NIP and by [8, Theorem 4.6], this implies monadic NIP. The first collapse result is monadic NIP and distality is equivalent to monadic distality, mirroring the collapse of monadic NIP and stability to monadic stability. The second is the collapse of distality to monadic distality in hereditary classes, mirroring similar results for monadic stability and monadic NIP from [8]. After these collapse results, we begin studying distality and distal expansions in hereditary classes, showing that various natural definitions of distality coincide and proving that classes of bounded twin-width admit distal expansions in a way that also lifts to infinite models of their universal theory.

3.1. **Monadic distality.**

Definition 3.1. A complete theory T is *distal* if for every indiscernible sequence $I = I_1 \frown I_2 \frown I_3$ such that I_1 has no maximum, I_2 no maximum or minimum, and I_3 no minimum, and for every $\bar{a}, \bar{b}$, if $I_1 \frown \bar{a} \frown I_2 \frown I_3$ and $I_1 \frown I_2 \frown \bar{b} \frown I_3$ are each indiscernible, then so is $I_1 \frown \bar{a} \frown I_2 \frown \bar{b} \frown I_3$.

An incomplete theory T is *distal* if $Th(M)$ is distal for every $M \models T$.

A (possibly incomplete) theory T is *monadically distal* if $Th(M^+)$ is distal for every monadic expansion M^+ of every $M \models T$.

Note that unlike many model-theoretic properties such as (monadic) stability/NIP, (monadic) distality is not preserved under reducts; for example, the theory of equality is not distal. Furthermore, the example below shows monadic distality is not preserved under passing to substructure. Nevertheless, we will recover various results for monadic distality paralleling those for monadic stability and monadic NIP.

Example 1 (A theory T that is monadically distal, but $T_\forall$ is not distal). Let $\mathcal{L} = \{U, V, <, E\}$ be a relational language with U, V unary and $<, E$ binary. Let M be an infinite $\mathcal{L}$-structure where U, V partition the domain, each point is E-related to a unique point, any two E-related points are in different unary relations, and $<$ is a linear order on $U(M)$. It is easy to check that M is monadically distal. However, the substructure $V(M)$ is a pure set and so is not distal.

Proposition 3.2. *If T is a theory that is monadically NIP but not monadically distal, then some $M \models T$ contains an infinite totally indiscernible set of singletons.*

Proof. Passing to an appropriate completion of T and renaming, we may assume T is complete. Let T' be a monadic expansion of T that is not distal. Since T is monadically NIP, so is T', and so T' is dp-minimal. By [24, Corollary 9.19], if a theory is dp-minimal, then it is not distal if and only if some unrealized global invariant 1-type is generically stable. Let $p'(x)$ be such a type in T', and consider a Morley sequence $\mathcal{I}'$ in p' (as described after [24, Lemma 2.23]). Then by [24, Theorem 2.29], $\mathcal{I}'$ is totally indiscernible, and thus so is the T-reduct $\mathcal{I}$. $\square$

Corollary 3.3. *Let T be a theory. Then T is monadically distal if and only if T is distal and monadically NIP.*

Proof. The forward direction is clear, so suppose that T is monadically NIP and distal. If T is not monadically distal, then by the previous lemma some $M \models T$ contains an infinite totally indiscernible set, contradicting that T is distal. $\square$

3.2. **Distality in hereditary classes.**

Definition 3.4. A *hereditary class* is a class of structures closed under substructure.

Given a class $\mathcal{C}$ of structures, we define $Th(\mathcal{C})$ as $\bigcap_{M \in \mathcal{C}} Th(M)$.

In addition to Definition 3.1, distality can be characterized in terms of strong honest definitions, which we now describe. This can be stated entirely at the level of a class of structures $\mathcal{C}$, rather than having to look at models of $Th(\mathcal{C})$.

Definition 3.5. Let $\mathcal{C}$ be a class of structures. A formula $\phi(\bar{x}, \bar{y})$ *admits a strong honest definition in $\mathcal{C}$* if there is a formula $\psi(\bar{x}, \bar{z})$ such that for all $M \in \mathcal{C}$, all finite $A \subset M$ with $|A| \geq 2$, and all $\bar{b} \in M^{|\bar{x}|}$, there is some $\bar{d} \in A^{|\bar{z}|}$ such that $M \models \psi(\bar{b}, \bar{d})$ and all elements of $\psi(M, \bar{d})$ realize the same ϕ-type over A.

