Interpretations in Trees with Countably Many Branches

Alexander Rabinovich
The Blavatnik School of Computer Science
Tel Aviv University
Tel Aviv, Israel
Email: rabinoa@post.tau.ac.il

Sasha Rubin
IST Austria and TU Vienna
Vienna, Austria

***Abstract*—We study the expressive power of logical interpretations on the class of *scattered* trees, namely those with countably many infinite branches. Scattered trees can be thought of as the tree analogue of scattered linear orders. Every scattered tree has an ordinal *rank* that reflects the structure of its infinite branches. We prove, roughly, that trees and orders of large rank cannot be interpreted in scattered trees of small rank. We consider a quite general notion of interpretation: each element of the interpreted structure is represented by a set of tuples of subsets of the interpreting tree. Our trees are countable, not necessarily finitely branching, and may have finitely many unary predicates as labellings. We also show how to replace injective set-interpretations in (not necessarily scattered) trees by 'finitary' set-interpretations.**

***Index Terms*—Composition method, finite-set interpretations, infinite scattered trees, monadic second order logic.**

I. Introduction

Monadic-second order logic (MSO) extends first-order logic with free variables that range over subsets of the domain, and allows quantification over them. When interpreted over trees MSO is expressive enough to capture interesting mathematics while still being manageable. Indeed, Rabin [20] proved that the MSO-theory of the full binary tree $\mathfrak{T}_2$ is decidable, and many other logical theories have been shown decidable by a reduction to this theory (see for instance the introductory sections in [20]). The *interpretation method* is a broad term that refers to effective reductions that are expressible logically, by a collection of formulas.[1]

We consider interpretations that define, by MSO-formulas, structures $\mathfrak{A}$ inside trees $\mathfrak{T}$. Our trees are subtrees of the countably-branching tree in the signature consisting of an order symbol (intended to be interpreted as the ancestor relation between nodes of the tree) and finitely many unary predicate symbols (intended to be interpreted as labellings of the nodes of the tree). In particular, nodes may have infinitely many children. A commonly occurring infinite tree is $(\mathbb{N}, <)$, also written ω. There are various kinds of interpretations depending on whether elements of the structure $\mathfrak{A}$ are represented by nodes or (finite) sets of nodes of $\mathfrak{T}$. The latter are called *(finite) set-interpretations* and the former we call *point-interpretations*[2] (Definitions 2.6 and 2.7). Moreover, each element of $\mathfrak{A}$ is coded by at least one *tuple* of sets (the common size of the tuples is called the *dimension* of the interpretation); and if every element of $\mathfrak{A}$ is coded by exactly one tuple of sets then the interpretation is called *injective*.

Why do interpretations in trees matter?

Interpretations allow one to transfer computational and logical properties from the interpreting structure to the interpreted structure.

Suppose that $\mathfrak{A}$ is 1-dim point-interpretable in $\mathfrak{T}$. Then there is a uniform way to translate MSO-formulas in the signature of $\mathfrak{A}$ to MSO-formulas in the signature of $\mathfrak{T}$ (just replace atoms by their definitions, and relativize quantifiers). So the MSO-theory of $\mathfrak{A}$ is computable in the MSO-theory of $\mathfrak{T}$. This explains the long term efforts to extend MSO-decidability from ω and $\mathfrak{T}_2$ to their expansions by unary predicates [8], [10], [22], [26].

Suppose $\mathfrak{A}$ is set-interpretable in $\mathfrak{T}$. Then there is a uniform way to translate FO-formulas in the signature of $\mathfrak{A}$ to MSO-formulas in the signature of $\mathfrak{T}$, and in this case the FO-theory of $\mathfrak{A}$ is computable in the MSO-theory of $\mathfrak{T}$. So the game here is to work out which structures $\mathfrak{A}$, known to have decidable FO-theory, are set-interpretable in trees with decidable MSO. Already Büchi noticed, in the language of automata, that the semigroup $(\mathbb{N}, +)$ is finite-set interpretable in ω. In modern terminology $(\mathbb{N}, +)$ is (finite-word) automatic. Also, the rational group $(\mathbb{Q}, +)$ is finite-set interpretable in a decidable expansion of ω (see [19]).

What is the expressive power of the interpretation method?

The research program that tries to outline the power of the interpretation method invariably has to prove non-interpretability results. Traditionally these were results about non-interpretability in expansions by unary predicates of linear orders (known as chains) [13], [18], [21]. However there are also non-interpretability results in graphs and trees. The Caucal hierarchy is a sequence $C_0, C_1 \dots$ of sets of graphs such that C_i is closed under 1-dim point-interpretations. There is a graph in C_i which is not 1-dim point-interpretable in any graph in C_j for $j < i$ [6].

[1] For instance, we learned in school that rational arithmetic is reducible to integer arithmetic by coding a rational by pairs of integers. In the terminology of this paper $(\mathbb{Q}, +, \times, =)$ is 2-dim point-interpretable in $(\mathbb{Z}, +, \times)$.

[2] In the literature these are sometimes called MSO-interpretations.

A consequence of a general result in [7] is that if $\mathbb{P}_f(\mathfrak{A})$ is 1-dim finite-set weak-MSO interpretable in binary-branching $\mathfrak{T}$ then already $\mathfrak{A}$ is 1-dim weak-MSO point-interpretable in $\mathfrak{T}$.[3] In the contrapositive this is a non-interpretability result; and indeed one of the main motivations in [7] is to reduce set-interpretability to the simpler point-interpretability. More recently, we learn that the real field $(\mathbb{R}, +, \times)$ is not set-interpretable in ω [1]. Looking at the proof we see that it goes through for any expansion of ω. Also, the rational group $(\mathbb{Q}, +)$ is not set-interpretable in ω (devoid of unary predicates) [27]. Much work in automatic structures is about proving that certain classes of structures are not (finite) set-interpretable in ω or $\mathfrak{T}_2$ (see [4], [15], [24]).

Overview

It is intuitively obvious that the more complex a tree the more it can interpret. We add weight to this contention by considering interpretations in trees with countably many infinite branches. We call these **scattered trees** since they are exactly the trees that do not embed the full binary tree. This name mimics the fact that linear orders that do not embed the rational order are called scattered orders, see [23]. The measure of complexity associates to every order and tree $\mathfrak{A}$ an ordinal $\mathrm{rank}(\mathfrak{A})$ (Definition 2.5). The rank of a tree reflects the structure of its infinite branches (similar to Cantor-Bendixson rank). All finite trees have rank 0, the line ω has rank 1, while any tree that embeds the full binary tree $\mathfrak{T}_2$ has rank ∞ (which is greater than all ordinals, and thus of maximum complexity). The rank of an order resembles Hausdorff ranks; thus ordinal ω^α has rank α. We prove, intuitively, that the rank of a tree $\mathfrak{T}$ limits the possible ranks of orders and trees interpretable in $\mathfrak{T}$. Compare this with the fact that every countable order is point interpretable in some expansion of the non-scattered tree $\mathfrak{T}_2$.

We emphasize that our results are of the form '$\mathfrak{A}$ is not interpretable in any expansion of $\mathfrak{T}$ by unary predicates', whether or not the expanded tree is decidable or even finitely presented. This is in line with previous investigations of the expressive power of interpretations in expansion of orders (called chains) [13], [18], [21] and in expansions of trees [7].

Technical contribution and related work

Point interpretations in scattered trees: These are simpler than set-interpretations and so proofs will not appear here. We prove that if an order or tree $\mathfrak{A}$ is 1-dim point-interpretable in a tree $\mathfrak{T}$ then $\mathrm{rank}(\mathfrak{A}) \leq \mathrm{rank}(\mathfrak{T})$. An immediate consequence is that neither $\mathbb{Q}$ nor $\mathfrak{T}_2$ is point-interpretable in any scattered tree. This is an analogue of the result that neither $\mathbb{Q}$ nor $\mathfrak{T}_2$ is point-interpretable in any scattered chain [21][Lemma 2.2].

[3] The structure $\mathbb{P}(\mathfrak{A})$ expands $(2^A, \subset)$ by the relations of $\mathfrak{A}$ on singleton sets. So $\mathbb{P}(\mathbb{Q})$ is $(2^\mathbb{Q}, \subset, <)$ where for $X, Y \in 2^\mathbb{Q}$, $X < Y$ if and only if $X = \{x\}$, $Y = \{y\}$ and the rational x is less than the rational y. The structure $\mathbb{P}_f(\mathfrak{A})$ is the substructure of $\mathbb{P}(\mathfrak{A})$ consisting of finite subsets of A. By weak-MSO interpretation we mean an interpretation in which additionally bound variables vary over finite sets.

Finite-set interpretations in expansions of ω: We prove that no scattered order or scattered tree of non-finite rank is finite-set interpretable in any expansion by unary predicates of ω (Section III). This generalizes an early breakthrough in the area of automatic structures that no scattered order or scattered tree of non-finite rank is finite-set interpretable in ω (devoid of any unary predicate) [9].

Finite-set interpretations in scattered trees: We prove that there is an ordinal function G such that no ordinal of rank $\geq G(\alpha)$ is finite-set interpretable in any scattered tree of rank $\leq \alpha$. We may take $G(n) = \omega^n$ for $0 \leq n < \omega$ and $G(\alpha) = \omega^{\alpha+1}$ for $\alpha \geq \omega$ (Theorem 4.1).

Injective set-interpretations in scattered trees: We prove that no ordinal of rank $\geq G(\alpha)$ is injectively set-interpretable in any scattered tree of rank $\leq \alpha$ (Corollary 5.3). So neither $\mathbb{P}(\mathbb{Q})$ nor $\mathbb{P}(\mathfrak{T}_2)$ are injectively set-interpretable in any scattered tree (Corollary 5.4). Compare this with the fact that $\mathbb{P}(\mathbb{Q})$ and $\mathbb{P}(\mathfrak{T}_2)$ are injectively set-interpretable in $\mathfrak{T}_2$. We conjecture that neither $\mathbb{P}(\mathbb{Q})$ nor $\mathbb{P}(\mathfrak{T}_2)$ are set-interpretable in any scattered tree.

Finitary-set interpretations in arbitrary trees: The previous result about injective set-interpretations follows from a theorem that is of independent interest. Even though set-interpretations allow one to interpret uncountable structures $\mathfrak{A}$, we do not study these. Instead consider the following general question: if $\mathfrak{A}$ is countable and set-interpretable in (not necessarily scattered) tree $\mathfrak{T}$ is $\mathfrak{A}$ finite-set interpretable in $\mathfrak{T}$? We do not solve this difficult problem. We establish a result of the same principle: if $\mathfrak{A}$ is countable and injectively set-interpretable in tree $\mathfrak{T}$ then there is an injective set interpretation of $\mathfrak{A}$ in $\mathfrak{T}$ for which the domain consists of tuples of finite sets and unlabelled-trees with finitely many infinite branches. Thus we manage to replace injective set-interpretations by these injective 'finitary'-set interpretations. Similar ideas also give: if $\mathfrak{A}$ is countable and injectively set-interpretable in a scattered tree $\mathfrak{T}$ of *finite rank* then $\mathfrak{A}$ is finite-set interpretable in $\mathfrak{T}$. We do not know if this holds for $\mathrm{rank}(\mathfrak{T}) = \omega$.

