Title

Distality in Combinatorics and Continuous Logic

Author

Anderson, Aaron William

Publication Date

2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Distality in Combinatorics and Continuous Logic

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Mathematics

by

Aaron William Anderson

2024

ABSTRACT OF THE DISSERTATION

Distality in Combinatorics and Continuous Logic

by

Aaron William Anderson
Doctor of Philosophy in Mathematics
University of California, Los Angeles, 2024
Professor Artem Chernikov, Chair

A central goal of modern model theory is to classify and understand first-order structures through studying the combinatorial properties of their definable sets, Conversely, model-theoretic properties can shed light on the behavior of interesting families of sets, such as semialgebraic families, which happen to be definable in well-behaved structures. This dissertation studies one such model-theoretic property, distality, as it manifests in the combinatorics of definable sets, and in continuous logic, where definable sets are generalized to real-valued functions.

In Chapter 2, we calculate explicit bounds on the sizes of distal cell decompositions for definable sets in a variety of distal structures. These lead to incidence combinatorics bounds in exponential-polynomial and p-adic settings. In Chapter 3 and Chapter 4, we develop a theory of distality in continuous logic. This begins with a study of the combinatorics of NIP metric structures, building on earlier work by Ben Yaacov and results from statistical learning theory. We also develop a theory of generically stable Keisler measures in continuous logic, allowing us to generalize combinatorial statements from just pertaining to finite counting measures. We then generalize many definitions of distal structures to the continuous logic context, showing that under an NIP assumption, they are all equivalent. These allow us to study distal metric structures through the perspectives of indiscernible sequences, strong honest definitions, distal cell decompositions, Keisler measures, and an analytic regularity lemma. Finally in Chapter 5, joint work with Itaï Ben Yaacov, we present examples of distal metric structures that are unique to continuous logic, including real closed metric valued fields and dual linear continua.

The dissertation of Aaron William Anderson is approved.

Igor Pak

Itay Neeman

Matthias J. Aschenbrenner

Artem Chernikov, Committee Chair

University of California, Los Angeles

2024

To Paula and Robert,

to Kate, Robert, Shirley, Irvin, Sally, and Albert

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Firstly, I would like to thank Artem Chernikov for years of advising. Starting when I was still an undergraduate, he has always been generous with his time, and our long meetings have fundamentally shaped my understanding of model theory. His career guidance has served me well, and I am deeply grateful.

I would also like to thank Itaï Ben Yaacov and the Institut Camille Jordan for hosting me in Lyon, and the French Embassy and their Chateaubriand fellowship for funding my visit. Those four months of meetings with Itaï helped immensely to clarify my continuous logic philosophy. This influence is present in all of the continuous logic chapters of this thesis, but most of all, Chapter 5, which Itaï co-authored. I thank Todor Tsankov for pointing me towards dual linear continua, and Itaï for exploring them with me in detail.

I would like to thank the logicians who have helped me on this journey. Those senior to me have always been supportive and provided useful advice, and my peers have patiently helped refine my ideas and restore my enthusiasm. All have welcomed me into this wonderful research community, whether at conferences or at logic coffee in the Montague room. Thank you in particular to my dissertation committee, to the cohort of logic grad students and postdocs who have made my time at UCLA so enjoyable, and to Diego Bejarano Rayo, James Hanson, Sergei Starchenko, Andrew Marks, Yiannis Moschovakis, and Pierre Touchard.

I am grateful to my professors from my undergraduate time at Caltech, including Alexander Kechris, Martino Lupini, and Adam Sheffer, for guiding my first research experiences and introducing me to ideas such as incidence combinatorics and continuous logic.

I thank Canada/USA Mathcamp and the Olga Radko Math Circle, for showing me both the wide variety in mathematics and the joy to be found in teaching it.

I am perpetually grateful to my family for setting me on this path in the first place, and supporting me in every endeavour I have considered. Thank you to all of you, and to Nadine and the Bradbury family, for supporting and encouraging me, particularly through the hardest times of the pandemic.

I also thank the UCLA Logic Center, the Chateaubriand Fellowship, the UCLA Dissertation Year Fellowship, and NSF grants DMS-1651321 and DMS-2246598 for financial support throughout my time in graduate school.

VITA

2018	Teaching assistant for Math 6a (Discrete Mathematics), California Institute of Technology.
2018	B.S. (Mathematics), California Institute of Technology.
2018-2022	Teaching assistant and graduate student researcher, Department of Mathematics, University of California, Los Angeles.
2020	M.A. (Mathematics), University of California, Los Angeles.
2023	Chateaubriand Fellowship, Institut Camille Jordan, Université Claude Bernard Lyon 1.
2023-2024	Dissertation Year Fellowship, University of California, Los Angeles.

PUBLICATIONS

Aaron Anderson. "The Fraïssé Limit of Matrix Algebras with the Rank Metric.", 2017.

Aaron Anderson. "Combinatorial Bounds in Distal Structures." The Journal of Symbolic Logic, pp. 1-29, 2023.

Aaron Anderson. "Fuzzy VC Combinatorics and Distality in Continuous Logic.", 2023.

Aaron Anderson. “Generically Stable Measures and Distal Regularity in Continuous Logic.”, 2023.

CHAPTER 1

Introduction

The school of model theory called *neostability* classifies first-order structures based on combinatorial properties. Several of these properties can be defined in terms of the graph-theoretic structure of *definable sets*. If M is a structure in a language $\mathcal{L}$, then a definable set is the solution set of a formula written with the symbols of $\mathcal{L}$ and logical connectives. For instance, when M is an algebraically closed field, these are the constructible sets, and when M is the ordered field $\mathbb{R}$, these are the semialgebraic sets. These definable sets are considerably better-behaved than arbitrary subsets of $\mathbb{C}^n$ or $\mathbb{R}^n$, as our deep understanding of algebraic geometry indicates. Neostability generalizes this setting with several dividing lines between such "tame" structures, where all definable sets are well-behaved, and "wild" structures, the domain of Gödelian self-reference and paradoxical decompositions.

Of these tame kinds of structure, the most work has been done studying stable structures. These are the structures such as algebraically closed fields, where no definable relation linearly orders an infinite set. Generalizing stability, both $\mathbb{R}$ and $\mathbb{C}$ are examples of NIP structures, defined by all definable binary relations having finite VC-dimension. The assumption of NIP is enough to prove a suite of combinatorial properties of definable sets, but there are dividing lines describing the combinatorics of semialgebraic sets in finer detail.

Some of these properties can be phrased in terms of *incidence problems*: Given a finite set of points and a finite set of curves in the plane or higher dimensional space, count the maximum number of *incidences*: times when one of the points lies on one of the curves. Most frequently, these curves are semialgebraic sets, given by a bounded number of polynomial inequalities of bounded degree - from a model-theoretic perspective, they are defined by the same formula. These incidence problems have been used to solve major problems such as Erdős's distinct distances problem.[GK15] They are best approached with variations on the polynomial method, where space is carefully partitioned into regions containing only a small subset of the points and curves. The incidences are counted region-by-region, and these contributions are summed together.

Model theory can be used throughout this process - for instance, the VC-dimension of the family of curves, which is finite by NIP, can be used to count incidences in each region of the partition. However, much of the combinatorial structure of semialgebraic sets appearing in incidence combinatorics comes from stronger model-theoretic properties connected to partitioning. One partitioning strategy starts with this fact about $\mathbb{R}$: given finitely many semialgebraic curves from a uniformly definable family, we can partition space with uniformly definable pieces so that the regions are not crossed by any of our curves. This partition is an example of a *distal cell decomposition* for the semialgebraic sets. We call a structure *distal* when all its definable sets have these cell decompositions. These can be used in a general polynomial method-like strategy, using a *distal cutting lemma*[CGS20] to prove combinatorial properties such as incidence bounds, NIP, and the *strong Erdős-Hajnal property.*[CS18]

The strong Erdős-Hajnal property can be viewed as a particularly strong version of Ramsey's theorem for definable relations, or as the first step towards a distal version of Szemerédi's regularity lemma. In its full generality, the regularity lemma partitions the vertices of an arbitrary finite graph into a *regularity partition.* The graph behaves like an easily understood pseudorandom bipartite graph between nearly every pair of pieces of this partition. In the case when the graph is definable in a tame structure, more can be said. Graphs definable in NIP structures require far fewer pieces to the partition, and the behavior of the graph on the "good" pieces is even more constrained.[CS21] If the structure is stable, no exceptional pairs of partition pieces are needed.[MS14] Distality, on the other hand, strengthens the NIP regularity lemma by guaranteeing that the graph is a complete bipartite graph, or has no edges at all, between most pairs of partition pieces. These regularity lemmas are versatile tools for understanding definable graphs, and provide one of the clearest ways to compare model-theoretic properties.

The combinatorial results above apply well to finite definable graphs, but these regularity lemmas and the distal cutting lemma apply in a wider context. Rather than counting elements of finite sets, we can measure arbitrary definable sets with respect to *Keisler measures*. These real-valued generalizations of types are natural to use in model-theoretic arguments. For instance, they are closed under ultraproducts, making them the logical generalization of counting measures to pseudofinite contexts. As with types, there are many different kinds of Keisler measures - to take full advantage of NIP or distality, we will focus on a particularly well-behaved class, the *generically stable* measures. These turn out to be the correct model-theoretic setting for many combinatorial results, and are of purely model-theoretic interest, as they can be used to characterize NIP and distal structures in their own right.

Keisler measures are not the only useful real-valued generalization of model-theoretic objects. The results of neostability can be applied to even more examples by extending the theory to *continuous logic*. This replaces the structures of classical first-order logic with *metric structures*: complete bounded metric spaces imbued once more with logical symbols in a specific language, but where formulas now evaluate to real number values in a continuous way. The monograph which first presented a full introduction to continuous logic, [BBH08], provides examples of stable metric structures, such as infinite-dimensional Hilbert spaces and atomless probability algebras. However, despite work extending the definition of NIP to continuous logic in [Ben09], there are still few examples of metric structures that are NIP and unstable, other than discrete structures and their randomizations. One goal of this dissertation is to change this, by introducing distality to continuous logic and providing some examples of distal metric structures.

Continuous logic also provides an even more natural setting for studying regularity lemmas. While the stable, NIP, and distal regularity lemmas were all originally phrased for graphs, they can be extended to *analytic regularity lemmas*, replacing the graphs with real-valued functions.[LS07] These are broken into a sum of a structured part, a pseudorandom part, and an error part. Now if we assume the real-valued function is a definable predicate in some metric structure, we are again able to deduce stronger regularity lemmas when the structure is NIP[LS10], n-dependent[CT20], or stable[CCP24]. In this dissertation, we will revisit the NIP analytic regularity lemma and prove a distal analytic regularity lemma.