We say ϕ *admits a strong honest definition in M* if it does so in $\{ M \}$. We say a class $\mathcal{C}$ of structures *admits strong honest definitions* if every formula admits strong honest definitions in $\mathcal{C}$, and that a (possibly incomplete) theory T *admits strong honest definitions* if the class of models of T does.

We remark that the definable cell decompositions described in [10, Definition 2.7] are nearly the same as strong honest definitions, but allow for more refined quantitative results, which we will not be concerned with.

Fact 3.6 ([11, Theorem 21]). *A complete theory T is distal if and only if it admits strong honest definitions.*

The following lemma extends this fact to incomplete theories of the form $Th(\mathcal{C})$. The main point is $(1) \Rightarrow (2)$ that from knowing only that ϕ admits a strong honest definition separately in each completion of $Th(\mathcal{C})$, we may deduce the uniformity result that there is a single strong honest definition

for ϕ that works across all completions of $Th(\mathcal{C})$. This shows that there is a single robust definition of distality for a class of structures.

Lemma 3.7. *Let $\mathcal{C}$ be a class of $\mathcal{L}$-structures. The following are equivalent.*

(1) $Th(\mathcal{C})$ is distal.
(2) $Th(\mathcal{C})$ admits strong honest definitions.
(3) $\mathcal{C}$ admits strong honest definitions.

Proof. The key point is that for a fixed formula $\phi(\bar{x}; \bar{y})$ and a fixed formula $\psi(\bar{x}; \bar{z})$, we may express by a set of sentences that ψ is a strong honest definition for $\phi(\bar{x}; \bar{y})$. Namely, for each $n \geq 1$, we may write a sentence θ_n stating that for every choice of n elements $a_1, \ldots, a_n$, for one of the finitely-many choices of ψ-definable families of sets over these elements, this family covers the whole structure and the truth value of $\phi(\bar{x}; \bar{a})$ is constant on each set in this family for each of the finitely many choices of $\bar{a} \in \{ a_1, \ldots, a_n \}^{|\bar{y}|}$. Using this, the equivalences follow by standard compactness arguments.

$(1) \Rightarrow (2)$ Suppose there is some $\phi(\bar{x}; \bar{y})$ such that there is no strong honest definition for $Th(\mathcal{C})$. So for each formula $\psi(\bar{x}, \bar{z})$, there is some $M \models Th(\mathcal{C})$ such that ψ fails over some parameter set $a_1, \ldots, a_{n_\Psi}$. As above, we may write a sentence χ_ψ saying that there exist elements $a_1, \ldots, a_{n_\Psi}$ over which ψ fails. Note that by the "standard coding trick" (as in [14, Lemma 2.5]) if there is some finite set Ψ of formulas such that for each $M \in \mathcal{C}$, some $\psi \in \Psi$ serves as a strong honest definition for ϕ in M, then ϕ admits a strong honest definition in $\mathcal{C}$. So $Th(\mathcal{C}) \cup \{ \chi_\psi \mid \text{formulas } \psi(\bar{x}, \bar{z}) \}$ is finitely satisfiable, and thus satisfiable. Thus there is some $M \models Th(\mathcal{C})$ such that there is no strong honest definition for $\phi(\bar{x}; \bar{y})$ in M.

$(2) \Rightarrow (1)$ is immediate from Fact 3.6.

$(3) \Rightarrow (2)$ Suppose $\mathcal{C}$ admits strong honest definitions. Fix a formula $\phi(\bar{x}; \bar{y})$ and let $\psi(\bar{x}; \bar{z})$ be strong honest definition. By assumption, every θ_n as described in the first paragraph is in $Th(\mathcal{C})$.