Hierachy strictness: For x a type of interpretation define $\mathcal{I}^x_\alpha$ as the set of structures that are x-interpretable in labelled trees of rank $\leq \alpha$. Clearly if $\alpha < \beta$ then $\mathcal{I}^x_\alpha \subseteq \mathcal{I}^x_\beta$. We have proven that these sets can be separated by ordinals; moreover, the bounds are tight. In summary, the hierarchies of injective set-interpretations, finite-set interpretations, and 1-dim point interpretations are strict: if $\alpha < \beta$ then $\mathcal{I}^x_\alpha \subsetneq \mathcal{I}^x_\beta$.

A note about technicalities. As far as the objects of study are concerned, the reader should have passing familiarity with linear orders and ordinal arithmetic (see [23]) and logical interpretations (see [14]). The central proof tool is Shelah's composition theorem (see [11] for a readable account). Ideas from the proof in Section III are used in Section IV. Section V can be read independently.

II. Definitions and Preliminaries

The structures in this paper have finite relational signatures Δ, typically of the form $\Delta_l := \{\prec, P_1, \ldots, P_l\}$ where $\prec$ is a binary predicate symbol and each P_i is a unary predicate symbol. In trees $\prec$ represents the ancestor relation and in orders $\prec$ represents element comparison. The tuple $\overline{P}$ represents a labelling of the domain by elements of $\{0,1\}^l$. In the next sections we define labelled linear orders (called chains) and labelled trees. Informally, the trees in this paper are subtrees of the countably-branching infinite tree of height ω, with unordered siblings, and expanded by finitely many unary predicates. The operations on these objects (sums) allow us to define the scattered orders and trees. Countable means finite or countably infinite. If unspecified structures in this paper are countable. We reserve $<$ for the ordering on ordinals. We write ω for the smallest infinite ordinal. The domain of a structure named $\mathfrak{A}$ is written A. The expansion of $\mathfrak{A}$ by predicates $\overline{V}$ is written $(\mathfrak{A}, \overline{V})$. If $B \subset A$ then write $(\mathfrak{B}, \overline{V})$ for the substructure of $(\mathfrak{A}, \overline{V})$ on domain B.

A. Labelled Orders

An *l-chain* is a labelled linear order $\mathfrak{L} = (L, \prec, P_1, \cdots, P_l)$. If $l = 0$ we talk about a linear order, or just order.

Definition 2.1 (Sums of l-chains): Given order $(\text{Ind}, \prec_{ind})$ and for every $i \in \text{Ind}$ an l-chain $\mathfrak{L}_i = (L_i, \prec_i, P_{i1}, \cdots, P_{il})$ the sum $\sum_{\text{Ind}} \mathfrak{L}_i$ is defined as the l-chain with domain $\bigcup_{i \in \text{Ind}} \{i\} \times L_i$, ordering $\prec$ defined by $(i, a) \prec (i', a')$ if and only if $i \prec_{ind} i'$ or $(i = i'$ and $a \prec_i a')$, and the kth unary predicate defined by $\bigcup_{i \in \text{Ind}} \{i\} \times P_{ik}$.

Write ω for the order type of the positive integers with the usual order, $\omega^\star$ for that of the negative integers, and $\mathbf{n}$ for the order type of the first n positive integers.

Definition 2.2 (scattered orders and ranks): Define sets of orders $\mathfrak{B}_\alpha$ and $\mathfrak{L}_\alpha$ by transfinite induction.

- $\mathfrak{B}_0 := \{\mathbf{1}\}$.
- $\mathfrak{L}_\alpha$ consists of $\sum_{\text{Ind}} \mathfrak{L}_i$ where Ind is finite and $\mathfrak{L}_i \in \mathfrak{B}_\alpha$.
- $\mathfrak{B}_\alpha$ consists of $\sum_{\text{Ind}} \mathfrak{L}_i$ where Ind has order type ω or $\omega^\star$ and for all $i \in \text{Ind}$, $\mathfrak{L}_i \in \bigcup_{\beta < \alpha} \mathfrak{L}_\beta$.

An order $\mathfrak{L}$ is *scattered* if it is in $\mathfrak{L}_\alpha$ for some α and the minimal such α is called the *rank* of $\mathfrak{L}$, written $\text{rank}(\mathfrak{L})$. The rank of a non-scattered order, written ∞, is defined to be greater than all countable ordinals.[4] The rank of a chain is defined as the rank of its underlying linear order.

The rank of a countable scattered order is countable and the rank of the ordinal written in Cantor-normal form $\sum_{i \leq m} \omega^{\alpha_i}$ is α_1. In particular ω^α is the least ordinal of rank α. The following pigeonhole principle for linear orders is used so often that we isolate it here.

Lemma 2.3 (partition property for orders): If the domain of an order $\mathfrak{L}$ is partitioned into finitely many pieces, then the order on at least one of the pieces has the same rank as that of $\mathfrak{L}$. If $\mathfrak{L}$ has order type ω^α then at least one of the pieces has order type ω^α.

B. Labelled Trees

An *l-labelled tree (or l-tree)* is a structure

$$\mathfrak{T} = (T, \prec, P_1, \ldots, P_l)$$

[4]Non-scattered orders can also be given an ordinal rank (see [23]) though we do not need this notion.

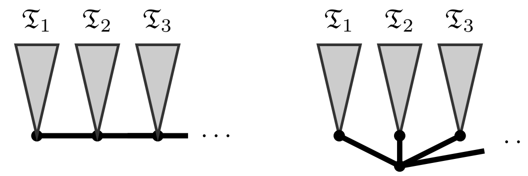

Fig. 1. An ω-sum (left), and ω-glueing.

where each $P_i \subseteq T$ and

- T is non-empty, partially ordered by $\prec$ with unique minimal element (the *root* r);
- every $\{y \in T \mid y \preceq x\}$ is a finite linear order.

Terminology. If $l = 0$ the tree is *unlabelled*. A node v is a *child* of a node u if $u \prec v$ and there is no z with $u \prec z \prec v$. If every node has finitely many children the tree is *finitely branching*; otherwise it is *countably-branching*. If b is the smallest integer with the property that every node has at most b children then the tree is *b-ary branching*. If for non-empty $T' \subset T$ the substructure $\mathfrak{T} \restriction T'$ is also a tree (ie. has a unique minimal element) then $\mathfrak{T}'$ *is a subtree of* $\mathfrak{T}$. A typical example of a subtree is *the subtree of* $\mathfrak{T}$ *rooted at* $x \in T$, written $\mathfrak{T}_{\succeq x}$ and defined by the domain $\{u \in T \mid u \succeq x\}$. Another example is given by a *downward-closed set* $I \subset T$ (ie. $t \prec i \in I \implies t \in I$). A subset X is a *branch* if it is linearly-ordered by $\prec$ and maximal with respect to set inclusion. Branches are subtrees and may be finite or infinite.

Definition 2.4 (tree sum): Given an unlabelled tree $(\text{Ind}, \prec_{ind})$ and for every $i \in \text{Ind}$ an l-tree $\mathfrak{T}_i = (T_i, \prec_i, P_{i1}, \cdots, P_{il})$ with root r_i, the *sum* $\sum_{Ind} \mathfrak{T}_i$ is the l-tree with domain $\bigcup_{i \in \text{Ind}} \{i\} \times T_i$; ancestor relation defined by $(i, a) \prec (i', a')$ if and only if $(i = i'$ and $a \prec_i a')$ or $(i \prec_{ind} i'$ and $a = r_i)$; and the kth unary predicate P_k defined by $\bigcup_{i \in \text{Ind}} \{i\} \times P_{ik}$.

Terminology. Suppose $\mathfrak{T} = \sum_{\text{Ind}} \mathfrak{T}_i$. If Ind is finite then $\mathfrak{T}$ is a *finite sum of* $\mathfrak{T}_i$*s*; if $(\text{Ind}, \prec_{ind})$ is a linear order of type ω then $\sum_{\text{Ind}} \mathfrak{T}_i$ is an *ω-sum of* $\mathfrak{T}_i$*s*. If $(\text{Ind}, \prec_{ind})$ consists of a root r and $n \leq \omega$ children, and $\mathfrak{T}_r$ is a singleton tree, then $\mathfrak{T}$ is an *(n)-glueing of* $\mathfrak{T}_i$*s*.

We now define ranks and scattered trees. The idea is that ω-sums and ω-glueings increase the rank while finite sums and finite glueings do not.

Definition 2.5 (scattered trees and ranks): Define families of unlabelled trees $\mathfrak{F}_\alpha$ and $\mathfrak{G}_\alpha$ by transfinite induction.

- $\mathfrak{F}_0$ consists of the unique tree with a single element.
- $\mathfrak{G}_\alpha$ consists of finite sums of $\mathfrak{T}_i$s with $\mathfrak{T}_i \in \mathfrak{F}_\alpha$.
- $\mathfrak{F}_\alpha$ consists of ω-sums and ω-glueings of $\mathfrak{T}_i$s where $\mathfrak{T}_i \in \bigcup_{\beta < \alpha} \mathfrak{G}_\beta$.

A tree $\mathfrak{T}$ is *scattered* if it is in $\mathfrak{G}_\alpha$ for some α and the minimal such α is called the *rank* of $\mathfrak{T}$. The rank of a non-scattered tree, written ∞, is defined to be greater than all ordinals. The rank of a labelled tree is defined as the rank of the unlabelled tree formed by removing the labels.

Thus the finite trees all have rank 0; the trees of rank 1 are ω-sums or ω-glueings of finite trees, and finite sums of these.

An *embedding* of unlabelled tree $\mathfrak{T}$ in unlabelled tree $\mathfrak{T}'$ is an injective function $f : T \to T'$ such that $x \prec y \implies f(x) \prec f(y)$. The following lore clarifies the status of scattered trees: tree $\mathfrak{T}$ is scattered if and only if the complete infinite binary tree does not embed in $\mathfrak{T}$ if and only if $\mathfrak{T}$ has countably many infinite branches.

C. MSO and Interpretations

Monadic second order (MSO) logic consists of *MSO-variables* $X, Y, Z, \ldots$ that are interpreted as subsets of the domain, and allows quantification over these variables. The non-logical symbols are those from signature Δ and the binary predicate symbol $\subseteq$ (representing set containment). Atomic formulas are those using symbols from Δ as well as those such as $X \subseteq Y$. Formulas are built from atomic formulas by applying Boolean connectives and universal and existential quantification of variables. We use abbreviations such as $X = Y, X \subset Y$ and $X \cap Y = \emptyset$. For convenience we may use FO-variables $x, y, z \cdots$ since 'X is a singleton' is definable in MSO.