1.1 Outline

In Chapter 2, we explore explicit combinatorial consequences of distality. In any distal structure, definable relations admit distal cell decompositions, which set up an array of combinatorial tools, including the distal cutting lemma.[CGS20] This cutting lemma can be used to bound the number of edges in a definable bipartite graph omitting a specific complete bipartite subgraph. This special case of Zaranciewicz's problem has been used, in the context of semialgebraic sets, to solve many combinatorial problems involving incidences (see for instance [She22]). In order to get practical combinatorial bounds, the number of cells in the distal cell decompositions must be estimated. In Chapter 2, we establish such bounds in a variety of distal structures. This begins with a method for extrapolating from one-dimensional distal cell decompositions to decompositions for higher dimensions. Then bounds are found for one-dimensional distal cell decompositions in (weakly) o-Minimal structures as well as both a strong vector space structure and the standard valued field structure on $\mathbb{Q}_p$.

Chapter 2 is a reprinting, with slight modifications, of [And23a] (Copyright © 2023, Cambridge University Press, reprinted with permission).

In the subsequent chapters, we study distality in the context of continuous logic. Chapter 3 starts by expanding the the study of NIP metric structures started in [Ben09]. In discrete logic, NIP is defined by all definable relations having finite VC dimension. There are several generalizations of VC dimension to real-valued functions, and we check that all of these give the same notion of NIP for metric structures. We then generalize other combinatorial properties of classes with finite VC dimension, such as ε-nets and the (p, q)-theorem to the real-valued case. These combinatorial facts let us prove the existence of uniform honest definitions for NIP metric structures.

This framework for studying NIP metric structures lets us approach distality. Distal structures can be defined using properties of indiscernible sequences or invariant types [Sim13], or the existence of strong honest definitions [CS15] or distal cell decompositions [CGS20]. We generalize these definitions of distality to continuous logic, and prove them equivalent, using the real-valued (p, q)-theorem again to find uniform strong honest definitions.

In Chapter 4, we extend more characterizations of distality to the metric structure setting: smoothness of generically stable measures, the definable strong Erdős-Hajnal property, and the distal regularity lemma. These are all stated in terms of Keisler measures, so we first establish more background on their behavior in NIP metric structures, including proving that the equivalence of the many definitions of generically stable measures holds in this setting.

Chapter 5, joint work with Itaï Ben Yaacov, finds some examples of distal metric structures. One source of examples are metric valued fields [Ben14]. By analyzing indiscernible sequences, we are able to show that real closed metric valued fields are distal, from which we conclude that algebraically closed metric valued fields, while stable, have the strong Erdős-Hajnal property. We find another example in topological dynamics. Specifically, we study a metric structure whose automorphism group is the well-understood Polish group $\mathrm{Hom}^+([0, 1])$ of increasing homeomorphisms of $[0, 1]$. It was shown in [Iba16] that any such structure would be NIP and highly unstable, and further properties of this structure were established in [Ben18]. In this chapter, we shed further light on this structure, including characterizing the models of its theory, which we call *Dual Linear Continua*, up to isomorphism. We are able to characterize their indiscernible sequences and prove that they are distal, as well as constructing explicit distal cell decompositions.

1.2 Preliminaries

Each chapter has been written to be mostly self-contained, but we will introduce some general notation and background in the rest of this introduction. The primary goal is to present a summary of the study of distal structures, which will be examined quantitatively in Chapter 2, and generalized to continuous logic in Chapters 3, 4, and 5.

1.2.1 Continuous Logic

Our continuous logic constructions follow the conventions of the monograph [BBH08]. We will give an overview of some of the basic definitions and notation we use, but the reader unfamiliar with continuous logic should consult [BBH08] for the full definitions. The first chapters of [Han20a] provide an alternate exposition, which we also recommend, although some terminology is different. For more background on the interaction between continuous logic and NIP, see Chapter 3. For more background on Keisler measures in continuous logic, see Chapter 4.

We start by reviewing the most basic definitions of continuous logic:

Definition 1.2.1 (Metric Structures)**.** A *metric language* is defined just as a language in classical logic, but each (function or relation) symbol is assigned a positive real Lipschitz constant.

A *metric structure* in a given language consists of:

- A complete bounded metric space M
- For each n-ary k-Lipschitz function symbol, a k-Lipschitz function $M^n \to M$

- For each n-ary k-Lipschitz relation symbol, a k-Lipschitz function $M^n \to [0, 1]$.

In discrete first-order logic, formulas are defined purely recursively. In many formalizations of continuous logic, it is easier to recursively construct a dense set of "formulas", and then close this set under uniform limits. There are different notations for these two classes of formulas, the dense set and its closure, but we will call the dense set, constructed recursively, *formulas*, while their closure will be the space of *definable predicates*. Because the interpretations of definable predicates can depend on infinitely many variables, we will frequently deal with variable tuples of countably infinite length. As if x, y are countably infinite tuples, $|x|$ equals $|x, y|$, we shall just refer to the relevant cartesian products of a set M as M^x and $M^x \times M^y$, rather than $M^{|x|}$ or $M^{|x,y|}$.

We use this recursive definition for formulas:

Definition 1.2.2 (Formulas)**.** A *term* is constructed by applying formula symbols to variables recursively, as in discrete logic.

An *atomic* formula is constructed by either applying an n-ary relation symbol to n terms, or applying the metric symbol $d(x, y)$ to two terms.

A *formula* on some tuple x of variables is something recursively constructed using the following operations:

- An atomic formula is a formula
- A combination $u(\phi_1, \ldots, \phi_n)$ of formulas is a formula, where $u : [0, 1]^n \to [0, 1]$ is a continuous function
- If $\phi(x, y)$ is a formula and y is a single variable, then $\sup_y \phi(x, y)$ and $\inf_y \phi(x, y)$ are formulas on x.

These symbols can be naturally interpreted in a given metric structure M - the Lipschitz continuity assumptions that we have made are one way of guaranteeing that these interpretations will be continuous. When we speak of continuity, we will understand M as having the metric topology, and M^x as having the product topology. If x is finite, we will also think of M^x as a metric space using the sup metric.

Fact 1.2.3 (Interpretations of Terms and Formulas are Continuous)**.** *Terms and formulas can be naturally interpreted in a given metric structure M. The interpretation of a term $t(x)$ is a continuous function $M^x \to M$, while the interpretation of a formula $\phi(x)$ is a continuous function $M^x \to [0,1]$.*

Also, the value of a given term or formula depends only on finitely many of the variables in the tuple x.

Given a metric structure M, some tuple $a \in M^x$, and some set $B \subseteq M$, the type $\mathrm{tp}(a/B)$ records the value $\phi(a,b)$ for every formula $\phi(x;y)$ and tuple $b \in B^y$. We refer to [BBH08] for the construction of the type space $S_x(B)$ in which such types belong, but we recall some of its properties:

Fact 1.2.4 (Type Spaces)**.** *For any metric structure M and set $B \subseteq M$, the type space $S_x(B)$ is a compact Hausdorff space.*

Each type in $S_x(B)$ can be realized in an elementary extension of M.

For each formula $\phi(x;y)$ and $b \in B^y$, the function $\phi(x;b)$ sending a type p to the value $\phi(a;b)$ where a realizes p in an elementary extension of M is well-defined and continuous. In fact, the topology on $S_x(B)$ is the coarsest such that all of these functions are continuous.

In discrete logic, each clopen set of a type space is defined by some formula. Characteristic functions of clopen sets are exactly continuous functions to the discrete space $\{0,1\}$, so we find that formulas with parameters in B (up to logical equivalence) are in correspondence with continuous functions $S_x(B) \to \{0,1\}$. Our definition of formulas will not precisely correspond to continuous functions $S_x(B) \to [0,1]$ - these are the larger, and arguably more important class of *definable predicates*:

Definition 1.2.5 (Definable Predicates)**.** If M is a metric structure and $B \subseteq M$, a *definable predicate* with parameters in B is a continuous function $S_x(B) \to [0,1]$.

These are precisely the uniform limits of formulas, when treated as functions $S_x(B) \to [0,1]$.

In classical model theory, we frequently use the notation $\phi(M;b)$ to indicate the subset of M^x defined by the formula $\phi(x;y)$ using the parameter $b \in M^y$. For a definable predicate $\phi(x;y)$ in continuous logic, we define $\phi(M;b)$ as the subset of M^x on which $\phi(x;b) = 0$. For other $r \in [0,1]$, we will use the notations $\phi_{\leq r}(M;b)$ and $\phi_{\geq r}(M;b)$ to denote the sets where $\phi(x;b) \leq r$ and $\phi(x;b) \geq r$. Given any condition (an inequality or equality of definable predicates), we will use notation such as $[\phi(x) \geq r]$ to denote the subset of a type space $S_x(B)$ where that condition is true.

1.2.2 Monster Models

Throughout this dissertation, we will usually work in the context of a monster model of any given theory, possibly in continuous logic. For any language $\mathcal{L}$ and $\mathcal{L}$-theory T, we let κ be a cardinal larger than $|\mathcal{L}|$, and let $\mathcal{U}$ be a κ-saturated and κ-strongly homogeneous model of T. Constructing an appropriate κ and the model $\mathcal{U}$ requires some set-theoretic technology which we will not recount here - see for instance [HK23] for details. When we refer to a *small set* $A \subseteq \mathcal{U}$ or a *small model* $M \preceq \mathcal{U}$, we are assuming that $|A|, |M| < \kappa$, or that $\mathcal{U}$ is $|A|^+, |M|^+$-saturated. In the context of continuous logic, the relevant measure of the "size" of a subset of $\mathcal{U}$ is its density character with respect to the metric on $\mathcal{U}$, so we define smallness as having a dense subset of size $< \kappa$.

1.2.3 NIP

Most of this thesis will take place in the combinatorially tame context of NIP structures. To approach these, we start by defining VC-dimension, and several associated definitions for arbitrary families of subsets of a given set. Some of these definitions will be revisited and extended to describe distal structures in Chapter 2, and Chapter 3 will examine fuzzy and real-valued generalizations of these concepts.

Definition 1.2.6 (Shatter functions and VC-dimension)**.** Let $\mathcal{F}$ be a family of subsets of a set X.

- For any finite $A \subseteq X$, let $\mathcal{F} \cap A = \{S \cap A : S \in \mathcal{F}\}$,
- For finite $A \subseteq X$, let $\pi_{\mathcal{F}}(A) = |\mathcal{F} \cap A|$
- For $n \in \mathbb{N}$, let $\pi_{\mathcal{F}}(n) = \max_{A \subseteq X, |A|=n} \pi_{\mathcal{F}}(A)$
- Define the *VC-dimension* of $\mathcal{F}$ to be the largest $n \in \mathbb{N}$ such that $\pi_{\mathcal{F}}(n) = 2^n$, or ∞ if this is true for all n.

The function $\pi_{\mathcal{F}}(n)$, called the *shatter function* of $\mathcal{F}$, can be understood through the dichotomy of the Sauer-Shelah lemma:

Fact 1.2.7 (Sauer-Shelah Lemma [Sau72, She72, CV71])**.** *Let $\mathcal{F}$ be a family of subsets of a set X. One of the following holds:*

- *For all n, $\pi_{\mathcal{F}}(n) = 2^n$ (that is $\mathcal{F}$ has infinite VC-dimension)*
- *For all n, $\pi_{\mathcal{F}}(n) \leq \sum_{k=0}^{d} \binom{n}{k} = O(n^d)$, where d is the VC-dimension of $\mathcal{F}$.*

This polynomial growth inspired another definition, named VC-density in [ADH16].