$(2) \Rightarrow (3)$ We may express the assumption of (2) via sentences θ_n as in the first paragraph. As $M \models Th(\mathcal{C})$ for each $M \in \mathcal{C}$, the same strong honest definitions will work for M. $\square$

The equivalences in the following two results are analogous to results for monadic NIP and monadic stability in [8]. For the following proposition, recall that Example 1 shows passing monadic distality from a theory to its universal part is not in general automatic.

Proposition 3.8. *Let $\mathcal{C}$ be a hereditary class of relational structures. Then the following are equivalent.*

(1) $Th(\mathcal{C})_\forall$ is monadically distal.
(2) $Th(\mathcal{C})_\forall$ is distal.
(3) $Th(\mathcal{C})$ is monadically distal.
(4) $Th(\mathcal{C})$ is distal.

Proof. For the equivalences (1) $\iff$ (2) and (3) $\iff$ (4), note that if the theory T in question is distal then it is NIP, and thus monadically NIP by [8, Theorem 4.6]. Thus T is monadically distal by Corollary 3.3.

Clearly (1) $\Rightarrow$ (3), so we show (3) $\Rightarrow$ (1). Suppose $Th(\mathcal{C})_\forall$ is not monadically distal. If it is not monadically NIP, then neither is $Th(\mathcal{C})$ by [8, Proposition 2.5], and we are finished, so assume it is monadically NIP. By Proposition 3.2, there is some $M \models Th(\mathcal{C})_\forall$ containing an infinite totally indiscernible set $\mathcal{I} = \{\, a_i \mid i \in \omega \,\}$. By induction on quantifier complexity, any finite subset of $\mathcal{I}$ considered as structure is still totally indiscernible. Thus $\mathcal{C}$ contains arbitrarily large totally indiscernible sets, and so by compactness $Th(\mathcal{C})$ has a model containing an infinite totally indiscernible set, contradicting distality. $\square$

Lemma 3.9. *Let $\mathcal{C}$ be a hereditary class of relational structures. Then the following are equivalent.*

(1) $Th(\mathcal{C})$ is monadically distal
(2) For every monadic expansion $\mathcal{C}^+$ of $\mathcal{C}$, $Th(\mathcal{C}^+)$ is distal

Proof. (1) $\Rightarrow$ (2) Let $\mathcal{C}^+$ be a monadic expansion of $\mathcal{C}$, i.e., a class where every element is a monadic expansion of some element of $\mathcal{C}$. Let $M^+ \models Th(\mathcal{C}^+)$ and let M be the reduct of M^+ to the language of $\mathcal{C}$. Then $M \models Th(\mathcal{C})$ is monadically distal, so $Th(M^+)$ is distal.

(2) $\Rightarrow$ (1) Let $M \models Th(\mathcal{C})$ and let M^+ be an arbitrary monadic expansion of M to an L^+-structure. Let $\mathcal{C}^+$ consist of all L^+-structures N^+ whose L-reduct is in $\mathcal{C}$. As $\mathcal{C}$ is hereditary, so is $\mathcal{C}^+$. Also, since $M \models Th(\mathcal{C})_\forall$, it follows that every finite L^+-substructure $N^+ \subseteq M^+$ is an element of $\mathcal{C}^+$. Thus, $M^+ \models Th(\mathcal{C}^+)_\forall$. But Proposition 3.8, combined with (2), implies that $Th(\mathcal{C}^+)_\forall$ is distal, so $Th(M^+)$ is distal as required. $\square$

Distality in hereditary classes very restrictive; for example, Ramsey's theorem implies that no hereditary graph class is distal. Since many of the good combinatorial properties of distal classes are preserved by passing to reducts, this motivates considering classes that have a distal expansion. As with distality, we are faced with choices about how much uniformity to require: we could ask only that every model of $Th(\mathcal{C})$ admit some distal expansion, or that there is a single expansion of $Th(\mathcal{C})$ with single uniform strong honest definition for every formula, or that there is an expansion $\mathcal{C}^+$ of $\mathcal{C}$ such that $Th(\mathcal{C}^+)$ is distal. There seems to be little reason to think these equivalent, but we will use the following implication.