A Δ-*structure* $\mathfrak{A}$ consists of a domain A and for each k-ary predicate symbol R from Δ a k-ary predicate $R^{\mathfrak{A}} \subseteq A^k$. We may drop the superscript in $R^{\mathfrak{A}}$ when we are only dealing with one structure. An *expansion* of Δ-structure $\mathfrak{A}$ by l many predicates is a structure $(\mathfrak{A}, P_1^{\mathfrak{A}}, \cdots, P_l^{\mathfrak{A}})$ in the signature $\Delta \cup \{P_1, \cdots, P_l\}$ where P_i are new unary predicate symbols. A Δ-*formula* is an MSO-formula in the signature Δ. Write $\mathbb{P}(A)$ for the set of all subsets of A. If $\varphi(X_1, \ldots, X_m)$ is a Δ-formula and $\mathfrak{A}$ is a Δ-structure then define

$$\varphi\mathfrak{A} := \{(S_1, \ldots, S_m) \in \mathbb{P}(A)^m \mid \mathfrak{A} \models \varphi(S_1, \ldots, S_m)\}.$$

What types of interpretations do we consider?

The most well studied interpretations in this field are point-interpretations — the interpreting formulas are MSO (bound variables vary over sets) but their free variables are, effectively, first order variables. We also consider a more general notion called *(finite-)set interpretations* — here the interpreting formulas have free monadic variables that vary over (finite) sets; bound variables vary over arbitrary subsets. So in (finite-) set interpretations elements are coded by (finite) sets. Moreover interpretations may be multi-dim (elements are coded by tuples of finite sets) and not necessarily injective (each element can be coded by more than one tuple).

Definition 2.6 (set interpretation): A *(d-dim) set interpretation* Γ *(in the signature Δ)* consists of Δ-formulas

$$\partial_\Gamma(\overline{X_1}), \mathrm{EQ}_\Gamma(\overline{X_1}, \overline{X_2}), \varphi^1_\Gamma(\overline{X_1}, \ldots, \overline{X_{r_1}}), \cdots, \varphi^k_\Gamma(\overline{X_1}, \ldots, \overline{X_{r_k}})$$

where k is an integer, $\overline{X_i}$ is a d-tuple of MSO-variables and in each formula all the variables named are distinct. Let $\mathfrak{T}$ be a Δ-structure. If $\mathrm{EQ}_\Gamma\mathfrak{T}$ is a congruence on the structure $(\partial_\Gamma\mathfrak{T}, \varphi^1_\Gamma\mathfrak{T}, \cdots, \varphi^k_\Gamma\mathfrak{T})$ then define

$$\Gamma\mathfrak{T} := (\partial_\Gamma\mathfrak{T}, \varphi^1_\Gamma\mathfrak{T}, \cdots, \varphi^k_\Gamma\mathfrak{T})_{/\mathrm{EQ}_\Gamma\mathfrak{T}},$$

and say that $\Gamma\mathfrak{T}$, and any structure isomorphic to it, is *set-interpretable in $\mathfrak{T}$ via* Γ. If $\mathrm{EQ}_\Gamma(\overline{X}, \overline{Y})$ is the formula $\bigwedge_{i \leq d} X_i = Y_i$ then Γ is an *injective* set-interpretation and $\Gamma\mathfrak{T}$ is *injectively* set-interpretable in $\mathfrak{T}$ via Γ. If $d = 1$ then we might stress that Γ is a 1-*dim set interpretation*.

We consider two particular types of interpretations depending, loosely speaking, on whether the free variables are taken to vary over elements of $\mathfrak{T}$ or finite subsets of $\mathfrak{T}$.

Definition 2.7 (point- and finite-set interpretations): Let Γ be a set interpretation. If in Definition 2.6 one replaces $\partial_\Gamma\mathfrak{T}$ everywhere it occurs by $\partial_\Gamma\mathfrak{T}$ restricted to d-tuples of finite sets (resp. singletons) then we say that $\Gamma\mathfrak{T}$ is (d-dim) *finite-set interpretable* (resp. *point-interpretable*) in $\mathfrak{T}$.

Remark 2.8: Note that because finiteness is not definable in our trees, $\mathfrak{A}$ may be finite-set interpretable in tree $\mathfrak{T}$ while not being set interpretable in $\mathfrak{T}$.

Example 2.9: For every $n < \omega$ there is an n-dim injective point-interpretation of ω^n in ω. For instance, $\delta_\Gamma(X_1, \cdots, X_n)$ states that each X_i is a singleton $\{x_i\}$; and $<_\Gamma (\overline{X}, \overline{Y})$ states that the least $i \leq n$ such that $x_i \neq y_i$ satisfies that $x_i < y_i$. Also, ω^n is injectively finite-set interpretable in ω. For instance, take $\delta_\Gamma(X)$ to be all sets of size n and order $X < Y$ if the smallest integer in the symmetric difference of X and Y is in X. The ordinal ω^ω is the least ordinal that is not finite-set interpretable in ω [9]. A corollary of Theorem 3.1 is that ω^ω is the least ordinal that is not finite-set interpretable in any expansion of ω by unary predicates.

Definition 2.10 (interpretation with parameters): Let Γ be an interpretation in the signature Δ such that every formula in Γ contains an additional m-tuple of free variables. Then Γ is called an *interpretation with m parameters*. Let $\overline{S}$ be an m-tuple of subsets of a tree $\mathfrak{T}$. Then $\Gamma(\mathfrak{T}, \overline{S})$, and any structure isomorphic to it, is said to be *interpretable in $\mathfrak{T}$ with m parameters via* Γ. A family $\{\mathfrak{B}_i\}$ is *interpretable in $\mathfrak{T}$ with m parameters via* Γ if for every member $\mathfrak{B}_i$ there exists an m-tuple $\overline{S}$ such that $\mathfrak{B}_i$ is isomorphic to $\Gamma(\mathfrak{T}, \overline{S})$.

Example 2.11: [20] The rational order $\mathbb{Q}$ is injectively 1-dim point-interpretable with one parameter in the full binary tree $\mathfrak{T}_2 := (\{0,1\}^*, \prec_{\text{pref}})$. The parameter may be taken as $R := \{0,1\}^*1$ and allows one to distinguish left and right children and so define the lexicographic ordering on $\{0,1\}^*$. Consequently, since every countable order embeds in the rational order, there is a 1-dim point-interpretation Γ such that the family of countable linear orders is injectively interpretable in $\mathfrak{T}_2$ with two parameters; one parameter is R and the other picks out the domain of the countable linear order.

The next proposition follows from the standard interpretation of the full countably-branching tree $\mathfrak{T}_\omega$ in $\mathfrak{T}_2$ [20].

Proposition 2.12: Every tree of rank α is 1-dim point-interpretable in a binary tree of rank α.

D. Composition Theorem for Tree-sums and Order-sums

Write Δ_l for the signature of order $\prec$ with l unary predicate $P_1, \cdots, P_l$ symbols. Thus a Δ_l-structure $\mathfrak{A}$ has the form $(A, \prec, P_1^{\mathfrak{A}}, \cdots, P_l^{\mathfrak{A}})$. The quantifier rank of a formula φ, denoted $\mathrm{qr}(\varphi)$, is the maximum depth of nesting of quantifiers in φ. For $r, l \in \mathbb{N}$ we denote by $\mathfrak{Form}^r_l$ the set of formulas of quantifier rank $\leq r$ and with free variables among $X_1, \ldots, X_l$ in signature $\{\prec\}$. For Δ_l-structures $\mathfrak{A}, \mathfrak{B}$ write $\mathfrak{A} \equiv_l^r \mathfrak{B}$ if for every $\varphi \in \mathfrak{Form}_l^r$,

$$\mathfrak{A} \models \varphi(P_1^{\mathfrak{A}}, \cdots, P_l^{\mathfrak{A}}) \text{ if and only if } \mathfrak{B} \models \varphi(P_1^{\mathfrak{B}}, \cdots, P_l^{\mathfrak{B}}).$$

Clearly $\equiv_l^r$ is an equivalence relation and the set $\mathfrak{Form}_l^r$ is infinite. Since the signature Δ_l is finite and relational the set $\mathfrak{Form}_l^r$ contains only finitely many semantically distinct formulas so there are only finitely many $\equiv_l^r$-classes of Δ_l-structures. The following lemma isolates maximally consistent formulas.

Lemma 2.13 (Hintikka lemma): For $r, l \in \mathbb{N}$, there is a *finite* set $H_l^r \subseteq \mathfrak{Form}_l^r$ such that:

1) For every Δ_l-structure $\mathfrak{A}$ there is a *unique* $\tau \in H_l^r$ with $\mathfrak{A} \models \tau(P_1^{\mathfrak{A}}, \cdots, P_l^{\mathfrak{A}})$.
2) If $\tau \in H_l^r$ and $\varphi \in \mathfrak{Form}_l^r$, then either $\tau \models \varphi$ or $\tau \models \neg\varphi$.[5]

Elements of H_l^r are called *(r,l)-Hintikka formulas*. For every r, l we fix an enumeration $\tau_1(\overline{X}), \ldots, \tau_{|H_l^r|}(\overline{X})$ of H_l^r.

Definition 2.14 (type of a structure): For Δ_l-structure $\mathfrak{A}$ write $\mathrm{Tp}_l^r(\mathfrak{A})$ for the unique $\tau(X_1, \cdots, X_l) \in H_l^r$ such that $\mathfrak{A} \models \tau(P_1^{\mathfrak{A}}, \cdots, P_l^{\mathfrak{A}})$, and call it the *$(r,l)$-type* of $\mathfrak{A}$.

Thus $\mathrm{Tp}_l^r(\mathfrak{A})$ effectively determines for which formulas $\varphi \in \mathfrak{Form}_l^r$ it holds that $\mathfrak{A} \models \varphi(P_1^{\mathfrak{A}}, \cdots, P_l^{\mathfrak{A}})$. Since l is often clear we may drop it and write $\mathrm{Tp}^r(\mathfrak{A})$ and r-type.

We now discuss increasingly informative versions of the composition theorem for MSO over tree-sums and order-sums, see [25] or [11], [12] for details. The first, lets call it *weak composition*, says that the (r,l)-type of a sum depends on the (r,l)-types of its summands.

Theorem 2.15 (weak composition for tree- and order-sums): For every Ind, if $\mathrm{Tp}_l^r(\mathfrak{A}_i) = \mathrm{Tp}_l^r(\mathfrak{A}'_i)$ for all $i \in$ Ind then $\mathrm{Tp}_l^r(\sum_{\text{Ind}} \mathfrak{A}_i) = \mathrm{Tp}_l^r(\sum_{\text{Ind}} \mathfrak{A}'_i)$.