Definition 1.2.8 (VC-density)**.** Let $\mathcal{F}$ be a family of subsets of a set X. The *VC-density* of $\mathcal{F}$ is

$$\limsup_{n\to\infty} \frac{\log \pi_{\mathcal{F}}(n)}{\log n}.$$

The VC-density is connected to the VC-dimension by the following corollary of Sauer-Shelah:

Corollary 1.2.9. *Let $\mathcal{F}$ be a family of subsets of a set X. One of the following holds:*

- *The VC-dimension and VC-density of $\mathcal{F}$ are both infinite*
- *The VC-dimension and VC-density of $\mathcal{F}$ are both finite, and the VC-density is bounded by the VC-dimension.*

We give a name to the finite-dimension side of this dichotomy:

Definition 1.2.10 (VC-classes)**.** Let $\mathcal{F}$ be a family of subsets of a set X. We call $\mathcal{F}$ a *VC-class*, or *dependent*, when the VC-dimension (or equivalently, the VC-density) is finite.

These VC-classes notably satisfy a uniform law of large numbers. To state them more easily, we consider ε-approximations, which are finite tuples whose counting measures approximate a given measure with respect to a specific class of sets.

Definition 1.2.11 (ε-approximations)**.** Let $\mathcal{F}$ be a family of subsets of a set X, let $\varepsilon > 0$, and let μ be a probability measure on X with respect to which every set of $\mathcal{F}$ is measurable.

For any tuple $(a_1, \ldots, a_n) \in X^n$ and $S \in \mathcal{F}$, let

$$\mathrm{Av}(a_1, \ldots, a_n; S) = \frac{1}{n} \left| \{i : a_i \in S\} \right|.$$

Then a tuple $(a_1, \ldots, a_n) \in X^n$ is an *ε-approximation* to μ when for all $S \in \mathcal{F}$,

$$|\mathrm{Av}(a_1, \ldots, a_n; S) - \mu(S)| \leq \varepsilon.$$

This uniform law of large numbers is usually called the VC-Theorem after its discoverers, the authors of [CV71].

Fact 1.2.12 (The VC-Theorem [CV71])**.** *Let (X, μ) be a finite probability space, and $\mathcal{F}$ a VC-class of subsets of X. Then for $\varepsilon > 0$, let A_n^ε be the set of ε-approximations to $\mathcal{F}$ with respect to μ. Then*

$$\lim_{n \to \infty} \mu^n(A_n^\varepsilon) = 1.$$

Logical formulas naturally induce families of sets, which may or may not be VC-classes. When they are, we use the term *NIP*, for "not the independence property."

Definition 1.2.13 (NIP)**.** In the context of a theory T, we say that a formula $\phi(x;y)$ is *NIP* when the family

$$\{\{a : M \vDash \phi(a;b)\} : b \in M^y\}$$

is a VC-class in every model $M \vDash T$.

Let $\pi_\phi(n)$ denote the shatter function for this class induced by ϕ - this implicitly depends on the choice of model M.

When every formula $\phi(x;y)$ is NIP, then we call the theory T NIP also.

Chapter 2 deals mostly with the dual case.

Definition 1.2.14 (Dual Shatter Function)**.** If $\phi(x;y)$ is a formula, let $\phi^*(y;x)$ be the *dual* formula, the same formula with the order of the variables reversed.

Also define $\pi^*_\phi(n)$ to be the shatter function for ϕ^*.

The following equivalent form is given as the definition of π^*_ϕ in Chapter 2:

Fact 1.2.15 (Dual Shatter Function Through Types)**.** *Given a formula $\phi(x;y)$ and a structure M, then*

$$\pi^*_\phi(n) = \max_{B \subseteq M^{|y|}, |B|=n} |S^\phi(B)|,$$

where $S^\phi(B)$ is the space of all complete ϕ-types over B.

For defining VC-classes and NIP, it does not matter whether we concern ourselves with a formula or its dual:

Fact 1.2.16 (Folklore, see [ADH16, Lemma 2.5].)**.** *The formula $\phi(x;y)$ is NIP if and only if its dual is.*

We can also define NIP structures in terms of indiscernible sequences:

Fact 1.2.17. *A theory T is NIP if and only for all models $M \vDash T$, indiscernible sequences $(a_n : n \in \mathbb{N})$ indexed by $\mathbb{N}$ in $M^{|x|}$, and all $\phi(x; b)$ with $b \in M^{|y|}$, the sequence $\phi(a_n; b)$ of truth values is eventually true or eventually false.*

This kind of definition in terms of indiscernible sequence is one of the easiest to translate to continuous logic, as seen in [Ben09], which we will use in Chapter 3.

For some examples, NIP structures include *stable* structures such as

- sets in the empty language
- vector spaces over a fixed field
- algebraically closed fields

and others, such as algebraically closed valued fields, and distal structures, which we will examine shortly in Subsection 1.2.5.

One relevant property of NIP formulas, which can be strengthened in the case of distal structures, is the presence of *honest definitions*.

Definition 1.2.18 (Honest Definitions)**.** Let $A \subseteq M^x$, and let $(M, A) \preceq (M', A')$ be an elementary extension of the structure (M, A) with a relation symbol for A. Let $\phi(x; b)$ be an M-formula, and let $\psi(x; d)$ be an A'-formula. We say that $\psi(x; d)$ is an *honest definition* for $\phi(x; b)$ over A when

- for $a \in A$, $M \vDash \psi(a; d) \iff M \vDash \phi(a; b)$
- for $a \in A'$, $M \vDash \psi(a; d) \implies M \vDash \phi(a; b)$.

If the same formula $\psi(y; z)$ works for any choice of M, A, b with $|A| \geq 2$, then we call $\psi(x; z)$ an honest definition for $\phi(x; y)$.

We will only need the version in [CS13, Theorem 11], which assumes the whole theory is NIP, but [BKS24, Corollary 5.23] constructs honest definitions for individual NIP formulas.

Fact 1.2.19 (Honest Definitions [CS13, Theorem 11]). *A theory T is NIP if and only if every formula $\phi(x; y)$ admits an honest definition $\psi(x; z)$.*

Honest definitions were originally used in a proof that the Shelah expansion of an NIP theory is NIP.

Fact 1.2.20 (The Shelah Expansion [CS13, Corollary 1.10]). *If M is an $\mathcal{L}$-structure, let M^{Sh} be the expansion of M by relations for all* externally definable sets*: sets of the form $\phi(M; b)$ where b lies in an elementary extension of M.*

If M is NIP, then M^{Sh} admits quantifier elimination in this language and is NIP.

Honest definitions can also be stated in a more finitary form, which gives the following useful corollary:

Fact 1.2.21 (UDTFS [CS15, Theorem 15]). *A theory T has NIP if and only if every formula has UDTFS (Uniform Definability of Types over Finite Sets):*

We say that $\phi(x; y)$ has UDTFS if there is $\theta(x; z)$ such that for every finite A and a there is $b \in A$ such that $\phi(A, a) = \theta(A, b)$.

1.2.4 Keisler Measures

Continuous logic is not the only way to generalize model theory to take real truth values. Types can also be generalized to *Keisler measures*. They can be defined as finitely-additive probability measures on boolean algebras of definable sets, but we will use this definition (see [Sim15, Section 7.1] for a discussion of the equivalence).

Definition 1.2.22 (Keisler Measures). A *Keisler measure* μ over a set $A \subseteq \mathcal{U}$ is a regular Borel probability measure on the type space $S_x(A)$. We call the space of such measures $\mathfrak{M}_x(A)$.

For a formula $\phi(x)$ with parameters in A, we will use the simple notation $\mu(\phi(x))$ for the measure of the subset of the type space $\phi(x)$ defines. We will use one consequence of the definition in terms of finitely-additive probability measures:

Fact 1.2.23. *A Keisler measure $\mu \in \mathfrak{M}_x(A)$ is uniquely determined by the measures $\mu(\phi(x))$ of formulas $\phi(x)$ with parameters in A.*

We say that these generalize types, because for any type $p \in S_x(A)$, there is a corresponding Dirac measure $\delta_p \in \mathfrak{M}_x(A)$. Much like types, these can be classified many ways, and we will frequently have to restrict our attention to particularly well-behaved classes of measures.

Definition 1.2.24 (Kinds of Keisler Measures)**.** Let μ be a global Keisler measure, and let $A \subseteq \mathcal{U}$ be a small set, and $M \preceq \mathcal{U}$ a small model.

- We say μ is A-*invariant* when for any tuples $a \equiv_A b$ in $\mathcal{U}^y$, and any formula $\phi(x; y) \in \mathcal{L}(A)$, $\mu(\phi(x; a)) = \mu(\phi(x; b))$. Equivalently, any automorphism of $\mathcal{U}$ fixing A preserves μ.
- If μ is A-invariant, define the map $F^{\phi}_{\mu,A} : S_y(A) \to [0, 1]$ by $F^{\phi}_{\mu,A}(p) = \mu(\phi(x; b))$ for $b \models p$.
- We say μ is A-*Borel definable* when it is A-*invariant* and for all $\phi(x; y) \in \mathcal{L}(A)$, the map $F^{\phi}_{\mu,A}$ is Borel.
- We say μ is A-*definable* when it is A-*invariant* and for all $\phi(x; y) \in \mathcal{L}(A)$, the map $F^{\phi}_{\mu,A}$ is continuous (and thus a definable predicate).
- We say μ is *finitely satisfiable* A when μ is in the topological closure of the convex hull of the Dirac measures at types of points in A.
- We call a definable, finitely satisfiable measure *dfs*.

- We say μ is *finitely approximated in* M when for every $\varphi(x;y) \in \mathcal{L}(M)$ and every $\varepsilon > 0$, there exists a tuple $(a_1, \ldots, a_n) \in (M^x)^n$ which is a ε-approximation for the family $\{\varphi(x;b) : b \in \mathcal{U}^y\}$ with respect to μ. We abbreviate this property as *fam*.

- We say μ is a *frequency interpretation measure over* M when for every $\varphi(x;y) \in \mathcal{L}(M)$, there is a family of formulas $(\theta_n(x_1, \ldots, x_n) : n \in \omega)$ with parameters in M such that $\lim_{n\to\infty} \mu^{(n)}(\theta_n(x_1, \ldots, x_n)) = 1$, and for every $\varepsilon > 0$, for large enough n, any $\bar{a} \in (\mathcal{U}^x)^n$ satisfying $\theta_n(\bar{a})$ is a ε-approximation to $\varphi(x;y)$ with respect to μ. We abbreviate this property as *fim*.

- We say μ is *smooth* over M when for every N with $M \preceq N$, there exists a unique extension $\mu' \in \mathfrak{M}_x(N)$ of $\mu|_M$.

We now define product measures.