Lemma 3.10. *Let $\mathcal{C}$ be a class of relational $\mathcal{L}$-structures. If $\mathcal{C}$ admits an expansion to a class of $\mathcal{L}^+$-structures $\mathcal{C}^+$ so that $Th(\mathcal{C}^+)_\forall$ is distal, then every model of $Th(\mathcal{C})_\forall$ admits a distal expansion to a model of $Th(\mathcal{C}^+)_\forall$.*

Proof. Let $M \models Th(\mathcal{C})$. Consider the theory given by the atomic diagram of M together with $Th(\mathcal{C}^+)_\forall$. This is finitely satisfiable as witnessed by structures in $\mathcal{C}^+$, and so has a model N^+, which embeds an $\mathcal{L}^+$-expansion M^+ of M. Then $M^+ \models Th(\mathcal{C}^+)_\forall$ is distal. $\square$

Lemma 3.11. *Let $\mathcal{C}$ be a class of structures such that $Th(\mathcal{C})$ is monadically NIP. If the structures in $\mathcal{C}$ admit a $\emptyset$-definable linear order, then T is distal.*

Proof. Since $\mathcal{C}$ is monadically NIP, it is dp-minimal. Since the structures in $\mathcal{C}$ are $\emptyset$-definably linearly ordered, they are distal by [24, Example 9.20]. □

We thus have a dichotomy for monadically NIP hereditary classes: either the structures in $\mathcal{C}$ admit an order-expansion (i.e., a class where every element is an expansion of some element of $\mathcal{C}$ by a linear order) that remains monadically NIP, which is also a distal expansion, or no distal expansion of $\mathcal{C}$ can have a $\emptyset$-definable linear order. In the case where $\mathcal{C}$ consists of binary relational structures, the classes that admit an order-expansion remaining monadically NIP coincides with the classes of *bounded twin-width*. The twin-width of a binary structure has a combinatorial definition (see [6]) that we will not use except in Example 2, and a class $\mathcal{C}$ has bounded twin-width if there is a uniform finite bound on the twin-width of all members.

Lemma 3.12. *Let $\mathcal{C}$ be a class of finite binary structures. Then $\mathcal{C}$ has bounded-twin width if and only if $\mathcal{C}$ admits an order-expansion $\mathcal{C}^<$ that is monadically NIP.*

Proof. Since bounded twin-width and monadic NIP are both preserved by closing under substructure, we may assume that $\mathcal{C}$ and $\mathcal{C}^<$ are hereditary. By [5, Theorem 3], a hereditary class of ordered binary structures is monadically NIP if and only if it has bounded twin-width.

($\Rightarrow$) Suppose $\mathcal{C}$ has bounded twin-width. Then it admits an order-expansion $\mathcal{C}^<$ with bounded twin-width (see [25, Fact 2]). Then $\mathcal{C}^<$ is monadically NIP, and thus so is $\mathcal{C}$.

($\Leftarrow$) Suppose $\mathcal{C}$ admits a monadically NIP order-expansion $\mathcal{C}^<$. Then $\mathcal{C}^<$ has bounded twin-width, which is preserved by reducts, so $\mathcal{C}$ has bounded twin-width. □

Classes with bounded twin-width include, for example, the class of planar graphs, and more generally every graph class determined by forbidden minors (other than the class of all graphs) [6]. The second part of the following result generalizes [22], which showed a strong honest definition for the edge relation, although [22] also gives an explicit bounds (as a function of the twin-width) on the complexity of the cell decompositions arising from the honest definition, which we do not.

Putting together the various results of this section, we obtain the following.

Theorem 3.13. *Let $\mathcal{C}$ be a hereditary class of binary relational structures with bounded twin-width. Then $\mathcal{C}$ admits a distal order-expansion, and thus so does every model of $Th(\mathcal{C})_\forall$. In the distal order-expansion of $\mathcal{C}$, every formula admits a strong honest definition.*

Proof. Let $\mathcal{C}$ be as assumed, and let $\mathcal{C}^<$ be an order-expansion that is monadically NIP. By [8, Proposition 2.5], every model of $Th(\mathcal{C}^<)_\forall$ is monadically NIP and so is distal by Lemma 3.11. By Lemma 3.10, every model of $Th(\mathcal{C})_\forall$

admits a distal order-expansion. Since $Th(\mathcal{C}^<)$ is distal, every formula admits a strong honest definition by Lemma 3.7. $\square$

Corollary 3.14. *Let G be an infinite planar graph. Then G admits a distal order-expansion.*

The distal expansion of such an infinite G is produced only by compactness, so in general even if we explicitly know the distal order-expansion of $Age(G)$, we do not explicitly obtain the distal order-expansion of G. However, in some cases we can obtain an explicit expansion, such as in the following example.