We use the following consequence of the weak composition for orders: for every r, l there is associative binary operation $+$ on the set H_l^r such that for all l-orders $\mathfrak{L}_1, \mathfrak{L}_2$ the formula $\mathrm{Tp}_l^r(\mathfrak{L}_1) + \mathrm{Tp}_l^r(\mathfrak{L}_2)$ is identical with the (r,l)-type of their sum $\mathrm{Tp}_l^r(\sum_{\mathbf{2}} \mathfrak{L}_i)$.

Definition 2.16 (partition of Ind): Let $\{\mathfrak{A}_i\}_{i \in \text{Ind}}$ be a family of l-structures. The *H_l^r-partition (of Ind) induced by* $\{\mathfrak{A}_i\}_{i \in \textit{Ind}}$ is the $|H_l^r|$-tuple $\overline{Q}$ where $Q_k := \{i \in \text{Ind} \mid \mathrm{Tp}_l^r(\mathfrak{A}_i) = \tau_k\}$ and τ_k is the kth formula in the enumeration of H_l^r. The *H_l^r-expansion of $(\textit{Ind}, \prec_{ind})$ induced by* $\{\mathfrak{A}_i\}_{i \in \textit{Ind}}$ is the structure $(\text{Ind}, \prec_{\text{Ind}}, \overline{Q})$.

With this notation weak composition states that $\mathrm{Tp}_l^r(\sum_{\text{Ind}} \mathfrak{A}_i)$ is determined by the H_l^r-expansion induced by $\{\mathfrak{A}_i\}_{\text{Ind}}$. It turns out that $\mathrm{Tp}_l^r(\sum_{\text{Ind}} \mathfrak{A}_i)$ is already determined by some q-type of the H_l^r-expansion. In the next version q does not even depend on Ind:

Theorem 2.17 (composition for tree-sums): For every formula $\varphi \in \mathfrak{Form}_l^r$ there exists a formula $\theta(Y_1, \ldots, Y_{|H_l^r|})$[6] such that for every unlabelled tree $(\text{Ind}, \prec_{ind})$ and family $\{\mathfrak{T}_i\}_{i \in \text{Ind}}$ of l-trees

$$\sum_{\text{Ind}} \mathfrak{T}_i \models \varphi \iff (\text{Ind}, \prec_{ind}) \models \theta(\overline{Q})$$

where $\overline{Q}$ is the H_l^r-partition induced by $\{\mathfrak{T}_i\}_{i \in \text{Ind}}$. The quantifier-rank of θ depends only on r and l and so is written $q(r,l)$.

[5]Furthermore, H_l^r is computable from r, l, and there is an algorithm that given τ and φ decides between $\tau \models \varphi$ and $\tau \models \neg\varphi$. We do not use these facts.

[6]Moreover, θ is computable from φ, although we do not use this fact.

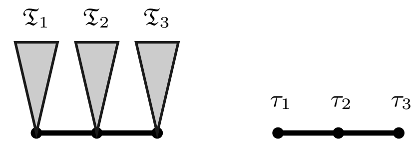

Fig. 2. Illustration of composition: The r-type of the tree (left) is determined by the q-type of the chain (right) where $\tau_i := \mathrm{Tp}^r(\mathfrak{T}_i)$. This q-type is called a projected type.

Projected Types: The notions in the remainder of this section are only used in Section IV. We visualise (Figure 2) the r-type of $\mathfrak{T}_i$ projected onto i and introduce notation to capture this. A *projected (r,l)-Hintikka formula* is a $(q(r,l), |H_l^r|)$-Hintikka formula. The *projected (r,l)-type of family $\{\mathfrak{T}_i\}_{i \in \textit{Ind}}$ of l-trees* is the $(q(r,l), |H_l^r|)$-type of the H_l^r-expansion of $(\text{Ind}, \prec_{ind})$ induced by $\{\mathfrak{T}_i\}_{\text{Ind}}$. Note that the projected type determines which quantifier-rank r formulas hold in $\sum_{\text{Ind}} \mathfrak{T}_i$.

When dealing with scattered trees Ind is often ω, so the type of an ω-sum of trees reduces to the type of an expansion of ω. Given a family $\{\mathfrak{T}_i\}_{i<n}$ of l-trees ($n \leq \omega$) and m-tuple $\overline{A}$ write $\mathrm{PrTp}(\overline{A})_{[i,j)}^r$ for the projected $(r, l+m)$-type of family $\{(\mathfrak{T}_k, \overline{A})\}_{k \in [i,j)}$. Under this notation we have

$$\mathrm{PrTp}(\overline{A})_{[0,n)}^r = \mathrm{PrTp}(\overline{A})_{[0,a)}^r + \mathrm{PrTp}(\overline{A})_{[a,n)}^r.$$

The following proposition says that if elements of $\Gamma \sum_\omega \mathfrak{T}_i$ are all in an 'interval sum' of $\mathfrak{T}$ then $\Gamma \sum_\omega \mathfrak{T}_i$ is interpretable in this interval.

Proposition 2.18: Suppose $\mathfrak{T} = \sum_\omega \mathfrak{T}_i$ is an l-tree and Γ is d-dim set (resp. point-, finite-set) interpretation. If there exists $a < b$ such that $\mathfrak{T} \models \partial_\Gamma(W_1, \cdots, W_d)$ implies $W_i \subseteq \sum_{i \in [a,b)} \mathfrak{T}_i$, then $\Gamma\mathfrak{T}$ is d-dim set (resp. point-, finite-set) interpretable in some expansion of $\sum_{i \in [a,b)} \mathfrak{T}_i$.

III. Finite-set Interpretations in expansions of ω

We write Δ_l for the signature consisting of $\prec$ and l predicate symbols.

Theorem 3.1:

1) For every **finite-set** interpretation Γ in the signature Δ_l there exists an integer N_Γ such that for every expansion $\mathcal{C}$ of ω by l unary predicates, if $\Gamma\mathcal{C}$ is a scattered order or scattered tree then its rank is at most N_Γ.
2) In particular, no scattered order or scattered tree of non-finite rank is finite-set interpretable in any expansion of ω by unary predicates.

The second item immediately follows from the first. Moreover, the bound is tight since every ordinal of finite rank is finite-set (even many-dim point-) interpretable in ω. It was shown in [9] that no ordinal of non-finite rank is finite-set interpretable in ω. That proof, which does not go through for expansions of ω, inspired Theorem 3.1.

From [3, Proposition 3.1] we can conclude that a countable structure that is set-interpretable with parameters in an expansion $\mathcal{C}$ of ω is already finite-set interpretable with parameters in $\mathcal{C}$. Thus no countable scattered order or tree of non-finite rank is *set* interpretable in any expansion of ω.

The proof of Theorem 3.1 uses the following *rank properties* on a class of structures C closed under isomorphism and under substructure:

1) Isomorphic structures in C have the same rank.
2) If $\mathfrak{A} \in$ C has rank α and A is partitioned into P_1, P_2, then at least one of $\mathfrak{A} \restriction P_1$ and $\mathfrak{A} \restriction P_2$ has rank α.
3) There is a 1-dim point-interpretation Γ with point parameters such that for every $\mathfrak{A}$ of finite rank k (infinite rank) there is a family of $k-1$ (infinitely many) structures of distinct ranks interpretable with parameters in $\mathfrak{A}$ via Γ.

Ranks on the class C of linear orders (Definition 2.2) satisfy these properties. For the third property use the fact that if $\mathfrak{L}$ has rank α then for every $\beta < \alpha$ there exist an open interval of $\mathfrak{L}$ of rank β. Infact the domain formula $\partial_\Gamma(x, p_1, p_2)$ may be defined as $p_1 \prec x \prec p_2$ where the p_is are parameters. Similarly the class C of forests (ie. sets of trees) is closed under substructure (unlike the class of trees) and we may define a ranking on forests (agreeing with the ranking on trees) satisfying these three properties as follows: the rank of a set $\mathfrak{S}$ of trees is the supremum of ranks of the trees in $\mathfrak{S}$. Theorem 3.1 is immediate from rank property 3) and the following.

Proposition 3.2: For every **set** interpretation Γ in the signature Δ_l with m parameters there exists an integer N_Γ such that if Γ interprets a family of scattered structures (orders or trees) $\{\mathcal{C}_i\}_{i \in I}$ with m-many **finite-set** parameters in some expansion $\mathcal{C}$ of ω then the number of distinct ranks amongst the ranks of $\{\mathcal{C}_i\}_{i \in I}$ is at most N_Γ.

Proof: To help readability we prove the proposition for 1-dim interpretations and $m = 1$ (one parameter). However the same proof goes through for d-dim interpretations and m parameters — replace variables and parameters ranging over subsets of ω by those ranging over d-tuples of subsets of ω. We sometimes mention m and d in the proof below to help the reader generalize.

Let q be an upper bound on the quantifier-rank of the formulas in the interpretation Γ. Define N_Γ greater than the number of $(q, l+m+2d)$-Hintikka formulas, namely $|H^q_{l+m+2d}|$. Take a family $\{\mathcal{C}_n\}_{n \in I}$ of scattered orders of distinct ranks that is interpreted with m-many parameters in an expansion $(\omega, \overline{p})$ via Γ. By assumption the m-many parameters are restricted to be finite subsets of ω.

Notation. For the rest of this proof we use lowercase $p_i, w_i, \ldots$ to refer to subsets of ω, and uppercase $L_i, D_i, \ldots$ to refer to sets of subsets of ω. For $z \subset \omega$ write $z[a,b)$ for $z \cap [a,b)$.

For every $n \in I$: fix a finite parameter $w_n \subset \omega$ so that $\Gamma(\omega, \overline{p}, w_n) \cong \mathcal{C}_n$; write D_n for the domain $\{x \mid (\omega, \overline{p}, w_n) \models \partial_\Gamma(x)\}$; write $=_n$ for the relation $\{(x,y) \mid (\omega, \overline{p}, w_n) \models \mathrm{EQ}_\Gamma(x,y)\}$; write $\prec_n$ for the ordering $\{(x,y) \mid (\omega, \overline{p}, w_n) \models x \prec_\Gamma y\}$; and write $\preceq_n$ for $\prec_n \cup =_n$.

For $z \subseteq \omega$ and $t \in \omega$ write $L_n(z,t)$ for the set of $x \in D_n$ such that $x[0,t) = z[0,t)$. That is, $L_n(z,t)$ consists of elements in the domain D_n that agree with z on the initial interval $[0,t)$. Note $L_n(z,0) = D_n$ for all z.

Claim 1. For every index $n \in I$ there is $z_n \subseteq \omega$ such that for every $t \in \omega$ the rank of $(L_n(z_n,t), \prec_n)/_{=_n}$ equals the rank of $\mathcal{C}_n$.