Definition 1.2.25 (Product Measures)**.** Let $\mu \in \mathfrak{M}_x(A), \nu \in \mathfrak{M}_y(A)$ be measures. Then a measure $\omega \in \mathfrak{M}_{xy}(A)$ is a *product measure* of μ and ν when for every formula $\phi(x)\psi(y)$ with parameters in A, we have

$$\omega(\phi(x)\psi(y)) = \mu(\phi(x))\nu(\psi(y)).$$

Note that not every measure on xy extending μ and ν is a product measure, a significant subtlety that doesn't arise when studying types. Also, this product is not usually uniquely determined. When it is, this is an important property of the measures:

Definition 1.2.26 ((Weak) Orthogonality)**.** Let $\mu \in \mathfrak{M}_x(A), \nu \in \mathfrak{M}_y(A)$ be measures. Then μ and ν are *weakly orthogonal* when they have a unique product measure $\omega \in \mathfrak{M}_{xy}(A)$. In the case of global measures (measures over $\mathcal{U}$), we simply call them *orthogonal*.

For a unique way of defining product measures without assuming any orthogonality, we turn to the Morley product:

Definition 1.2.27 (The Morley Product)**.** Given an A-Borel definable measure μ and a global measure ν, let $\mu \otimes \nu$ be the measure such that

$$\mu \otimes \nu(\phi(x; y)) = \int_{S_y(A')} F^{\phi}_{\mu,A'}(y)\, d\nu|_{A'}$$

where $\phi(x; y)$ is a formula, and A' contains A and the parameters of ϕ.

In the NIP context, many important definitions coincide, and we call measures satisfying those properties *generically stable*. These turn out to be the natural generalization of counting measures in many combinatorial statements, as they satisfy a definable version of the VC-theorem (fim). They also are simple in basic model-theoretic terms (dfs), and their Morley products act much like probability measures of independent identically distributed variables.

Fact 1.2.28 (Generically Stable Measures [HPS13, Theorem 3.2])**.** *Assume T is NIP. For any small model $M \subseteq \mathcal{U}$, if μ is a global M-invariant measure, the following are equivalent:*

1. *μ is fim over M*
2. *μ is fam over M*
3. *μ is dfs over M*
4. *$\mu(x) \otimes \mu(y) = \mu(y) \otimes \mu(x)$*
5. *$\mu^{(\omega)}(x_0, x_1, \dots)|_M$ is totally indiscernible.*

Perhaps the most important tool in model theory is the ability to realize types. The closest analog we have for measures to a realized type is a smooth measure, as realized types are exactly the types that admit unique extensions. In the NIP context, we can find smooth extensions, allowing us to generalize many arguments involving realizing types:

Fact 1.2.29 (Smooth Extensions [Kei87, Theorem 3.26])**.** *Each Keisler measure $\mu \in \mathfrak{M}_x(M)$ over a small model M admits a smooth extension over some $M \preceq N$.*

Smooth measures are weakly orthogonal to all other measures, and in fact, this characterizes smoothness:

Fact 1.2.30 (Orthogonality and Smoothness [Sim16, Lemma 1.6])**.** *Let $\mu \in \mathfrak{M}_x(\mathcal{U})$ be a global measure, and let $M \subset \mathcal{U}$ be small. The following are equivalent:*

- *μ is smooth over M*
- *$\mu|_M$ is weakly orthogonal to all types $p(y) \in S_y(M)$*
- *$\mu|_M$ is weakly orthogonal to all measures $\nu(y) \in \mathfrak{M}_y(M)$.*

1.2.5 Distality

The main subject of this dissertation is distality, a stronger property than NIP.

Definition 1.2.31 (Distality)**.** A theory T is *distal* when for every dense linear order I, indiscernible sequence $(a_i : i \in I)$ in a model $M \vDash T$, and $B \subseteq M$, then if $(a_i : i \neq i_0)$ is indiscernible over B, so is $(a_i : i \in I)$.

Distality also strengthens NIP by being equivalent to a stronger version of honest definitions. These can be posed in two essentially equivalent ways. In the first, a strong honest definition for $\phi(x; y)$ is another formula which gives control over ϕ-types:

Fact 1.2.32 (Strong Honest Definitions [CS15, Theorem 21])**.** *An NIP theory T is distal if and only if every formula admits a strong honest definition:*

Let $A \subseteq M^x$, and let $(M, A) \preceq (M', A')$ be an elementary extension of the structure (M, A) with a relation symbol for A. Let $\phi(x; y)$ be an M-formula, and let $\theta(x; d)$ be an A'-formula. We say that $\theta(x; d)$ is a strong honest definition *for $\phi(a; y)$ over A when*

- *$M' \vDash \theta(a; d)$*
- *$\theta(x; d) \vdash \mathrm{tp}_\phi(a/A)$.*

For either of these definitions, if the same formula $\theta(x; z)$ works for any choice of M, A, a, then we call $\theta(x; z)$ a strong honest definition for $\phi(x; y)$.

and the other more obviously strengthens honest definitions:

Definition 1.2.33 (Strong* Honest Definitions)**.** Let $A \subseteq M^x$, and let $(M, A) \preceq (M', A')$ be an elementary extension of the structure (M, A) with a relation symbol for A. Let $\phi(x; b)$ be an M-formula, and let $\psi(x; d)$ be an A'-formula. We say that $\psi(x; d)$ is a *strong* honest definition* for $\phi(x; b)$ over A when

- for $a \in A$, $M \vDash \psi(a; d) \iff M \vDash \phi(a; b)$
- for $a \in M'^x$, $M \vDash \psi(a; d) \implies M \vDash \phi(a; b)$.

If the same formula $\psi(y; z)$ works for any choice of M, A, b with $|A| \geq 2$, then we call $\psi(x; z)$ a strong* honest definition for $\phi(x; y)$.

In this dissertation, we refer to this second form as *strong** honest definitions, chiefly because we needed a name to differentiate these from the other form. This name was chosen because strong* honest definitions are dual to strong honest definitions in that they refer to different variable tuples, and they are also slightly less strong, as encapsulated in the equivalence of existence of strong(*) honest definitions for individual formulas:

Fact 1.2.34 (Strong vs. Strong* Honest Definitions [Sim15, See Subsection 9.3.1])**.** *Let $\phi(x; y)$ be a formula. If $\phi(x; y)$ admits a strong honest definition, then both $\phi^*(y; x)$ and its negation admit strong* honest definitions.*

Conversely, if both $\phi(x; y)$ and its negation admit strong honest definitions, then the dual formula $\phi^*(y; x)$ admits a strong honest definition.*

Strong honest definitions are the bridge between the model-theoretic property of distality and the combinatorial properties of a structure. To better take advantage of combinatorial intuition, they were equivalently reformulated as *distal cell decompositions* in [CGS20]. Distal cell decompositions, along with other kinds of combinatorial cell decompositions, are defined and explored in detail in Chapter 2. We will state one important combinatorial consequence here, the *distal cutting lemma*. While distal cell decompositions essentially cover a space M^x with pieces on which ϕ-types over a particular finite set do not vary at all, a *cutting* is a cover with pieces on which these types are allowed to vary a limited amount:

Definition 1.2.35. Let $\mathcal{F}$ be a finite family of subsets of a set X with $|\mathcal{F}| = n$. Given a real $1 < r < n$, we say that a family $\mathcal{C}$ of subsets of X is a $\frac{1}{r}$*-cutting for* $\mathcal{F}$ when $\mathcal{C}$ forms a cover of X and each set $C \in \mathcal{C}$ is crossed by at most $\frac{n}{r}$ elements of $\mathcal{F}$.

Fact 1.2.36 (Distal Cutting Lemma [CGS20, Theorem 3.2])**.** *Let $\phi(x; y)$ be a formula admitting a distal cell decomposition of exponent d. Then for any natural n and any real $1 < r < n$, there exists $t = \mathcal{O}(r^d)$ such that for any finite $H \subseteq M^{|y|}$ of size n, there are uniformly definable sets $X_1, \dots, X_t \subseteq M^{|x|}$ which form an $\frac{1}{r}$-cutting for $\{\phi(x; h) : h \in H\}$.*

Distality can also be defined through properties of types. One way is with *compressible types*, which are essentially types where a limited version of strong honest definitions apply:

Fact 1.2.37 (Compressible Types [Sim19, Definition 3.1])**.** *An NIP theory T is distal if and only if every type is* compressible*:*

A type $p(x) = \mathrm{tp}(a/A)$ is compressible *if given an $|A|^+$-saturated elementary extension $(A, a) \preceq (A', a)$, for any formula $\phi(x; y)$, there is some $\zeta(x; e) \in \mathrm{tp}(a/A')$ such that $\zeta(x; e) \vdash \mathrm{tp}_\phi(a/A)$.*

These types are also used to prove facts about general NIP theories, such as explicit constructions of honest definitions in [BKS24].

Meanwhile, [Sim15] introduces distality in terms of a different property of types, more related to the indiscernible sequence definition of distality:

Fact 1.2.38 (Distal Types [Sim15, Def. 9.3])**.** *An NIP theory T is distal if and only if every global A-invariant type is* distal*:*

Let p be a global A-invariant type. Then p is distal *over A when for any tuple b, if $I \models p^{(\omega)}|_{Ab}$, then $p|_{AI}$ and $\mathrm{tp}(b/AI)$ are weakly orthogonal. (That means that there is a unique complete type over A extending $p(x) \cup q(y)$.)*

If p is distal over all A such that p is invariant over A, then we just say that p is distal, without specifying A.

Distality can also be defined through properties of measures, as generically stable measures in a distal theory are as well-behaved as possible: they are smooth.

Fact 1.2.39 (Distality through Measures [Sim13, Theorem 1.1])**.** *An NIP theory T is distal if and only if all generically stable global measures are smooth.*

Examples Distal structures include o-minimal structures and their generalizations, such as weakly o-minimal and quasi-o-minimal structures. In addition to these explicitly ordered structures, there are also structures dominated by other ordered behavior, such as valuations. Specifically, P-minimal structures such as the valued field $\mathbb{Q}_p$ and a linear reduct of $\mathbb{Q}_p$, which we will examine in more detail in Chapter 2.

1.2.6 Regularity

Szemerédi's regularity lemma is a powerful tool for understanding the structure of arbitrary finite graphs. It partitions the vertices of a graph into pieces such that between most pairs of pieces, the graph acts like a random bipartite graph.

Fact 1.2.40 (Szemerédi's Regularity Lemma [Sze75])**.** *For every $\varepsilon > 0$, there is some $N \in \mathbb{N}$ such that every finite graph (V, E) admits a ε-*regularity partition *of size N: Specifically, V can be partitioned into disjoint sets $V = V_1 \cup \cdots \cup V_N$, and there are reals δ_{ij} for $i, j \leq N$,* *and some set* $\Sigma \subseteq \{1, \ldots, N\}^2$ *of* exceptional pairs*, such that*

$$\sum_{(i,j)\in\Sigma} |V_i||V_j| \leq \varepsilon |V|^2$$

and for each $(i,j) \notin \Sigma$*, and* $A \subseteq V_i, B \subseteq V_j$*, if* $E(A,B)$ *is the number of edges between sets* A *and* B*, then*

$$|E(A,B) - \delta_{ij}|A||B|| \leq \varepsilon |V_i||V_j|.$$

Unfortunately, in the case of an arbitrary graph, the number M can be enormous. If, however, the graph happens to be definable in a tame structure, more can be said. In Chapter 4, we will examine continuous logic versions of the NIP and distal regularity lemmas, which we will first review in the discrete setting here. They can be stated in terms of counting measures, or rather, the sizes of vertex sets, but they are more naturally proven in the more general context of generically stable Keisler measures.