Example 2 (A distal expansion of the everywhere infinite forest)**.** We produce a distal expansion of the everywhere infinite forest, i.e. of the theory of an infinite acyclic graph where every point has infinite degree. For a finite forest, it is easy to describe a contraction sequence witnessing bounded twin-width (see [3, Lemma 5]), which can be used as described in the proof of [25, Fact 2] to construct the following distal order-expansion for the class of finite forests. Each finite tree is ordered by choosing an arbitrary root, and then linearly ordering the vertices so that they satisfy the following condition ($*$): for every vertex x, for each child y of x, the set of consisting of y and its descendants form a convex set in the order. A finite forest is then ordered by separately ordering each tree to satisfy ($*$), and then making each tree a convex set. We can choose an arbitrary root for each tree in the everywhere infinite forest and then construct from the root upwards a dense linear order that satisfies ($*$), and completing the order to make each tree convex. This will be a distal expansion, since the induced order on every finite substructure will be as described above.

Note that the proof that this is a distal expansion is reduced to combinatorics on finite forests, and does not rely on fine analysis of definability or indiscernible sequences in the infinite model.

Since distality is supposed to capture order-like behavior, it may be that the second possibility in the dichotomy discussed above never occurs for a class $\mathcal{C}$ admitting a distal expansion (i.e. it cannot be that no distal expansion of $\mathcal{C}$ has a definable linear order). In the case of binary structures, this would yield a converse to the previous theorem, showing every hereditary class admitting a distal expansion must have bounded twin-width.

Question 1. *Let $\mathcal{C}$ be a hereditary class of relational structures. If $\mathcal{C}$ admits an expansion $\mathcal{C}^+$ such that $Th(\mathcal{C}^+)$ is distal, can we take the additional relations in $\mathcal{C}^+$ to be solely a linear order (equivalently, to include a linear order)?*

This raises a related question in the setting of complete theories.

Question 2. *Let T be a complete theory and T^+ a distal expansion of T. Does T^+ admits an order-expansion that remains distal (or at least NIP)?*

4. Not interpreting an infinite group

In this section, we show a strengthening of the statement that if T is monadically NIP, then T does not interpret an infinite group. We then use this to show the existence of a (monadically) stable theory that has no distal expansion but does not interpret an infinite group. We begin with the following weakening of the notion of interpretation.

Definition 4.1 ([26]). Let M, N be structures (in possibly different languages). Let $\tau \colon N \to M^m$ be an injection. We say that M *trace defines* N *(via τ)* if for every N-definable subset X of N^n there is an M-definable subset Y of M^{mn} such that:

$$\text{for all } a_1, \ldots, a_n \in N, (a_1, \ldots, a_n) \in X \iff (\tau(a_1), \ldots, \tau(a_n)) \in Y$$

Let T, T^* be theories. Then T *trace defines* T^* if every T^*-model is trace definable in a T model, and T *trace defines* M if some T-model trace defines M.

In particular, if M interprets N then M trace defines N. However, if M trace defines N via $\tau \colon N \to M^m$, then the image $\tau[N]$ need not be a definable subset of M^m.

In [26], it was conjectured that monadically NIP theories cannot trace define an infinite group. We confirm this with the following more general result.

Definition 4.2. Let $(G, \cdot)$ be a set equipped with a binary function. Then $(G, \cdot)$ is a *cancellative magma* if it satisfies the following conditions: for every $a, b \in G$, there is at most one $x \in G$ such that $a \cdot x = b$ and at most one $y \in G$ such that $y \cdot a = b$.

The notion of dp-rank is characterized in [24, Theorem 4.18], and dp-minimality corresponds to the case of dp-rank 1. The only result about finite dp-rank that we need will be recalled within the following proof.