Proof of Claim 1. Fix n and suppose z_n has already been defined on the interval $[0,k)$ and for all $t \leq k$ the rank of $(L_n(z_n,t), \prec_n)/_{=_n}$ equals the rank of $\mathcal{C}_n$. Partition the sets in $L_n(z_n,k)$ into two classes V_0, V_1 depending on whether or not k is in the set. By rank property 2) above at least one of $(V_i, \prec_n)/_{=_n}$ has the same rank as $\mathcal{C}_n$; say the class is represented by $\epsilon \in \{0,1\}$. Put integer k in z_n if and only if $\epsilon = 1$. •

For every $n \in I$ fix z_n by Claim 1. Suppose, for a contradiction, there were more than N_Γ distinct ranks amongst the $\mathcal{C}_i$s. By Claim 1, and rank property 1) above, for every t there are more than N_Γ structures up to isomorphism amongst $(L_n(z_n,t), \prec_n)/_{=_n}$ $(n \in I)$. Pick finite $J_t \subset I$ of size greater than N_Γ indexing non-isomorphic structures (that is, $j \neq k \in J_t$ implies $(L_j(z_j,t), \prec_j)/_{=_j}$ is not isomorphic to $(L_k(z_k,t), \prec_k)/_{=_k}$). The following claim shows that the choice of N_Γ ensures that this is impossible.

Claim 2. There exists integer t and distinct indices $n, k \in J_t$ such that $(L_n(z_n,t), \prec_n)/_{=_n}$ and $(L_k(z_k,t), \prec_k)/_{=k}$ are isomorphic.

Proof of Claim 2. Write $(\omega, \overline{v})[0,t)$ for the structure $(\omega, \overline{v})$ restricted to domain $[0,t)$. By choice of N_Γ, for every t there exist distinct $n, k \in J_t$ such that $\mathrm{Tp}^q(\omega, \overline{p}, w_n, z_n, z_n)[0,t) = \mathrm{Tp}^q(\omega, \overline{p}, w_k, z_k, z_k)[0,t)$. Fix an integer t that is greater than all the integers in all the w_is for $i \in J_t$; and take n, k as in the previous sentence. So $w_n, w_k \subset [0,t)$. For $x, y \subset [t, \omega)$,

$$\begin{array}{ll} \mathrm{Tp}^q(\omega, \overline{p}, w_n, z_n[0,t) \cup x, z_n[0,t) \cup y) & = \\ \mathrm{Tp}^q(\omega, \overline{p}, w_n, z_n, z_n)[0,t) + \mathrm{Tp}^q(\omega, \overline{p}, \emptyset, x, y)[t, \omega) & = \\ \mathrm{Tp}^q(\omega, \overline{p}, w_k, z_k, z_k)[0,t) + \mathrm{Tp}^q(\omega, \overline{p}, \emptyset, x, y)[t, \omega) & = \\ \mathrm{Tp}^q(\omega, \overline{p}, w_k, z_k[0,t) \cup x, z_k[0,t) \cup y). & \end{array}$$

Immediately then

1) $(z_n[0,t) \cup x) \in L_n(z_n,t)$ iff $(z_k[0,t) \cup x) \in L_k(z_k,t)$
2) $(z_n[0,t) \cup x) \prec_n (z_n[0,t) \cup y)$ iff $(z_k[0,t) \cup x) \prec_k (z_k[0,t) \cup y)$, and
3) $(z_n[0,t) \cup x) =_n (z_n[0,t) \cup y)$ iff $(z_k[0,t) \cup x) =_k (z_k[0,t) \cup y)$.

These properties ensure that the map $z_n[0,t) \cup x \overset{\phi}{\mapsto} z_k[0,t) \cup x$ (where x ranges over subsets of $[t, \omega)$) induces an isomorphism $\Phi : (L_n(z_n,t), \prec_n)/_{=_n} \to (L_k(z_k,t), \prec_k)/_{=_k}$. Indeed Φ is a well-defined function by item 3); it is onto by item 1); and order-preserving by item 2). • ■

IV. Finite-set Interpretations in Scattered Trees

Theorem 4.1: There is an ordinal function G such that no ordinal of rank $\geq G(\alpha)$ is finite-set interpretable in any labelled tree of rank $\leq \alpha$. We may take $G(n) = \omega^n$ for $0 \leq n < \omega$ and $G(\alpha) = \omega^{\alpha+1}$ for $\alpha \geq \omega$.

A. Natural Sum and Product on Ordinals

The proof has some similarities with that of Proposition 3.2 but requires additional machinery, including the use of the natural-sum $\oplus$ (also called Hessenberg-sum) and natural-product $\otimes$ (also called Hausdorff-product) on ordinals. These operations were introduced in [5] and can be thought of as addition and multiplication of polynomials in ω. The natural-sum is a commutative, associative binary operation $\oplus$ on ordinals. Suppose $\alpha = \sum_{i<m} \omega^{\alpha_i}$ and $\beta = \sum_{j<n} \omega^{\beta_j}$ are in Cantor-normal-form. Then $\alpha \oplus \beta$ is defined as the sum (as in Definition 2.1) of all ω^{α_i} and ω^{β_j} arranged in non-increasing order. Similarly the natural-product is a commutative, associative binary operation $\otimes$ on ordinals. Define $\alpha \otimes \beta$ as the natural sum of all $\omega^{\alpha_i \oplus \beta_j}$. We implicitly use easy properties of natural ordinal arithmetic: for instance, the natural sum or natural product of countable ordinals is countable; if $\alpha, \beta < \omega^\gamma$ then $\alpha \oplus \beta < \omega^\gamma$; if $\alpha, \beta < \omega^{\omega^\gamma}$ then $\alpha \otimes \beta < \omega^{\omega^\gamma}$. Here is a central property of $\otimes$. A function $f : (\alpha_1 \times \cdots \times \alpha_k) \to \gamma$ is *coordinate-wise non-decreasing* if for all $(\delta_1, \cdots, \delta_k)$ in the domain of f and all $n \leq k$ and all ordinals δ with $\delta_n \leq \delta < \alpha_n$, it holds that $f(\delta_1, \cdots, \delta_k) \leq f(\delta_1, \cdots, \delta_{n-1}, \delta, \delta_{n+1}, \ldots, \delta_k)$.

Lemma 4.2: [5] If $f : (\alpha_1 \times \cdots \times \alpha_k) \to \gamma$ is onto and coordinate-wise non-decreasing then $\gamma \leq \alpha_1 \otimes \cdots \otimes \alpha_k$.

B. Proof of Theorem 4.1

We illustrate the proof for dimension 1 to get a function G_1 and remark at the end how to deal with dimension d. It is enough to find a function G_1 such that if ω^δ is 1-dim finite set interpretable in a tree of rank α then $\delta < G(\alpha)$ (this is because $\omega^{\text{rank}(\beta)} \leq \beta$ and the ordinals that are finite-set interpretable in expansions of $\mathfrak{T}$ are closed downwards). To this end, say ω^δ is finite-set interpretable in the l-tree $\mathfrak{T}$ of rank α via Γ. Write D for the domain of the interpretation; $\prec$ for $\prec_\Gamma \mathfrak{T}$, and $=_\Gamma$ for $\mathrm{EQ}_\Gamma \mathfrak{T}$. Since $(D, \prec)/_{=_\Gamma}$ is isomorphic to ω^δ, say via bijection f, for every ordinal $x < \omega^\delta$ pick a unique element in D from the equivalence class $f^{-1}(x)$ and call it the *code* of x. Let r be the largest quantifier-rank appearing in formulas of Γ. For the remainder write $\mathrm{PrTp}(\overline{A})_{[i,j)}$ instead of $\mathrm{PrTp}(\overline{A})^r_{[i,j)}$.

A note on the structure of the proof. We induct on α to bound δ. For the base case $G_1(0) := 1$ (since no infinite structure is interpretable in a finite one). For the remainder suppose $\alpha > 0$. We consider two cases: the first is that $\mathfrak{T}$ is an ω-sum of $\mathfrak{T}_i$s of lower rank (these we call Case 1 trees); and the second is that $\mathfrak{T}$ is a finite sum of Case 1 trees (called Case 2 trees). The case that $\mathfrak{T}$ is an ω-glueing is reduced to these cases by Proposition 2.12.

Notation. For the rest of this proof we use lowercase $p_i, w_i, \ldots$ to refer to subsets of T, and uppercase $L, R, \ldots$ to refer to sets of subsets of T.

Case 1. Say $\mathfrak{T}$ is an ω-sum of $\mathfrak{T}_i s$ of lower rank. For interval $[a, b)$ write $T[a, b)$ for the set $\cup_{i\in[a,b)} T_i$.

The **aim** is to bound every $\beta < \delta$ in terms of r, l, d and $G_1(\alpha')$ (for $\alpha' < \alpha$). So take arbitrary $\beta < \delta$ and write $w \in D$ for the code of ω^β. By Lemma 2.3 (partition property for orders) for all t there exists projected $(r, l + 2d)$-Hintikka formulas $\lambda_{w,t}$ and $\rho_{w,t}$ such that

$$\begin{aligned} D_{w,t} := \{x \in D \mid x \prec w \quad &\text{and} \quad \mathrm{PrTp}(w, x)_{[0,t)} = \lambda_{w,t} \\ &\text{and} \quad \mathrm{PrTp}(w, x)_{[t,\omega)} = \rho_{w,t}\} \end{aligned}$$

ordered by $\prec$ has order type ω^β. Define sets $L_{w,t}$ as

$$\{y \subset T[0, t) \mid y \text{ finite and } \mathrm{PrTp}(w, y)_{[0,t)} = \lambda_{w,t}\}$$

and $R_{w,t} := \{z \subset T[t, \omega) \mid z \text{ finite and } \mathrm{PrTp}(w, z)_{[t,\omega)} = \rho_{w,t}\}$.

In other words, every $x \in D_{w,t}$ satisfies that $x \cap [0, t) \in L_{w,t}$ and $x \cap [t, \omega) \in R_{w,t}$. Define binary relations $\prec_{L,w,t}$ and $=_{L,w,t}$ on $L_{w,t}$ by

$$\begin{aligned} y \prec_{L,w,t} y' &\text{ if } \exists z \in R_{w,t} \quad y \cup z \prec y' \cup z \\ y =_{L,w,t} y' &\text{ if } \exists z \in R_{w,t} \quad y \cup z =_\Gamma y' \cup z \end{aligned}$$

and binary relations $\prec_{R,w,t}$ and $=_{R,w,t}$ on $R_{w,t}$ by

$$\begin{aligned} y \prec_{R,w,t} y' &\text{ if } \exists z \in L_{w,t} \quad y \cup z \prec y' \cup z \\ y =_{R,w,t} y' &\text{ if } \exists z \in L_{w,t} \quad y \cup z =_\Gamma y' \cup z \end{aligned}$$

Let $\#w$ denote the smallest s such that $w \subset T[0, s]$. We will show that (‡) for $t > \#w$ we can replace $\exists z$ by $\forall z$ in the above definitions.

Lemma A. For $t > \#w$, both $(L_{w,t}, \prec_{L,w,t})/_{=_{L,w,t}}$ and $(R_{w,t}, \prec_{R,w,t})/_{=_{L,w,t}}$ are well-orders.