Fact 1.2.41 (NIP Regularity Lemma [CS21,])**.** *Let* M *be an NIP structure, and let* $\phi(x;y)$ *be a formula. Then there is some* $c > 0$ *such that for every* $\varepsilon > 0$ *and all generically stable measures* $\mu \in \mathfrak{M}_x(M), \nu \in \mathfrak{M}_y(M)$*, there are partitions* $M^x = A_1 \cup \cdots \cup A_N$ *and* $M^y = B_1 \cup \cdots \cup B_N$ *and a set* $\Sigma \subseteq \{1, \ldots, N\}^2$ *such that, if* $\omega = \mu \otimes \nu$*,*

- $K \leq \varepsilon^{-c}$
- *Each* A_i *and* B_j *is defined by a boolean combination of instances of* ϕ *of complexity depending only on* ϕ, ε
- $\sum_{(i,j)\in\Sigma} \omega(A_i \times B_j) \leq \varepsilon$
- *For all* $(i,j) \notin \Sigma$*, there is* $\delta_{i,j} \in \{0,1\}$ *such that*

$$|\omega(\phi(x;y) \cap A_i \times B_j) - \delta_{i,j}\omega(A_i \times B_j)| \leq \varepsilon\omega(A_i \times B_j).$$

The characterization of distality in terms of generically stable measures being smooth was used in [Sim16] to prove a regularity lemma for distal structures. Relative to the NIP regularity lemma, this strengthens the control over the behavior of ϕ on the non-exceptional pairs of the regularity partition to the following strong property:

Definition 1.2.42. If $\phi(x; y)$ is a formula, then subsets $A \subseteq M^x, B \subseteq M^y$ are *ϕ-homogeneous* when either for all $(a, b) \in A \times B$, $M \vDash \phi(a, b)$, or for all $(a, b) \in A \times B$, $M \vDash \neg\phi(a, b)$.

Fact 1.2.43 (Distal Regularity Lemma [CS18,], with an alternate proof in [Sim16])**.** *Let M be a distal structure, and let $\phi(x; y)$ be a formula. Then there is some $c > 0$ such that for every $\varepsilon > 0$ and all generically stable measures $\mu \in \mathfrak{M}_x(M), \nu \in \mathfrak{M}_y(M)$, there are partitions $M^x = A_1 \cup \cdots \cup A_N$ and $M^y = B_1 \cup \cdots \cup B_N$ and a set $\Sigma \subseteq \{1, \ldots, N\}^2$ such that, if $\omega = \mu \otimes \nu$,*

- $K \leq \varepsilon^{-c}$
- *Each A_i and B_j is defined by a boolean combination of instances of ϕ of complexity depending only on ϕ, ε*
- $\sum_{(i,j)\in\Sigma} \omega(A_i \times B_j) \leq \varepsilon$
- *For all $(i, j) \notin \Sigma$, the pair A_i, B_j is ϕ-homogeneous.*

However, that was not the first way the lemma was proven. The original proof, in [CS18], uses the distal cutting lemma to prove that bipartite graphs definable in distal structures satisfy a Ramsey-theoretic statement, called the Definable Strong Erdős-Hajnal property:

Definition 1.2.44 ((Definable) Strong Erdős-Hajnal)**.** We say that a formula $\phi(x; y)$ has the *strong Erdős-Hajnal property* when there exists $\delta > 0$ such that for any finite sets $A \subseteq M^x, B \subseteq M^y$, there are subsets $A' \subseteq A, B' \subseteq B$ such that $|A'| \geq \delta|A|, |B'| \geq \delta|B|$ and (A', B') is ϕ-homogeneous.

A formula $\phi(x; y; w)$ has the *definable* strong Erdős-Hajnal property when there are formulas $\psi(x; z), \theta(y; z)$ and $\delta > 0$ such that for all parameters $b \in M^w$ and all generically stable Keisler measures $\mu \in \mathfrak{M}_x(M), \nu \in \mathfrak{M}_y(M)$, there are parameters $d \in M^z$ such that $\mu(\psi(x; d)) \geq \delta, \nu(\theta(y; d)) \geq \delta$, and the sets $\psi(M; d), \theta(M; d)$ are (ϕ, ε)-homogeneous.

Fact 1.2.45 ([CS18, Theorem 3.1]). *An NIP theory T is distal if and only if every formula has the definable strong Erdős-Hajnal property.*

As a corollary, any theory with a distal expansion has the strong Erdős-Hajnal property.

This property was then recursively applied to find a regularity partition, proving the distal regularity lemma. This proof method has the advantage of producing combinatorial bounds lacking in the proof from [Sim16], but both ways of connecting distality to regularity are useful, and generalized to continuous logic in Chapter 4.

It should also be noted that all of the above results extend in some form to hypergraphs.

CHAPTER 2

Combinatorial Bounds in Distal Structures

This chapter is a modified version of [And23a], copyright © 2023, Cambridge University Press, reprinted with permission. In this chapter, we provide polynomial upper bounds for the minimal sizes of distal cell decompositions in several kinds of distal structures, particularly weakly o-minimal and P-minimal structures. The bound in general weakly o-minimal structures generalizes the vertical cell decomposition for semialgebraic sets, and the bounds for vector spaces in both o-minimal and p-adic cases are tight. We apply these bounds to Zarankiewicz's problem and sum-product bounds in distal structures.

2.1 Introduction

Some of the strongest tools in geometric combinatorics revolve around partitioning space. These techniques fall largely into two categories, the polynomial partitioning method developed by Guth and Katz [GK15], and versions of the cutting lemma for various cell decompositions [CEG91]. While the polynomial method has yielded impressive results, its reliance on Bézout's Theorem limits its scope to questions about algebraic and semialgebraic sets. If one tries to generalize it to sets definable in o-minimal structures other than real closed fields, Bézout's theorem can fail [GKP99]. The cutting lemma method, however, can be generalized to more complicated sets using the language of model theory. Distal cell decompositions, defined in [CGS20], provide an analogous definition to the stratification or vertical cell decomposition results known for $\mathbb{R}$, with a similar cutting lemma, for families of sets definable in a suitable first-order structure, known as a distal structure.

We then study distal cell decompositions through the lens of shatter functions. In [ADH16], the *dual shatter function* π^*_Φ of a set Φ of formulas is defined so that $\pi^*_\Phi(n)$ is the maximum cardinality of the set of Φ-types over a parameter set of size n. We define an analogous shatter function $\pi_{\mathcal{T}}(n)$ for each distal cell decomposition $\mathcal{T}$, where instead of counting all Φ-types, we count the maximum number of cells needed for a distal cell decomposition against n sets (See Definition 2.2.11). This shatter function grows polynomially in a distal structure, so each $\mathcal{T}$ has some exponent $t \in \mathbb{R}$ such that $\pi_{\mathcal{T}}(n) = \mathcal{O}(n^t)$. This exponent is what determines the effectiveness of the cutting lemma for combinatorial applications. Just as the dual VC density of Φ is defined to be the rate of growth of π^*_Φ, we define the *distal density* of Φ to be the infimum of the exponents of all distal cell decompositions $\mathcal{T}$ for Φ.

In this chapter, we construct and bound the sizes of distal cell decompositions for definable families in several distal structures, namely the weakly o-minimal structures, including a better bound on ordered vector spaces, the field $\mathbb{Q}_p$, and its linear reduct. Then we apply these bounds to some combinatorial problems.

2.1.1 Main Results

Our first theorem constructs distal cell decompositions (see Definition 2.2.8) for all sets of formulas $\Phi(x; y)$, with x and y tuples of variables of arbitrary finite length, in some structure $\mathcal{M}$, given a distal cell decomposition for all sets of formulas $\Phi(x; y)$, with with $|x| = 1$. This construction by inducting on the dimension generalizes the stratification result in [CEG91], which essentially constructs distal cell decompositions for $\mathbb{R}$ as an ordered field. It is also similar to Theorem 7.1 in [ADH16], which provides an analogous bound for the VC density of a set of formulas in many dimensions assuming the strong VCd property in dimension 1.

Theorem (Theorem 2.3.1)**.** *Let $\mathcal{M}$ be a structure in which all finite sets $\Phi(x; y)$ of formulas with $|x| = 1$ admit a distal cell decomposition with k parameters (see Definition 2.2.10),* *and for some $d_0 \in \mathbb{N}$, all finite sets $\Phi(x; y)$ of formulas with $|x| = d_0$ admit distal cell decompositions of exponent at most r. Then all finite sets $\Phi(x; y)$ of formulas with $|x| = d \geq d_0$ admit distal cell decompositions of exponent $k(d - d_0) + r$.*

In sections 2.4, 2.5, 2.6, and 2.7, we prove upper bounds on the exponents of distal cell decompositions in weakly o-minimal structures, as well as the field $\mathbb{Q}_p$ and its linear reduct. Those results are summarized and contrasted with the best-known bounds for the dual VC density, in the following theorem:

Theorem 2.1.1. *Let $\mathcal{M}$ be a structure from the first column of this table. Then any formula $\phi(x; y)$ has dual VC density bounded by the corresponding value in the second column, and admits a distal cell decomposition with exponent bounded by the value in the third column. Thus also its distal density is bounded by the value in the third column.*

$\mathcal{M}$	Dual VC density	Distal Density
o-minimal expansions of groups	$\lvert x\rvert$	$2\lvert x\rvert - 2$ (1 if $\lvert x\rvert = 1$)
weakly o-minimal structures	$\lvert x\rvert$	$2\lvert x\rvert - 1$
ordered vector spaces over ordered division rings	$\lvert x\rvert$	$\lvert x\rvert$
Presburger arithmetic	$\lvert x\rvert$	$\lvert x\rvert$
$\mathbb{Q}_p$ the valued field	$2\lvert x\rvert - 1$	$3\lvert x\rvert - 2$
$\mathbb{Q}_p$ in the linear reduct	$\lvert x\rvert$	$\lvert x\rvert$

Table 2.1: Distal Density and Dual VC Density of Formulas in Distal Structures

Proof. The Dual VC density bounds are from [ADH16], except for the bound for the linear reduct of $\mathbb{Q}_p$, which is from [Bob17].

Theorem 2.4.1 establishes the bound for weakly o-minimal structures by constructing a distal cell decomposition in the 1-dimensional case, and then applying Theorem 2.3.1. Taking into account [CGS20], we improve that bound for o-minimal expansions of fields to match the bound from [CEG91] for the case of $\mathbb{R}$ as an ordered field. This improves [Bar13, Theorem 4.0.9], which provides a cell decomposition with $\mathcal{O}(|B|^{2|x|-1})$ uniformly definable cells for $\mathcal{M}$ an o-minimal expansion of a real closed field.