The following argument is similar to that given in [15, A.6.9] that a colored linear order cannot interpret an infinite group (first given in [21], in the more general setting of "local theories").

Proposition 4.3. *Let T be a theory with finite dp-rank and endless indiscernible triviality. Then T does not trace define an infinite cancellative magma.*

Proof. Let $M \models T$ and suppose M trace defines an infinite cancellative magma $(G, \cdot)$ via $\tau \colon G \to M^n$. Let M^+ be the expansion of M by an n-ary relation R naming $\tau(G)$. By passing to an elementary extension (and renaming M and G), we may suppose there is an indiscernible sequence $\mathcal{I} = (\bar{a}_i \in R(M^+) : i \in \mathbb{Q})$. We now return to working in the original language.

Since T has finite dp-rank, there is some fixed $K \in \omega$, depending only on n and the dp-rank of the theory, such that for every $\bar{p} \in M^n$, there is

a finite subset $F \subseteq \mathbb{R}$ with $|F| < K$ such that in the induced partition of $\mathbb{Q} \setminus F = \bigsqcup \{ I_i \mid i \leq |F| \}$ into open convex sets, each of the subsequences $\mathcal{I}_i := (\bar{a}_q : q \in I_i)$ is indiscernible over $\bar{p}$.

Let $\bar{b} \in \tau(G)$ correspond to the product $(((\bar{a}_1 \cdot \bar{a}_2) \cdot \bar{a}_3) \cdot \ldots \cdot \bar{a}_K)$. From above, find a finite subset $F \subseteq \mathbb{R}$ with $|F| < K$ for this $\bar{b}$ and choose an integer i^*, $1 \leq i^* \leq K$ with $i^* \notin F$. Choose an open interval $J \subseteq \mathbb{Q}$ for which $J \cap \{1, \ldots, K\} = \{i^*\}$ and for which $\mathcal{J} := (\bar{a}_j : j \in J)$ is indiscernible over $\bar{b}$. Since $\mathcal{I}$ was chosen to be indiscernible over $\emptyset$, $\mathcal{J} = (\bar{a}_j : j \in J)$ is indiscernible over $\{\bar{a}_q \mid q \notin J\}$. By endless indiscernible triviality, $\mathcal{J}$ would be indiscernible over $\bar{b} \cup \{\bar{a}_i \mid 1 \leq i \leq K, i \neq i^*\}$. But this is impossible. Choose any $j \in J \setminus \{i^*\}$. It follows from the indiscernibility that $\bar{b}'$, which is the same product but with $\bar{a}_{i^*}$ replaced by $\bar{a}_j$, must be equal to $\bar{b}$. By two applications of cancellation, we would obtain $\bar{a}_j = \bar{a}_{i^*}$, which is a contradiction. $\qquad\square$

Corollary 4.4. *If T is a monadically NIP theory, then T does not trace define an infinite cancellative magma.*

We remark that monadically NIP theories can have non-trivial binary functions, so long as they are not cancellative. For example, the theory of dense meet-trees in [24, Section 2.3.1] is monadically NIP by [18].

The following characterizes which o-minimal theories are monadically NIP. It is noteworthy that for this class of theories, indiscernible triviality and endless indiscernible triviality coincide.

Definition 4.5. An o-minimal structure $(M, <, \ldots)$ whose ordering is dense is *non-trivial* if there is a non-empty open interval $I \subseteq M$ and a definable, continuous function $f \colon I \times I \to M$ that is strictly monotone in both variables.

Proposition 4.6. *The following are equivalent for a complete, o-minimal theory T whose underlying order is dense.*

(1) *T is trivial.*
(2) *T has indiscernible triviality.*
(3) *T has endless indiscernible triviality.*
(4) *T is monadically NIP.*

Proof. (1) $\Rightarrow$ (2) Suppose that T is o-minimal and trivial. It follows from Lemmas 2.1 and 2.2 of [17] that T is *binary*, i.e. for every $M \models T$, the type of any n-tuple of elements is determined by the union of the types of its 2-element subsequences. Clearly, any binary theory T must satisfy indiscernible triviality.