We defer the proof of Lemma A. Write $\beta_{L,w,t}$ and $\beta_{R,w,t}$ for their respective order types and note ($\star$) that each is at most ω^β. The map $\Phi : \beta_{L,w,t} \times \beta_{R,w,t} \to \omega^\beta$ (induced by $(y, z) \mapsto y \cup z$) is surjective by definition of $L_{w,t}$ and $R_{w,t}$. It is co-ordinate wise non-decreasing by (‡). Apply Lemma 4.2 to conclude ($\star\star$) that ordinal ω^β is at most $\beta_{L,w,t} \otimes \beta_{R,w,t}$.

The order $(L_{w,t}, \prec_{L,w,t})/_{=_L}$ is finite-set interpretable in $\mathfrak{T}[0, t)$ (by Proposition 2.18). This tree has some rank $\alpha' < \alpha$, so apply induction and conclude ($\star\star\star$) that $\beta_{L,w,t} < \omega^{G_1(\alpha')}$. All that is left is to bound $\beta_{R,w,t}$. For this we use a pigeonhole argument.

Lemma B. If w' codes $\omega^{\beta'}$ ($\beta' < \delta$) then for $t > \#w, \#w'$, if $\lambda_{w,t} = \lambda_{w',t}$ and $\rho_{w,t} = \rho_{w',t}$ then $\beta_{R,w,t} = \beta_{R,w',t}$.

We defer the proof of Lemma B. Define $\gamma(\alpha) := \sup_{\alpha'<\alpha} G_1(\alpha')$ and pick a sequence of ordinals γ_i such that $\gamma_0 := \gamma(\alpha)$ and (†) $\gamma_{i+1} > \gamma_0 \oplus \gamma_i$.

Let n be the number of projected $(r, l + 2d)$-Hintikka formulas and set $N := n^2 + 1$. Suppose, for a contradiction, that $\gamma_N \leq \beta$. Then each ordinal ω^{γ_i} ($0 \leq i \leq N$) has a code, say $w_i \in D$. For all $t > \max_{0\leq i\leq N} \#w_i$ there exist two indices $c < d \leq N$ such that $\lambda_{w_c,t} = \lambda_{w_d,t}$ and $\rho_{w_c,t} = \rho_{w_d,t}$ (by choice of N). Then

$$\begin{aligned} \omega^{\gamma_d} &\leq \beta_{L,w_d,t} \otimes \beta_{R,w_d,t} \text{ (by } \star\star) \\ &< \omega^{\gamma_0} \otimes \beta_{R,w_d,t} \text{ (by } \star\star\star) \\ &= \omega^{\gamma_0} \otimes \beta_{R,w_c,t} \text{ (by Lemma B)} \\ &\leq \omega^{\gamma_0} \otimes \omega^{\gamma_c} \text{ (by } \star) \end{aligned}$$

which equals $\omega^{\gamma_0 \oplus \gamma_c}$, contradicting (†). Thus $\beta < \gamma_N$ and we have bound the arbitrarily chosen β in terms of r, l, d and $G_1(\alpha')$ for $\alpha' < \alpha$. This achieves the **aim.** We conclude that the $\delta \leq \gamma_N$. Of course as r, l and d vary there is no bound on N. Thus define $G_1(\alpha) := \sup_i \gamma_i$ so that $\delta < G_1(\alpha)$.

To be concrete, take $\gamma_{i+1} := (\gamma_0 \oplus \gamma_i) + 1$, so $\sup_i \gamma_i = \gamma_0 \times \omega$. There are four cases: 1) if $1 \leq \alpha < \omega$ then $\gamma(\alpha) = G_1(\alpha - 1) = \omega^{\alpha-1}$ and so $G_1(\alpha) = \omega^\alpha$; 2) if $\alpha = \omega$ then $\gamma(\omega) = \sup_{n<\omega} \omega^n = \omega^\omega$ and so $G_1(\omega) = \omega^{\omega+1}$; 3) if $\alpha > \omega$ is a successor $\beta + 1$ then $\gamma(\alpha) = G_1(\beta) = \omega^{\beta+1} = \omega^\alpha$ and so $G_1(\alpha) = \omega^{\alpha+1}$; 4) if $\alpha > \omega$ is a limit ordinal then $\gamma(\alpha) = \sup_{\alpha'<\alpha} \omega^{\alpha'+1} = \omega^\alpha$ and so $G_1(\alpha) = \omega^{\alpha+1}$.

Proof of Lemma A. For $t > \#w$, $y, y' \in L_{w,t}$ and $z, z' \in R_{w,t}$ we claim $\mathrm{PrTp}(w, y, y', z)_{[0,\omega)} = \mathrm{PrTp}(w, y, y', z')_{[0,\omega)}$ and $\mathrm{PrTp}(w, y, z, z')_{[0,\omega)} = \mathrm{PrTp}(w, y', z, z')_{[0,\omega)}$. We prove the first equality (the second is similar). Recall that $+$ is the operation summing types of chains. Then $\mathrm{PrTp}(w, y, y', z)_{[0,\omega)}$ equals

$$\begin{aligned}
&\mathrm{PrTp}(w, y, y', z)_{[0,t)} + \mathrm{PrTp}(w, y, y', z)_{[t,\omega)} &=\\
&\mathrm{PrTp}(w, y, y', \emptyset)_{[0,t)} + \mathrm{PrTp}(\emptyset, \emptyset, \emptyset, z)_{[t,\omega)} &=\\
&\mathrm{PrTp}(w, y, y', z')_{[0,t)} + \mathrm{PrTp}(\emptyset, \emptyset, \emptyset, z')_{[t,\omega)} &=\\
&\mathrm{PrTp}(w, y, y', z')_{[0,t)} + \mathrm{PrTp}(w, y, y', z')_{[t,\omega)} &=\\
&\mathrm{PrTp}(w, y, y', z')_{[0,\omega)}.
\end{aligned}$$

To go from the second line to the third line use that $\mathrm{PrTp}(\emptyset, \emptyset, \emptyset, z)_{[t,\omega)}$ is determined by $\mathrm{PrTp}(\emptyset, z)_{[t,\omega)}$ which is $\rho_{w,t}$ by definition of $R_{w,t}$. Since $z' \in R_{w,t}$ then also $\mathrm{PrTp}(\emptyset, \emptyset, \emptyset, z')_{[t,\omega)}$ is determined by $\rho_{w,t}$.

Thus for $t > \#w$, we can replace $\exists$ by $\forall$ in the definitions of $\prec_{L,w,t}$ and $=_{L,w,t}$, and $\prec_{R,w,t}$ and $=_{R,w,t}$. For example, if $y, y' \in L_{w,t}$ and $z \in R_{w,t}$ and $y \cup z =_\Gamma y' \cup z$, then for all $z' \in R_{w,t}$ it holds that $y \cup z' =_\Gamma y' \cup z'$. It is now immediate that both $(L_{w,t}, \prec_{L,w,t})/_{=_{L,w,t}}$ and $(R_{w,t}, \prec_{R,w,t})/_{=_{L,w,t}}$ are well-defined well-orders.

Proof of Lemma B. First note $R_{w,t} = R_{w',t}$ since $\mathrm{PrTp}(w, z)_{[t,\omega)} = \mathrm{PrTp}(\emptyset, z)_{[t,\omega)} = \mathrm{PrTp}(w', z)_{[t,\omega)}$. Second by the reasoning in Lemma A and using $\lambda_{w,t} = \lambda_{w',t}$ we see that if $y_1, y_2 \in R_{w,t}$ and $z \in L_{w,t}$ with $y_1 \cup z \preceq y_2 \cup z$ then for all $z' \in L_{w',t}$ it holds that $y_1 \cup z' \preceq y_2 \cup z'$.

Case 2. Say $\mathfrak{T} = \sum_{i \in \mathrm{Ind}} \mathfrak{T}_i$ is a finite-sum of type 1 trees each of rank α. We prove $\omega^\delta < \omega^{G_1(\alpha)}$. By Lemma 2.3 (partition property for orders) we may assume that for every $i \in \mathrm{Ind}$ there is a type τ_i such that if $x \in D$ then $(\mathfrak{T}_i, x \cap T_i)$ has type τ_i. Define $D_i := \{x \cap T_i \mid x \in D\}$ and a binary relation $\prec_i$ on D_i by $x \prec_i y$ if $\exists z \in D\ x \cup (z \setminus T_i) \prec y \cup (z \setminus T_i)$, and similarly a binary relation $=_i$. By the same compositional reasoning as above $(D_i, \prec_i)/_{=_i}$ is well-ordered and $\leq \omega^\delta$, say of type η_i. By a fact similar to Proposition 2.18 ordinal η_i is finite-set interpretable in (Case 1 tree) $\mathfrak{T}_i$, thus $\eta_i < \omega^{G_1(\alpha)}$. The function sending $(x_1, \cdots, x_{|\mathrm{Ind}|}) \mapsto \cup x_i$ from $D_1 \times \cdots \times D_{|\mathrm{Ind}|} \to D$ induces a surjective co-ordinate wise non-decreasing function $\eta_1 \times \cdots \times \eta_{|\mathrm{Ind}|} \to \omega^\delta$. Thus $\omega^\delta \leq \eta_1 \otimes \cdots \otimes \eta_{|\mathrm{Ind}|}$. But since each $\eta_i < \omega^{G_1(\alpha)}$ and since $G_1(\alpha)$ is a power of ω, we see that $\omega^\delta < \omega^{G_1(\alpha)}$. Thus $\mathrm{rank}(\omega^\delta) = \delta < G_1(\alpha)$.

Finally, the same proof goes through for d-dim interpretations (replace variables by tuples of variables and make minor changes in notation). And the dimension d has no effect on G_d; that is $G_d = G_1$. Thus define $G := G_1$ to complete the proof. ■

The proof just presented can be adapted to scattered rank α trees of height $\omega+1$, in particular to completions $\widehat{\mathfrak{T}}$. We explain the terminology. A 'well-founded tree' is one in which every set of the form $\{y \mid y \preceq x\}$ is a (not necessarily finite) well-founded set. The *height* of a well-found tree is the supremum of the order types of these sets. Thus the trees as defined in Section II-B have height $\leq \omega$. Writing $[\mathfrak{T}]$ for the infinite branches of $\mathfrak{T}$ define the *completion* of a tree $\mathfrak{T}$, written $\widehat{\mathfrak{T}}$, as the partial order whose domain is $T \cup [\mathfrak{T}]$ and for which u is below v if either $u, v \in T$ and $u \prec^{\mathfrak{T}} v$, or $u \in T, v \in [\mathfrak{T}]$ and $u \in v$ (that is, u is a node on infinite branch v). If $\mathfrak{T}$ has height $\leq \omega$ then $[\mathfrak{T}]$ has height $\leq \omega + 1$. To define scattered trees of height $\omega + 1$ we replace ω-sums by $\omega + 1$-sums $\sum_{i<\omega+1} \mathfrak{T}_i$ where $\mathfrak{T}_\omega$ is a tree with exactly one element.