Theorem 2.4.2 shows that the distal density of any finite set of formulas $\Phi(x;y)$ in an ordered vector space over an ordered division ring matches the VC density. In particular, the distal exponent of Φ is bounded by $|x|$, which is optimal. This also works for any o-minimal locally modular expansion of an abelian group, and Theorem 2.5 shows the same results for $\mathbb{Z}$ in Presburger's language.

Theorem 2.6.1 shows that the distal density matches the VC density for any finite set of formulas $\Phi(x;y)$ in $\mathbb{Q}_p$ equipped with its reduced linear structure in the language $\mathcal{L}_{\text{aff}}$ described by Leenknegt in [Lee14]. The proof adapts Bobkov's bound on VC density in the same structure [Bob17].

Theorem 2.7.1 establishes the bound for $\mathbb{Q}_p$ or any other P-minimal field with quantifier-elimination and definable Skolem functions in Macintyre's language by constructing a distal cell decomposition in the 1-dimensional case and applying Theorem 2.3.1. $\square$

Finally in Section 2.8 we apply these results to combinatorics. We combine them with the results on Zarankiewicz's problem from [CGS20] to prove a bound on the number of edges in bipartite graphs definable in distal structures which omit some (oriented) complete bipartite graph $K_{s,u}$, similar to the bound given by Theorem 1.2 from [FPS17].

Corollary (Corollary 2.8.7, expressed in terms of distal density)**.** *Let $\mathcal{M}$ be a structure and $t \in \mathbb{N}_{\geq 2}$. Assume that $E(x,y) \subseteq M^{|x|} \times M^{|y|}$ is a definable relation given by an instance of a formula $\theta(x,y;z) \in \mathcal{L}$, such that the formula $\theta'(x;y,z) := \theta(x,y;z)$ has distal density at most t, and the graph $E(x,y)$ does not contain $K_{s,u}$. Then for every $\varepsilon \in \mathbb{R}_{>0}$, there is a constant $\alpha = \alpha(\theta, s, u, \varepsilon)$ satisfying the following.*

For any finite $P \subseteq M^{|x|}, Q \subseteq M^{|y|}$, $|P| = m, |Q| = n$, we have:

$$|E(P,Q)| \leq \alpha \left(m^{\frac{(t-1)s}{ts-1}+\varepsilon} n^{\frac{t(s-1)}{ts-1}} + m + n \right).$$

This corollary then lets us quickly prove bounds on graphs in the following contexts:

Corollary (Corollary 2.8.8)**.** *Assume that $E(x,y) \subseteq \mathbb{R}^{|x|} \times \mathbb{R}^{|y|}$ is a relation given by a boolean combination of exponential-polynomial (in)equalities, and the graph $E(x,y)$ does not contain $K_{s,u}$. Then there is a constant $\alpha = \alpha(\theta, s, u)$ satisfying the following.*

For any finite $P \subseteq \mathbb{R}^{|x|}, Q \subseteq \mathbb{R}^{|y|}$, $|P| = m, |Q| = n$, we have:

$$|E(P,Q)| \leq \alpha \left(m^{\frac{(2|x|-2)s}{(2|x|-1)s-1}} n^{\frac{(2|x|-1)(s-1)}{(2|x|-1)s-1}+\varepsilon} + m + n \right).$$

(Here an *exponential-polynomial (in)equality* is an (in)equality between functions $\mathbb{R}^n \to \mathbb{R}$ in $\mathbb{Z}[x_1, \ldots, x_n, e^{x_1}, \ldots, e^{x_n}]$ as in [BKW10].)

Corollary (Corollary 2.8.10)**.** *Assume that $E(x,y) \subseteq \mathbb{Z}_p^{|x|} \times \mathbb{Z}_p^{|y|}$ is a subanalytic relation, and the graph $E(x,y)$ does not contain $K_{s,u}$. Then there is a constant $\alpha = \alpha(\theta, s, u)$ satisfying the following.*

For any finite $P \subseteq \mathbb{Z}_p^{|x|}, Q \subseteq \mathbb{Z}_p^{|y|}$, $|P| = m, |Q| = n$, we have:

$$|E(P,Q)| \leq \alpha \left(m^{\frac{(3|x|-3)s}{(3|x|-2)s-1}} n^{\frac{(3|x|-2)(s-1)}{(3|x|-2)s-1}+\varepsilon} + m + n \right).$$

Here subanalytic relations are defined in the sense of [DHM99].

2.2 Preliminaries

In this section, we review the notation and model-theoretic framework necessary to understand distal cell decompositions. For further background on these definitions, see [CS18] and

[CGS20].

Firstly, we review asymptotic notation:

Definition 2.2.1. Let $f, g : \mathbb{N} \to \mathbb{R}_{\geq 0}$.

- We will say $f(x) = \mathcal{O}(g(x))$ to indicate that there exists $C \in \mathbb{R}_{>0}$ such that for $n \in \mathbb{N}_{>0}$, $f(n) \leq Cg(x)$.
- We will say $f(x) = \Omega(g(x))$ to indicate that there exists $C \in \mathbb{R}_{>0}$ such that for $n \in \mathbb{N}_{>0}$, $f(n) \geq Cg(x)$.

If $f, g : \mathbb{N} \times \mathbb{N} \to \mathbb{R}_{\geq 0}$, then $f(x, y) = \mathcal{O}(g(x, y))$ indicates that there is a constant $C \in \mathbb{R}_{>0}$ such that for all $m, n \in \mathbb{N}_{>0}$, $f(m, n) \leq Cg(m, n)$.

Throughout this section, let $\mathcal{M}$ be a first-order structure in the language $\mathcal{L}$. We will frequently refer to $\Phi(x; y)$ as a set of formulas, which will implicitly be in the language $\mathcal{L}$. Each formula in Φ will have the same variables, split into a tuple x and a tuple y, where, for instance, $|x|$ represents the length of the tuple x. We use M to refer to the universe, or underlying set, of $\mathcal{M}$, and M^n to refer to its nth Cartesian power. If $\phi(x; y)$ is a formula with its variables partitioned into x and y, and $b \in M^{|y|}$, then $\phi(M^{|x|}; b)$ refers to the definable set $\{a \in M^{|x|} : \mathcal{M} \models \phi(a, b)\}$. We also define the dual formula of $\phi(x; y)$ to be $\phi^*(y; x)$ such that $\mathcal{M} \models \forall x \forall y \phi(x; y) \leftrightarrow \phi^*(y; x)$, and similarly define $\Phi^*(y; x)$ to be the set $\{\phi^*(y; x) : \phi(x; y) \in \Phi(x; y)\}$.

Definition 2.2.2. For sets $A, X \subseteq M^d$, we say that A *crosses* X if both $X \cap A$ and $X \cap \neg A$ are nonempty.

Definition 2.2.3. Let $B \subseteq M^t$.

- For $\phi(x; y)$ with $|y| = t$, we say that $\phi(x; B)$ *crosses* $X \subseteq M^{|x|}$ when there is some $b \in B$ such that $\phi(M^{|x|}; b)$ crosses X.

- For $\Phi(x;y)$ with $|y| = t$, we say that $X \subseteq M^{|x|}$ *is crossed by* $\Phi(x;B)$ when there is some $\phi \in B$ such that $\phi(x;B)$ crosses X.

Definition 2.2.4. We define $S^{\Phi}(B)$ to be the set of complete Φ-types over a set $B \subseteq M^{|y|}$ of parameters, or alternately, the set of maximal consistent subsets of $\{\varphi(x;b) : \varphi \in \Phi, b \in B\} \cup \{\neg\varphi(x;b) : \varphi \in \Phi, b \in B\}$.

Throughout this chapter, we will want to use the concepts of VC density and dual VC density.

Definition 2.2.5. Let $\Phi(x;y)$ be a finite set of formulas.

- For $B \subseteq M^{|y|}$, define $\pi^*_{\Phi}(B) := \left|S^{\Phi}(B)\right|$.
- For $n \in \mathbb{N}$, define $\pi^*_{\Phi}(n) := \max_{B \subseteq M^{|y|}, |B|=n} \pi^*_{\Phi}(B)$.
- Define the *dual VC density* of Φ, $\mathrm{vc}^*(\Phi)$, to be the infimum of all $r \in \mathbb{R}_{>0}$ such that there exists $C \in \mathbb{R}$ with $\left|S^{\Phi}(B)\right| \leq C|B|^r$ for all choices of B. Equivalently, we can define $\mathrm{vc}^*(\Phi)$ to be
$$\limsup_{n \to \infty} \frac{\log \pi^*_{\Phi}(n)}{\log n}.$$
- Dually, we define $\pi_{\Phi} := \pi^*_{(\Phi^*)}$ and define the *VC density* of Φ to be $\mathrm{vc}(\Phi) = \mathrm{vc}^*(\Phi^*)$.

This definition of (dual) VC density of sets of formulas comes from Section 3.4 of [ADH16], which relates it to the other definitions of VC density.

Definition 2.2.6. An *abstract cell decomposition* for $\Phi(x;y)$ is a function $\mathcal{T}$ that assigns to each finite $B \subset M^{|y|}$ a set $\mathcal{T}(B)$ whose elements, called cells, are subsets of $M^{|x|}$ not crossed by $\Phi(x;B)$, and cover $M^{|x|}$ so that $M^{|x|} = \bigcup \mathcal{T}(B)$.

Example 1. Fix $\Phi(x;y)$. For each type $p(x) \in S^{\Phi}(B)$, the set $p(M^{|x|})$ is a definable subset of $M^{|x|}$, as $p(x)$ is equivalent to a boolean combination of formulas $\phi(x;b)$ for $\phi \in \Phi$ and $b \in B$. Define $\mathcal{T}_{\mathrm{vc}}(B) := \{p(M^{|x|}) : p \in S^{\Phi}(B)\}$. Then $\mathcal{T}_{\mathrm{vc}}$ is an abstract cell decomposition with $|\mathcal{T}_{\mathrm{vc}}(B)| = \left|S^{\Phi}(B)\right| = \pi^*_{\Phi}(B)$.

Proposition 2.2.7. *For any abstract cell decomposition $\mathcal{T}$ of $\Phi(x; y)$ and any finite $B \subseteq M^{|y|}$, $|\mathcal{T}(B)| \geq \pi^*_\Phi(B)$.*

Proof. As each cell $\Delta \in \mathcal{T}(B)$ is not crossed by $\Phi(x; B)$, its elements must all have the same Φ-types over B. Thus there is a function $f : \mathcal{T}(B) \to S^\Phi(B)$ sending each cell to the Φ-type over B of its elements. Each type in $S^\Phi(B)$ is consistent and definable by a formula, and thus must be realized in M, so there must be at least one cell of $\mathcal{T}(B)$ containing formulas of that type. Thus f is a surjection, and $|\mathcal{T}(B)| \geq \left|S^\Phi(B)\right|$. $\square$

Definition 2.2.8. Let $\Phi(x; y)$ be a finite set of formulas without parameters. Then a *distal cell decomposition* $\mathcal{T}$ for Φ is an abstract cell decomposition defined using the following data:

- A finite set $\Psi(x; y_1, \ldots, y_k)$ of formulas (without parameters) where $|y_1| = \cdots = |y_k| = |y|$.
- For each $\psi \in \Psi$, a formula (without parameters) $\theta_\psi(y; y_1, \ldots, y_k)$.