(2) $\Rightarrow$ (3) is immediate. As T is o-minimal, it is dp-minimal, and so (3) $\Rightarrow$ (4) follows from Theorem 2.12.

(4) $\Rightarrow$ (1) Assume that (1) fails, i.e., there is a model $M \models T$ and a definable, continuous function $f \colon I^2 \to M$ that is strictly monotone in both variables. There are two arguments to show that T is not monadically NIP. On the one hand, note that one can construct, element by element, an

infinite $G \subseteq I$ for which $f\!\restriction_{G^2}$ is 1-1. [If $G_n \subseteq I$ is finite and $f\!\restriction_{G_n^2}$ is 1-1, then $f\!\restriction_{(G_n \cup \{g\})^2}$ is 1-1 for all but finitely many $g \in I$.] This directly describes a tuple-coding configuration as in Definition 2.11, contradicting monadic NIP. Alternately, quoting Theorem 1.1 of [19], we see that M trace defines an infinite group, hence T is not monadically NIP by Corollary 4.4. $\qquad\square$

We now turn to the existence of a monadically stable theory with no distal expansion that does not interpret an infinite group. This example seems to be folklore to some extent ([16] credits Pierre Simon with asking for the relevant graph class), and for its analysis it is already enough to know that monadically stable theories have trivial forking [2]. But since it has not appeared explicitly in the literature, we record it here as a corollary of Corollary 4.4.

Definition 4.7. Let $\mathcal{C}$ be a class of graphs. Then $\mathcal{C}$ is *somewhere dense* if for some d, there are d-subdivisions of arbitrarily large complete graphs appearing as (not necessarily induced) subgraphs of graphs in $\mathcal{C}$. Otherwise, $\mathcal{C}$ is *nowhere dense*. We will say a single infinite graph G is nowhere dense if $\{G\}$ is.

In [20], it is shown that if G is nowhere dense (there called superflat), then it is monadically stable. We now introduce a combinatorial condition that is necessary for M to have a distal expansion [12, Corollary 4.10].

Definition 4.8. Given a class of structures $\mathcal{C}$ and a formula $\phi(\bar{x}, \bar{y})$, we say ϕ has the *strong Erdős-Hajnal property in* $\mathcal{C}$ there is a fixed $\delta > 0$ such that for every $M \in \mathcal{C}$ and finite $A \subset M^{|\bar{x}|}, B \subset M^{|\bar{y}|}$, there are $A' \subset A, B' \subset B$ such that $|A'| \geq \delta|A|, |B'| \geq \delta|B|$ and either $A' \times B' \subset \phi(M)$ or $(A' \times B') \cap \phi(M) = \emptyset$.

Lemma 4.9 ([16, Theorem 50], [9, Proposition 5]). *There is a nowhere dense class of finite graphs $\mathscr{R}$ such that the edge relation does not have the strong Erdős-Hajnal property.*

We remark that the class constructed in [16, Theorem 50] comes from taking expander graphs of increasing expansion, while the class in [9, Proposition 5] comes from a probabilistic construction.

Corollary 4.10. *There is a monadically stable theory without a distal expansion that does not trace define an infinite cancellative magma (so in particular, does not interpret an infinite group).*

Proof. Let $\mathscr{R}$ be a graph class as in Lemma 4.9, let $\sqcup\mathscr{R}$ be the disjoint union of the graphs in $\mathscr{R}$, and let $T := Th(\sqcup\mathscr{R})$. Then $\sqcup\mathscr{R}$ is nowhere dense, and so T is monadically stable, and so does not trace define an infinite cancellative magma by Corollary 4.4. The edge relation in $\sqcup\mathscr{R}$ also does not have the strong Erdős-Hajnal property, and so T does not admit a distal expansion. $\qquad\square$

References

[1] Matthias Aschenbrenner, Artem Chernikov, Allen Gehret, and Martin Ziegler, *Distality in valued fields and related structures*, Transactions of the American Mathematical Society **375** (2022), no. 7, 4641–4710.

[2] John T Baldwin and Saharon Shelah, *Second-order quantifiers and the complexity of theories.*, Notre Dame Journal of Formal Logic **26** (1985), no. 3, 229–303.