Corollary 4.3: Let G be the function from Theorem 4.1. No ordinal of rank $\geq G(\alpha)$ is finite-set interpretable in the completion of any labelled tree of rank $\leq \alpha$.

Proof Sketch: The composition theorem holds for well-founded trees, so we can run the proof of Theorem 4.1 with the following modifications: at the start of Case 1, partition the domain depending on whether the set hits T_ω or not. It is sufficient to deal with each of these domains. The latter case is as before. For the former case replace $[t, \omega)$ by $[t, \omega]$, and define $\#s$ as the smallest integer (exclude ω). In Lemma A for instance $\mathrm{PrTp}(w, y, y', z)_{[\omega,\omega]} = \mathrm{PrTp}(w, \emptyset, \emptyset, z)_{[\omega,\omega]}$, which is, now, also independent of the set w. This yields the same function G. ■

V. Replacing Set-interpretations by Simpler Interpretations

If $\mathfrak{A}$ is finite-set interpretable in $\mathfrak{T}$ then $\mathfrak{A}$ is necessarily countable. A general problem, that we do not solve, states:

Problem 5.1: If $\mathfrak{A}$ is countable and set-interpretable in (not necessarily scattered) tree $\mathfrak{T}$, is $\mathfrak{A}$ finite-set interpretable in $\mathfrak{T}$?

Here is our contribution.

Theorem 5.2: For every injective set interpretation Γ there exists injective set interpretation Γ_f such that (for labelled tree $\mathfrak{T}$ that is not necessarily scattered) if $\Gamma\mathfrak{T}$ is countable then

1) $\Gamma\mathfrak{T}$ is set interpretable in $\mathfrak{T}$ via Γ_f, and
2) every set in every tuple in the domain of $\Gamma_f\mathfrak{T}$ is either a finite subset of T or a finite union of infinite branches of $\mathfrak{T}$.

Corollary 5.3: Let G be the function from Theorem 4.1. No ordinal of rank $\geq G(\alpha)$ is injectively set-interpretable in any labelled tree of rank $\leq \alpha$.

Proof: Since a finite subset of $\widehat{\mathfrak{T}}$ is, modulo interpretation, a union of finite sets and finitely many infinite branches, Theorem 5.2 states that if $\mathfrak{A}$ is countable and injectively set interpretable in $\mathfrak{T}$ then $\mathfrak{A}$ is finite-set interpretable in the completion $\widehat{\mathfrak{T}}$. Apply Corollary 4.3. ■

Corollary 5.4: Neither $\mathbb{P}(\mathbb{Q})$ nor $\mathbb{P}(\mathfrak{T}_2)$ is injectively set-interpretable in any scattered tree.

Conjecture 5.5: Neither $\mathbb{P}(\mathbb{Q})$ nor $\mathbb{P}(\mathfrak{T}_2)$ is set-interpretable in any scattered tree.

A. Proof Plan

Given an MSO-formula φ the aim is to define an MSO-formula CODE such that for every tree $\mathfrak{T}$ for which $\varphi\mathfrak{T}$ is countable:

- CODE is an injective function with domain $\varphi\mathfrak{T}$,
- the range of CODE consists of tuples whose sets are either finite subsets of T or finite unions of finitely many infinite branches of $\mathfrak{T}$.

If φ is the domain formula of an injective set-interpretation Γ then define finitary interpretation Γ_f as follows: its domain formula expresses that $\overline{X}$ is in the range of CODE, its ith relation formula, say of arity n, expresses that there exist $\overline{Y}_j$s such that $\text{CODE}(\overline{Y}_j, \overline{X}_j)$ and $\phi^i_\Gamma(\overline{Y}_1, \cdots, \overline{Y}_n)$ (where ϕ^i_Γ is the ith relation formula in the interpretation Γ). Injectivity ensures that CODE is an isomorphism between $\Gamma\mathfrak{T}$ and $\Gamma_f\mathfrak{T}$.

In section V-B we discuss structural properties of $\varphi\mathfrak{T}$. In section V-C we provide a first coding that when applied to finitely-branching tree $\mathfrak{T}$ codes $\overline{V}$ (for $\mathfrak{T} \models \varphi(\overline{V})$) by a subtree with finitely many (finite and infinite) branches as well as a labelling of this subtree. In section V-D we sketch how to replace the labelling of the finitely many infinite branches by a tuple of finite sets. If the first coding is applied to a countably-branching tree $\mathfrak{T}$ we still obtain a subtree with finitely many infinite branches, but now it may also contain infinitely many finite branches. In the full version of the paper we show how to replace the labelled subtree consisting of the finite branches with a tuple of finite sets.

B. Structural Properties

The definitions and ideas of this section are from [2].

Definition 5.6 (U-trees and D-trees): Let $\mathfrak{T}$ be an l-tree, $\overline{V}$ a k-tuple and r an integer. If there exists $\overline{W} \neq \overline{V}$ (tuples of subsets of T) with $\mathrm{Tp}^r(\mathfrak{T}, \overline{W}) = \mathrm{Tp}^r(\mathfrak{T}, \overline{V})$ then call $(\mathfrak{T}, \overline{V})$ a *D-tree wrt.* r, k. Otherwise call $(\mathfrak{T}, \overline{V})$ a *U-tree wrt.* r, k.

Definition 5.7 (trunk): Define $\mathrm{trunk}^r(\mathfrak{T}, \overline{V})$ as the set of nodes $u \in T$ such that the subtree of $(\mathfrak{T}, \overline{V})$ rooted at u is a D-tree wrt. r, k.

Lemma 5.8: The set $\mathrm{trunk}^r(\mathfrak{T}, \overline{V})$ is MSO-definable in $(\mathfrak{T}, \overline{V})$ and downward closed.

We can decompose a tree along a downward closed set:

Definition 5.9 (tree decomposition): Let $\mathfrak{T}$ be an l-tree and $I \subset T$ a downward closed set. For $i \in I$ define T_i as those $t \in T$ such that $i \preceq t$ and there is no $i' \in I$ with $i \prec i' \preceq t$. As usual write $\mathfrak{T}_i$ for the substructure of $\mathfrak{T}$ restricted to T_i. We call the family $\{\mathfrak{T}_i\}_{i \in I}$ the *I-decomposition of* $\mathfrak{T}$.

If $\{\mathfrak{T}_i\}_{i \in I}$ is the I-decomposition of $\mathfrak{T}$ then $\mathfrak{T}$ is isomorphic to $\sum_{i \in I} \mathfrak{T}_i$ and the H^r_l-partition of I induced by $\{\mathfrak{T}_i\}_{i \in I}$ is definable in $\mathfrak{T}$ expanded by I.

Lemma 5.10 (interpretability of H^r_l-expansion): For every r, l there is a 1-dim injective point interpretation Γ such that for every tree $\mathfrak{T}$ and downward closed $I \subset T$ — writing $\{\mathfrak{T}_i\}_{i \in I}$ for the I-decomposition of $\mathfrak{T}$ — $\Gamma(\mathfrak{T}, I)$ is isomorphic to the H^r_l-expansion of $(I, \prec)$ induced by $\{\mathfrak{T}_i\}_{i \in I}$.

Proposition 5.11 (trunk is finitary): Let φ be a formula of quantifier-rank r and $\mathfrak{T}$ a labelled tree. If $\varphi\mathfrak{T}$ is countable then for every $\overline{V}$ satisfying φ in $\mathfrak{T}$ — writing $\{\mathfrak{T}_i\}$ for the $\mathrm{trunk}^r(\mathfrak{T}, \overline{V})$-decomposition of $(\mathfrak{T}, \overline{V})$ —

1) All but finitely many $\mathfrak{T}_i$s are U-trees.
2) The set $\mathrm{trunk}^r(\mathfrak{T}, \overline{V})$ is a union of a finite set and a finite set of infinite branches.

C. First Coding

Suppose $\mathfrak{T}$ is an l-tree, $\varphi\mathfrak{T}$ is countable and r is the quantifier-rank of $\varphi(X_1, \cdots, X_m)$. For $\overline{V}$ such that $\mathfrak{T} \models \varphi(\overline{V})$ let $\{\mathfrak{T}_i\}$ be the $\mathrm{trunk} := \mathrm{trunk}^r(\mathfrak{T}, \overline{V})$-decomposition of tree $(\mathfrak{T}, \overline{V})$. Write E for finite set of $i \in \mathrm{trunk}$ such that $\mathfrak{T}_i$ is a D-tree. Write BUDS for the set of children of the root of $\mathfrak{T}_i$ for $i \in E$. We can code m-tuple $\overline{V}$ by the following data:

1) a pair (F, B) where F, B partition trunk, F is a finite set, and B is a finite set of infinite branches,
2) the H^r_{l+m}-partition of F induced by $\{\mathfrak{T}_i\}_{i \in F}$,
3) the H^r_{l+m}-partition of B induced by $\{\mathfrak{T}_i\}_{i \in B}$,
4) the H^r_{l+m}-partition of BUDS induced by $\{(\mathfrak{T}_{\succeq s}, \overline{V})\}_{s \in \text{BUDS}}$.

This coding is injective: we argue that the coding of $\overline{V}$ uniquely determines $\overline{V}$. Consider $j \in \mathrm{trunk}$. If $\mathfrak{T}_j$ is a U-tree then $\overline{V} \cap T_j$ is determined by the data in 2) and 3); if $\mathfrak{T}_j$ is a D-tree then consider $i \in \mathfrak{T}_{\succeq s}$ for some child $s \in$ BUDS of the root of $\mathfrak{T}_j$. Then $\overline{V} \cap T_{\succeq s}$ is determined by the data in 4) since it is a U-tree. Moreover $\overline{V} \cap \{j\}$ is determined by 2) and 3). The coding is MSO-definable: indeed, F, B are definable from trunk which is definable by Lemma 5.8, partitions are definable by Lemma 5.10, and E and BUDS are definable since the set of D-trees (wrt. r, k) are definable.

The predicates in 1) and 2) are finitary: F is a finite set, its partition is a tuple of finite sets, and B is a finite set of infinite branches. Two tasks remain.

Task 1**.** The predicates in 3) label the subtree on domain B. In Section V-D we sketch how to code this labelling by a tuple of *finite* sets.

Task 2**.** If $\mathfrak{T}$ is finitely-branching then each predicate in 4) is a finite set. In the full version we show how, if $\mathfrak{T}$ is countably-branching, to code the possibly infinite set BUDS and its H^r_{l+m}-partition by a tuple of finite sets.