Given a finite set $B \subseteq M^{|y|}$, let $\Psi(B) := \{\psi(M^{|x|}; b_1, \ldots, b_k) : \psi \in \Psi, b_1, \ldots, b_k \in B\}$. This is the set of potential cells from which the cells of the decomposition are chosen. Then for each potential cell $\Delta = \psi(M^{|x|}; b_1, \ldots, b_k)$, we let $\mathcal{I}(\Delta) = \theta_\psi(M^{|y|}; b_1, \ldots, b_k)$. Then we define $\mathcal{T}(B)$ by choosing the cells $\Delta \in \Psi(B)$ such that $B \cap \mathcal{I}(\Delta) = \emptyset$, that is, $\mathcal{T}(B) = \{\Delta \in \Psi(B) : B \cap \mathcal{I}(\Delta) = \emptyset\}$.

In the rest of this chapter, when $\Phi(x; y)$ is a finite set of formulas, we will assume that Φ is defined without parameters.

The following lemma will be useful in defining distal cell decompositions later on:

Lemma 2.2.9. *Let $\Phi(x; y)$ be a finite set of formulas, and let $\Phi'(x; y)$ be a finite set of formulas such that each formula in Φ is a boolean combination of formulas in Φ'. Then if $\mathcal{T}$ is a distal cell decomposition for Φ', it is also a distal cell decomposition for Φ.*

Proof. The definability requirements for a distal cell decomposition do not depend on the set of formulas Φ, so it suffices to show that $\mathcal{T}$ is an abstract cell decomposition for Φ, or that for a given B, each cell $\Delta \in \mathcal{T}(B)$ is not crossed by $\Phi(x; B)$. As for any $\varphi \in \Phi$, $b \in B$, $\varphi(x; b)$ is a boolean combination of formulas in $\Phi'(x; B)$, and all of these have a fixed truth value on Δ, so does $\varphi(x; b)$. $\square$

We now consider a few ways of counting the sizes of distal cell decompositions:

Definition 2.2.10. Let $\mathcal{T}$ be a distal cell decomposition for the finite set of formulas $\Phi(x; y)$, whose cells are defined by formulas in the set Ψ.

- We say that $\mathcal{T}$ *has k parameters* if every formula in Ψ is of the form $\psi(x; y_1, \ldots, y_k)$.
- We say that $\mathcal{T}$ *has exponent r* if $|\mathcal{T}(B)| = \mathcal{O}(|B|^r)$ for all finite $B \subseteq M^{|y|}$.

Note that even if $\mathcal{T}$ has k parameters, not every formula ψ used to define $\mathcal{T}$ needs to use all k parameters. In practice, we will sometimes define distal cell decompositions using formulas with different numbers of variables, but as each distal cell decomposition is defined using finitely many formulas, we can just take k to be the maximum number of parameters used by any one formula, and add implicit variables to the rest.

Definition 2.2.11. Let $\Phi(x; y)$ be a finite set of formulas. Then define the *distal density* of Φ to be the infimum of all reals $r \geq 0$ such that there exists a distal cell decomposition $\mathcal{T}$ of Φ of exponent r. If no $\mathcal{T}$ exists for Φ, the distal density is defined to be ∞.

Problem 2.2.12. Note that if Φ has distal density t, it is not known if θ must have a distal cell decomposition of exponent precisely t.

Definition 2.2.13. We also define a shatter function $\pi_{\mathcal{T}}(n) := \max_{|B|=n} |\mathcal{T}(B)|$. The distal density of Φ can equivalently be defined as the infimum of

$$\limsup_{n\to\infty} \frac{\log \pi_{\mathcal{T}}(n)}{\log n}$$

over all distal cell decompositions $\mathcal{T}$ of Φ, if any exist.

Proposition 2.2.14. *For any finite set of formulas $\Phi(x; y)$, $\pi_{\mathcal{T}}(n) \geq \pi^*_\Phi(n)$ for all $n \in \mathbb{N}$, and the distal density of Φ is at least $\mathrm{vc}^*(\Phi)$.*

Proof. By Proposition 2.2.7, for every distal cell decomposition $\mathcal{T}$, $|\mathcal{T}(B)| \geq \left|S^\Phi(B)\right|$. Thus

$$\mathrm{vc}^*(\Phi) \leq \limsup_{n\to\infty} \frac{\log \pi^*_\Phi(n)}{\log n} \leq \limsup_{n\to\infty} \frac{\log \pi_{\mathcal{T}}(n)}{\log n}$$

so after taking the infimum over all $\mathcal{T}$, the distal density is at least $\mathrm{vc}^*(\Phi)$. □

Also, just by defining $\Phi(x; y)$ to be $\{x = y\}$, where $|x| = |y| = d$, we see that $\left|S^\Phi(B)\right| \geq |B|^d$, so we see that for every d, there is a Φ with both VC- and distal densities at least d in any structure.

Example 2. Chernikov, Galvin and Starchenko found that if $\mathcal{M}$ is an *o*-minimal expansion of a field, and $|x| = 2$, then any $\Phi(x; y)$ admits a distal cell decomposition with $|\mathcal{T}(B)| = \mathcal{O}(|B|^2)$ for all finite B [CGS20]. Thus the distal density of such a Φ is at most 2.

So far, we have defined distal cell decompositions and distal density in the context of a particular structure. In fact, if $\Phi(x; y)$ is a finite set of $\mathcal{L}$-formulas, and T a complete $\mathcal{L}$-theory, we will show that the distal density of $\Phi(x; y)$ is the same in every model of T, so we can define *the distal density of* Φ *over* T to be the distal density of Φ in any model of T. (This uses the fact that the formulas in Φ and the formulas defining a distal cell decomposition are required to be parameter-free.)

Proposition 2.2.15. *Let $\Phi(x; y)$ be a finite set of $\mathcal{L}$-formulas, and $\mathcal{M} \equiv \mathcal{M}'$ be elementarily equivalent $\mathcal{L}$-structures. Then if Φ admits a distal cell decomposition $\mathcal{T}$ in $\mathcal{M}$, the same formulas define a distal cell decomposition for Φ in $\mathcal{M}'$. Thus we can refer to $\mathcal{T}$ as being a distal cell decomposition for Φ over the theory $T = \mathrm{Th}(\mathcal{M})$. Also, the shatter function $\pi_{\mathcal{T}}$, and thus the distal exponent of $\mathcal{T}$ and the distal density of Φ, will be equal for $\mathcal{M}$ and $\mathcal{M}'$, and can be viewed as properties of the theory T.*

Proof. Let $\mathcal{T}$ be a distal cell decomposition for Φ over $\mathcal{M}$, consisting of a set $\Psi(x; y_1, \ldots, y_k)$ of formulas, and a formula θ_ψ for each $\psi \in \Psi$ (as in Definition 2.2.8). Then to verify that the same formulas define a distal cell decomposition for Φ over $\mathcal{M}'$, we must simply check that for all finite $B \subset M'^{|y|}$, the set of cells $\mathcal{T}(B)$ covers $M'^{|x|}$, and that no cell of $\mathcal{T}(B)$ is crossed by $\Phi(x; B)$.

It is enough to show that these facts can be described with first-order sentences. Fix some natural number n, and we will find a first-order sentence that shows that for all $B = \{b_1, \ldots, b_n\}$, the cells of $\mathcal{T}(B)$ cover the space and are not crossed. We can encode that the cells of $\mathcal{T}(B)$ cover $M'^{|x|}$ with the sentence

$$\forall y_1, \ldots, y_n, \forall x, \bigwedge_{\psi \in \Psi, i_1, \ldots, i_k \in \{1, \ldots, n\}} \psi(x; y_{i_1}, \ldots, y_{i_k}) \wedge \bigwedge_{i=1}^{n} \neg\theta_\psi(y_i; y_{i_1}, \ldots, y_{i_k}).$$

When interpreted over $\mathcal{M}'$, this simply states that for any choice of n parameters $b_1, \ldots, b_n$ and any $x_0 \in M'^{|x|}$, there is some $\psi, i_1, \ldots, i_k$ such that $\psi(x; b_{i_1}, \ldots, b_{i_k})$ defines a valid cell, which contains x_0. Similarly, to show that the cell defined by $\psi(x; b_{i_1}, \ldots, b_{i_k})$, if included in the cell decomposition, is not crossed by $\Phi(x; B)$, we can use the following sentence, showing that for all $B = \{b_1, \ldots, b_n\}$, if for some i and some $\varphi \in \Phi$, $\phi(x; b_i)$ crosses $\psi(x; b_{i_1}, \ldots, b_{i_k})$, then $\psi(x; b_{i_1}, \ldots, b_{i_k})$ is not a valid cell:

$$\begin{aligned}
&\forall y_1, \ldots, y_n, \\
&\quad \left(\bigvee_{\varphi \in \Phi, 1 \leq i \leq n} \exists x_1, x_2, \varphi(x_1; y_i) \wedge \neg\varphi(x_2; y_i) \wedge \psi(x_1; y_{i_1}, \ldots, y_{i_k}) \wedge \psi(x_2; y_{i_1}, \ldots, y_{i_k}) \right) \\
&\quad \to \bigvee_{i=1}^{n} \theta_\psi(y_i; y_{i_1}, \ldots, y_{i_k}).
\end{aligned}$$

Now it suffices to show that the shatter function $\pi_{\mathcal{T}}$ is the same in both models, as the distal exponent of $\mathcal{T}$ and distal density of Φ are defined in terms of these shatter functions.

To say that $\pi_{\mathcal{T}}(n) \leq m$ in $\mathcal{M}$ is to say that for all $b_1, \ldots, b_n \in M^{|y|}$, there are at most

m cells in $\mathcal{T}(B)$. This is the disjunction of a finite number of cases, which we will index by $A_1, \ldots, A_m$, where each $A_i \subset \Psi \times \{1, \ldots, n\}^k$, as each tuple $t = (\psi_t, t_1, \ldots, t_n) \in \Psi \times \{1, \ldots, n\}^k$ corresponds to a potential cell $\Delta_s = \psi_t(x; b_{t_1}, \ldots, b_{t_n})$. Then in the case indexed by $A_1, \ldots, A_m$, there is a first-order sentence stating that for all $1 \leq i \leq n$ and $s, t \in A_i$, the formulas Δ_s and Δ_t are equivalent, and for all tuples $t = (\psi, i_1, \ldots, i_n)$ not contained in any A_i, t is not a valid cell, as implied by $\bigvee_{j=1}^{n} \theta_\psi(b_j; b_{i_1}, \ldots, b_{i_n})$. The disjunction of all these sentences states that there are at most m distinct cells in $\mathcal{T}(\{b_1, \ldots, b_n\})$, and if $b_1, \ldots, b_n$ are replaced with universally-quantified variables, we find a sentence that states that $\pi_{\mathcal{T}}(n) \leq m$. Thus for all n, $\pi_{\mathcal{T}}(n)$ evaluates to the same number over any model of the theory of $\mathcal{M}$. $\square$

Distality of a theory was defined originally in terms of indiscernible sequences in [Sim13]. We will not present that definition here, but we will take the following equivalence as a definition:

Fact 2.2.16. *The following are equivalent for any first-order structure $\mathcal{M}$:*

1. *$\mathcal{M}$ is distal.*
2. *For every formula $\phi(x; y)$, $\{\phi\}$ admits a distal cell decomposition.*
3. *For every finite set of formulas $\Phi(x; y)$, Φ admits a distal cell decomposition.*

Proof. The equivalence of (1) and (2) is from [CS15] (see [CGS20, Fact 2.9] for a discussion). Clearly (3) implies (2), so it suffices to show that (2) implies (3).