[3] Pierre Bergé, Édouard Bonnet, and Hugues Déprés, *Deciding twin-width at most 4 is np-complete*, arXiv preprint arXiv:2112.08953 (2021).

[4] Achim Blumensath, *Simple monadic theories and indiscernibles*, Mathematical Logic Quarterly **57** (2011), no. 1, 65–86.

[5] Édouard Bonnet, Ugo Giocanti, Patrice Ossona de Mendez, Pierre Simon, Stéphan Thomassé, and Szymon Toruńczyk, *Twin-width IV: ordered graphs and matrices*, Journal of the ACM **71** (2024), no. 3, 1–45.

[6] Édouard Bonnet, Eun Jung Kim, Stéphan Thomassé, and Rémi Watrigant, *Twin-width I: tractable FO model checking*, Journal of the ACM (JACM) **69** (2021), no. 1, 1–46.

[7] Samuel Braunfeld and Michael C Laskowski, *Characterizations of monadic NIP*, Transactions of the American Mathematical Society, Series B **8** (2021), no. 30, 948–970.

[8] ———, *Existential characterizations of monadic NIP*, arXiv preprint arXiv:2209.05120 (2022).

[9] Marcin Brianski, Piotr Micek, Michał Pilipczuk, and Michał T Seweryn, *Erdös–Hajnal properties for powers of sparse graphs*, SIAM Journal on Discrete Mathematics **35** (2021), no. 1, 447–464.

[10] Artem Chernikov, David Galvin, and Sergei Starchenko, *Cutting lemma and Zarankiewicz?s problem in distal structures*, Selecta Mathematica **26** (2020), 1–27.

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

[12] Artem Chernikov and Sergei Starchenko, *Regularity lemma for distal structures*, Journal of the European Mathematical Society **20** (2018), no. 10, 2437–2466.

[13] Jan Dreier, Nikolas Mählmann, and Szymon Toruńczyk, *Flip-breakability: A combinatorial dichotomy for monadically dependent graph classes*, Proceedings of the 56th Annual ACM Symposium on Theory of Computing, 2024, pp. 1550?1560.

[14] Vincent Guingona, *On uniform definability of types over finite sets*, The Journal of Symbolic Logic **77** (2012), no. 2, 499–514.

[15] Wilfrid Hodges, *Model theory*, Cambridge University Press, 1993.

[16] Yiting Jiang, Jaroslav Nešetřil, Patrice Ossona de Mendez, and Sebastian Siebertz, *Regular partitions of gentle graphs*, Acta Mathematica Hungarica **161** (2020), no. 2, 719–755.

[17] Alan Mekler, Matatyahu Rubin, and Charles Steinhorn, *Dedekind completeness and the algebraic complexity of o-minimal structures*, Canadian Journal of Mathematics **44** (1992), no. 4, 843–855.

[18] Michel Parigot, *Théories d'arbres*, The Journal of Symbolic Logic **47** (1982), no. 4, 841–853.

[19] Ya'acov Peterzil and Sergei Starchenko, *A trichotomy theorem for o-minimal structures*, Proc. London Math. Soc. **77** (1998), no. 3, 481–523.

[20] Klaus-Peter Podewski and Martin Ziegler, *Stable graphs*, Fund. Math **100** (1978), no. 2, 101–107.

[21] Bruno Poizat, *A propos de groupes stables*, Studies in Logic and the Foundations of Mathematics, 1987, pp. 245–265.

[22] Wojciech Przybyszewski, *Distal combinatorial tools for graphs of bounded twin-width*, 38th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), 2023, pp. 1–13.

[23] Saharon Shelah, *Strongly dependent theories*, Israel Journal of Mathematics **204** (2014), no. 1, 1–83.

[24] Pierre Simon, *A Guide to NIP Theories*, Cambridge University Press, 2015.

[25] Pierre Simon and Szymon Toruńczyk, *Ordered graphs of bounded twin-width*, arXiv preprint arXiv:2102.06881 (2021).

[26] Erik Walsberg, *Notes on trace equivalence*, arXiv preprint arXiv:2101.12194 (2021).