D. Dealing with Infinite Labelled Branches (Task 1)

Proposition 5.12: [2] For Δ_l-formula $\varphi(X_1, \cdots, X_m)$ there is ψ such that for every l-tree $\mathfrak{T}$ and every branch I_b of $\mathfrak{T}$ — writing $\mathcal{C} = (I_b, \prec, \overline{Q})$ for the H^{r+m}_l-expansion of I_b induced by the I_b-decomposition of $\mathfrak{T}$ — the following holds for all $\overline{W}$: $\mathcal{C} \models \psi(\overline{W})$ if and only if there exists $\overline{V}$ such that

1) $\mathfrak{T} \models \varphi(\overline{V})$, and
2) $\overline{W}$ is the H^r_{l+m}- partition of I_b induced by the I_b-decomposition of $(\mathfrak{T}, \overline{V})$.

In particular if $\varphi\mathfrak{T}$ is countable then $\psi\mathcal{C}$ is countable.

Definition 5.13: For sets $X, Y \subset T$, write $X =_{\text{end}} Y$ to mean that the symmetric difference of X and Y is finite (and say that X and Y have the same *end*). This notion extends to k-tuples: write $\overline{X} =_{\text{end}} \overline{Y}$ if $X_i =_{\text{end}} Y_i$ for all $i \leq k$.

Proposition 5.14 (definable ends in ω): For every Δ_s-formula $\psi(X_1, \cdots, X_n)$ there exist a constant $M := M(s, n)$ and formulas $\Psi_1(\overline{X}), \cdots, \Psi_M(\overline{X})$ such that for every s-chain $\mathcal{C}$ over ω there exist M-many tuples $\overline{W}_1, \cdots, \overline{W}_M$ such that if $\varphi\mathcal{C}$ is countable then

1) $\mathcal{C} \models \psi(\overline{V})$ implies there is $i \leq M$ with $\overline{W}_i =_{\text{end}} \overline{V}$.
2) the only tuple satisfied by Ψ_i in $\mathcal{C}$ is $\overline{W}_i$.

The first item appears in [16]. The second uses the selection property for ω-chains: a formula $\alpha(\overline{X})$ is a selector for formula $\beta(\overline{X})$ over a class of structures $\mathcal{C}$ if the following conditions hold in $\mathcal{C}$: 1) there is at most one $\overline{X}$ with $\alpha(\overline{X})$; 2) for all $\overline{X}$ if $\alpha(\overline{X})$ then $\beta(\overline{X})$; 3) if there exists $\overline{Y}$ with $\beta(\overline{Y})$ then there exists $\overline{X}$ with $\alpha(\overline{X})$. Every MSO-formula β has a selector α, also an MSO-formula, over the class of all expansions of ω by unary predicates, see [17], [22]. Since a branch of $\mathfrak{T}$ is isomorphic to ω, from Propositions 5.12 and 5.14, and Lemma 5.10 we get:

Proposition 5.15 (definable ends along a branch): For every Δ_l-formula $\varphi(X_1, \cdots, X_m)$ of quantifier rank r there exists a constant M and MSO-formulas $\Phi_1, \cdots, \Phi_M$ such that for every l-tree $\mathfrak{T}$ with $\varphi\mathfrak{T}$ countable, if $I_b \subset T$ is an infinite branch of $\mathfrak{T}$ and $\{\mathfrak{T}_i\}_{i \in I_b}$ is the I_b-decomposition of $\mathfrak{T}$ then there exist M-many tuples $\overline{W}_1, \cdots, \overline{W}_M$ over I_b such that

1) $\mathfrak{T} \models \varphi(\overline{V})$ implies some $\overline{W}_j$ has the same end as the H^r_{l+m}-partition of I_b induced by $\{(\mathfrak{T}_i, \overline{V})\}_{i \in I_b}$.
2) $\overline{W}_i$ is the unique tuple satisfied by Φ_i in $(\mathfrak{T}, I_b)$.

We sketch how to finish Task 1. Recall we have to encode, by a tuple of finite sets, the H^r_{l+m}-partition of B induced by $\{\mathfrak{T}_i\}_{i \in B}$ (where $\{\mathfrak{T}_i\}$ is the $\text{trunk}^r(\mathfrak{T}, \overline{V})$-decomposition of $(\mathfrak{T}, \overline{V})$). One set stores, for each of the finitely many branches I in B, an index $n \leq M$ such that the H^r_{l+m}-partition of I induced by $\{\mathfrak{T}_i\}_{i \in I}$ has the same end as tuple defined by Φ_n. The same set stores from which point of I onwards the tuples agree. In fact the index for I can be coded as a label of a definable node y of I that is on no other branch of B (see formula ϵ below). Also mark the $\prec$-least node z of I above y from which point on the tuples agree. Finally we need to store the restriction of the partition to all nodes below z. This data can be stored in a tuple of finite sets, and determines the H^r_{l+m}-partition of B. We now argue that it is MSO-definable.

Formally, apply Proposition 5.15 to the domain formula of Γ. This gives M and $\Phi_1, \cdots, \Phi_M$. For $n \leq M$ define formula $\mu_n(\overline{V}, I)$ stating that I is an infinite branch of $\text{trunk}^r(\mathfrak{T}, \overline{V})$ and n is the least integer with the property that the unique tuple $\overline{W}_n$ over I satisfying Φ_n has the same end as the H^r_{l+m}-partition of I induced by $\{\mathfrak{T}_i\}_{i \in I}$; if furthermore $\overline{W}_n$ and the mentioned H^r_{l+m}-partition of I agree on $\{i \in I \mid z \preceq i\}$ then write $\nu_n(\overline{V}, I, z)$. Define an auxiliary formula $\epsilon(X, x)$ stating that X is an infinite branch and x is the $\prec$-minimal element such that x is on X and no two elements of B above x are $\prec$-incomparable.

Finally, code the H^r_{l+m}-partition of B induced by $\{\mathfrak{T}_i\}_{i \in B}$ by $|H^r_{l+m}|$-tuple of finite sets $\overline{H}$ and M-tuple of finite sets $\overline{G}$:

1) For every $n \leq M$: $z \in G_n$ if and only there exists I, y such that $\mu_n(\overline{V}, I)$ and $\epsilon(I, y)$ and z is the $\prec$-minimal element such that $y \preceq z$ and $\nu_n(\overline{V}, I, z)$;
2) $\overline{H}$ is the restriction of the H^r_{l+m}-partition of B induced by $\{\mathfrak{T}_i\}_{i \in B}$ to the finite set $\bigvee_{n \leq M} \{u \mid \exists z \in G_n u \preceq z\}$.

ACKNOWLEDGMENTS

Vince Bárány, Lukasz Kaiser, Phillip Schlict, funding from the European Science Foundation for the activity entitled 'Games for Design and Verification', FWF NFN Nr. S11407-N23 (RiSE), ERC 279307: Graph Games, PROSEED WWTF Nr. ICT 10-050, ARISE FWF Nr. S11403-N23.

REFERENCES

[1] F. Abu Zaid, E. Grädel, and Ł. Kaiser, "The field of reals is not omega-automatic," in *STACS*, 2012.
[2] V. Bárány, L. Kaiser, and A. Rabinovich, "Expressing cardinality quantifiers in monadic second-order logic over trees," *Fundamentae Informaticae*, vol. 100, pp. 1–18, 2010.
[3] V. Bárány, L. Kaiser, and S. Rubin, "Cardinality and counting quantifiers on omega-automatic structures," in *STACS*, 2008.
[4] A. Blumensath and E. Grädel, "Automatic structures," in *15th Symposium on Logic in Computer Science* (LICS), 2000, pp. 51–62.
[5] P. W. Carruth, "Arithmetic of ordinals with applications to the theory of ordered abelian groups," *Bull. AMS*, vol. 48, pp. 262–271, 1942.
[6] D. Caucal, "On infinite terms having a decidable monadic theory," in *MFCS*, ser. Lecture Notes in Comput. Sci., vol. 2420, 2002, pp. 165–176.
[7] T. Colcombet and C. Löding, "Transforming structures by set interpretations," *Logical Methods in Computer Science*, vol. 3, no. 2, 2007.
[8] B. Courcelle, "The monadic second-order logic of graphs ix: Machines and their behaviours," *Theoretical Computer Science*, vol. 151, no. 1, pp. 125–162, 1995.
[9] C. Delhommé, "Automaticité des ordinaux et des graphes homogènes," *Comptes Rendus Mathematique*, vol. 339, no. 1, pp. 5–10, 2004.
[10] C. Elgot and M. Rabin, "Decidability of extensions of theory of successor," *J. Symb. Log.*, vol. 31, no. 2, pp. 169–181, 1966.
[11] Y. Gurevich, "Monadic second-order theories," in *Model-Theoretic Logics*. Springer, 1985, pp. 479–506.
[12] Y. Gurevich and S. Shelah, "Rabin's uniformization problem," *J. Symb. Log.*, vol. 48, pp. 1105–1119, 1983.
[13] ——, "On the strength of the interpretation method," *J. Symb. Log.*, vol. 54, no. 2, pp. 305–323, 1989.
[14] W. A. Hodges, *Model Theory*. Cambridge University Press, 1993.
[15] B. Khoussainov and A. Nerode, "Automatic presentations of structures," *Lecture Notes in Computer Science*, vol. 960, pp. 367–392, 1995.
[16] D. Kuske and M. Lohrey, "First-order and counting theories of omega-automatic structures," *J. Symb. Log.*, vol. 73, no. 1, pp. 129–150, 2008.
[17] S. Lifsches and S. Shelah, "Uniformization and skolem functions in the class of trees," *J. Symb. Log.*, vol. 63, pp. 103–127, 1998.
[18] ——, "Random graphs in the monadic theory of order," *Archive for Math Logic*, pp. 273–312, 1999.
[19] A. Nies, "Describing groups," *Bull. Symb. Log.*, vol. 13, no. 3, pp. 305–339, 2007.
[20] M. Rabin, "Decidability of second-order theories and automata on infinite trees," *AMS Transactions*, vol. 141, pp. 1–35, 1969.
[21] A. Rabinovich, "The full binary tree cannot be interpreted in a chain," *J. Symb. Log.*, vol. 75, pp. 1489–1498, 2010.
[22] ——, "On decidability of monadic logic of order over the naturals extended by monadic predicates," *Inf. Comput.*, vol. 205, no. 6, pp. 870–889, 2007.
[23] J. Rosenstein, *Linear Orderings*. Academic Press, 1982.
[24] S. Rubin, "Automata presenting structures: a survey of the finite string case," *Bull. Symb. Log.*, vol. 14, no. 2, pp. 169–209, 2008.
[25] S. Shelah, "The monadic theory of order," *Annals of Mathematics*, pp. 349–419, 1975.
[26] W. Thomas, "Constructing infinite graphs with a decidable MSO-theory," in *MFCS 2003*, ser. LNCS, vol. 2747, 2003, pp. 113–124.
[27] T. Tsankov, "The additive group of the rationals does not have an automatic presentation," *JSL*, vol. 76, no. 4, pp. 1341–1351, 2011.