For a given $\Phi(x; y)$, assume each $\phi \in \Phi$ admits a distal cell decomposition $\mathcal{T}_\phi$. Then for finite $B \subseteq M^{|y|}$, we define $\mathcal{T}(B)$ to consist of all nonempty intersections $\bigcap_{\phi \in \Phi} \Delta_\phi$, where each Δ_ϕ is chosen from $\mathcal{T}_\phi(B)$. These cells will cover $M^{|x|}$, as each $a \in M^{|x|}$ belongs to some Δ_ϕ for each ϕ, and thus belongs to their intersection. Any cell $\Delta = \bigcap_{\phi \in \Phi} \Delta_\phi$ will not be crossed by $\Phi(x; B)$, as for each $\phi \in \Phi$, as $\Delta \subset \Delta_\phi$, and Δ_ϕ is not crossed by $\phi(x; B)$.

Now we check that this cell decomposition is uniformly definable. For each $\phi \in \Phi$, let $\mathcal{T}_\phi$ consist of Ψ_ϕ and $\{\theta_\psi : \psi \in \Psi_\phi\}$. Then $\mathcal{T}$ can be defined by the set of formulas Ψ consisting of all conjunctions $\bigwedge_{\phi\in\Phi} \psi_\phi$ where $\psi_\phi \in \Psi_\phi$ for each ϕ. For a given $\Delta = \bigcap_{\phi\in\Phi} \Delta_\phi$, we can let $\mathcal{I}(\Delta) = \bigcup_{\phi\in\Phi} \mathcal{I}(\Delta_\phi)$. $\square$

Examples of distal structures include:

- o-minimal structures
- Presburger arithmetic $(\mathbb{Z}, 0, +, <)$
- The field of p-adics $\mathbb{Q}_p$ and other P-minimal fields.
- The linear reduct of $\mathbb{Q}_p$, in the language $\mathcal{L}_{\text{aff}}$.

For justification of the first three of these, see [CS18]. The distality of these structures is established using the indiscernible sequence definition, which does not provide good bounds. In what follows, we will construct explicit distal cell decompositions for all of these examples.

2.3 Dimension Induction

In this section, we provide a bound on the size of distal cell decompositions for all dimensions, given a bound for distal cell decompositions for a fixed dimension in an arbitrary distal structure. This allows us to bound the size of a distal cell decomposition for any finite family of formulas in several kinds of distal structures, including any o-minimal structures. This approach is inspired by the partition construction in [CEG91], which can be interpreted as constructing distal cell decompositions in the context of $\mathbb{R}$ as an ordered field. (It also improves the bound in [ACG22, Proposition 1.9].)

Theorem 2.3.1. *Let $\mathcal{M}$ be a structure in which all finite sets $\Phi(x; y)$ of formulas with $|x| = 1$ admit a distal cell decomposition with k parameters (see Definition 2.2.10), and for* *some $d_0 \in \mathbb{N}$, all finite sets $\Phi(x;y)$ of formulas with $|x| = d_0$ admit distal cell decompositions of exponent at most r. Then all finite sets $\Phi(x;y)$ of formulas with $|x| = d \geq d_0$ admit distal cell decompositions of exponent $k(d-d_0)+r$.*

Proof. The case with $d = d_0$ follows directly from the assumptions, so we can proceed by induction. Assume the result for all finite sets of formulas with $|x| = d-1 \geq d_0$. Then we will build a distall cell decomposition for a $\Phi(x;y)$ with $|x| = d$. Where $x = (x_1, \ldots, x_d)$, let $x' = (x_2, \ldots, x_d)$. We start by fixing a distal cell decomposition $\mathcal{T}_1$ for the set of formulas $\Phi_1(x_1; x', y) := \{\phi(x_1; x', y) : \phi(x;y) \in \Phi\}$. Let the cells of $\mathcal{T}_1$ be defined by $\Psi_1(x_1; x'_1, y_1, \ldots, x'_k, y_k)$ and a formula $\theta_\psi(x', y; x'_1, y_1, \ldots, x'_k, y_k)$ for each $\psi \in \Psi_1$. For this construction, we will only use $\mathcal{T}_1$ to define Φ_1-types over sets of the form $\{a'\} \times B$. Because each element of that set has the same first coordinate, we will abbreviate the formula $\psi(x_1; x'_1, y_1, x'_2, y_2, \ldots, x'_1, y_k)$ as $\psi(x_1; x', y_1, y_2, \ldots, y_k)$, assuming all the variables x'_i are equal. Similarly, we abbreviate $\theta_\psi(x', y; x'_1, y_1, \ldots, x'_k, y_k)$ as $\theta_\psi(x', y; y_1, \ldots, y_k)$, setting each x'_i equal to x'. We will also want to repartition the variables, setting $\theta_\psi*(x'; y_1, \ldots, y_k, y) := \theta_\psi(x', y; y_1, \ldots, y_k)$.

For each $\psi \in \Psi_1$, let $\Phi_\psi(x'; y_1, \ldots, y_k, y)$ be the set of formulas consisting of $\theta_\psi*$ and all formulas of the form $\forall x_1, \psi(x_1; x', y_1, \ldots, y_k) \to \Box\phi(x_1, x'; y)$ where $\phi \in \Phi$, and $\Box$ is either $\neg$ or nothing.

Then let $\mathcal{T}_\psi$ be a distal cell decomposition for Φ_ψ, consisting of Ψ_ψ and a formula $\theta_{\psi'}$ for each $\psi' \in \Psi_\psi$. As before, we will assume some of the variables are equal, and write these formulas more succinctly, assuming that our set of parameters is of the form $\{(b_1, \ldots, b_k)\} \times B$ for some $b_1, \ldots, b_k \in M$ and finite $B \subseteq M^{|y|}$. This allows us to write each $\psi' \in \Psi_\psi$ as $\psi'(x'; y_1, \ldots, y_k, y'_1, \ldots, y'_m)$, and write $\theta_{\psi'}$ as $\theta_{\psi'}(y; y_1, \ldots, y_k, y'_1, \ldots, y'_m)$.

For each $\psi \in \Psi_1$ and $\psi' \in \Psi_\psi$, let $\psi \otimes \psi'(x; y_1, \ldots, y_k, y'_1, \ldots, y'_m)$ be the formula

$$\psi'(x'; y_1, \ldots, y_k, y'_1, \ldots, y'_m) \wedge \psi(x_1; x', y_1, \ldots, y_k).$$

(Intuitively, this defines a sort of cylindrical cell in $M^{|x|}$, where x' is in a cell of one cell decomposition of $M^{|x'|}$, and x_1 is in a cell of a cell decomposition of M, defined using x' as a parameter.) Let $\Psi(x; y_1, \ldots, y_k, y'_1, \ldots, y'_m) = \{\psi \otimes \psi' : \psi \in \Psi_1, \psi' \in \Psi_\psi\}$. We will use Ψ to define a distal cell decomposition $\mathcal{T}$ for $\Phi(x; y)$.

To define $\mathcal{T}$, it suffices to define $\theta_{\psi\otimes\psi'}$ for each $\psi \in \Psi_1, \psi' \in \Psi_\psi$. Define

$$\begin{aligned}&\theta_{\psi\otimes\psi'}(y; y_1, \ldots, y_k, y'_1, \ldots, y'_m) := \\ &\qquad \theta_{\psi'}(y; y_1, \ldots, y_k, y'_1, \ldots, y'_m) \wedge (\exists x', \psi'(x'; y_1, \ldots, y_k, y'_1, \ldots, y'_m) \wedge \theta_\psi(x'; y_1, \ldots, y_k, y)).\end{aligned}$$

This means that if Δ is the cell $\psi \otimes \psi'(M^d; b_1, \ldots, b_k, b'_1, \ldots, b'_m)$, then

$$\begin{aligned}\mathcal{I}(\Delta) &:= \{b \in M^{|y|} : \mathcal{M} \models \theta_{\psi\otimes\psi'}(b; b_1, \ldots, b_k, b'_1, \ldots, b'_m)\} \\ &= \{b \in M^{|y|} : \exists (a_1, a') \in \Delta, \mathcal{M} \models \theta_\psi(a', b; b_1, \ldots, b_k)\}.\end{aligned}$$

Thus for all a' in the projection of Δ onto M^{d-1}, the fiber $\{a_1 \in M : (a_1, a') \in \Delta\}$ is a cell of $\mathcal{T}_1(\{a'\} \times B)$ if and only if $B \cap \mathcal{I}(\Delta) = \emptyset$.

Now we show that this definition of $\mathcal{T}$ gives a valid distal cell decomposition for $\Phi(x; y)$. Fix a finite $B \subset M^{|y|}$ and let $a \in M^d$ be given. Firstly, each element of M^d is contained in a cell. If $a = (a_1, a')$ with $a_1 \in M, a' \in M^{d-1}$, then a_1 is in some cell of $\mathcal{T}_1(\{a'\} \times B)$, and that cell is defined by some $\psi(x_1; a', b_1, \ldots, b_k)$, so for all $b \in B$, $\mathcal{M} \models \neg\theta_\psi * (a'; b_1, \ldots, b_k, b)$. Therefore a' is in some cell of $\mathcal{T}_\psi(\{(b_1, \ldots, b_k)\} \times B)$ on which $\mathcal{M} \models \neg\theta_\psi * (x'; b_1, \ldots, b_k, b)$. If that cell is defined by $\psi'(b_1, \ldots, b_k, b'_1, \ldots, b'_m)$, then we can now define a cell containing a by $\psi \otimes \psi'(x; b_1, \ldots, b_k, b'_1, \ldots, b'_m)$.

Secondly, we show that each cell of $\mathcal{T}(B)$ is not crossed by $\Phi(x; B)$. Fix a cell $\Delta \in \mathcal{T}(B)$, and fix $\phi \in \Phi$, $b \in B$. We know that for each a' in the projection of Δ onto M^{d-1}, the fiber $\{a_1 \in M : (a_1, a') \in \Delta\}$ is a cell of $\mathcal{T}_1(\{a'\} \times B)$, so that fiber is not crossed by $\phi(x; B)$. We also guaranteed that if Δ is defined by the formula $\psi \otimes \psi'(x, b_1, \ldots, b_k, b'_1, \ldots, b'_m)$, then the projection of Δ onto M^{d-1} is a cell of $\mathcal{T}_\psi(B)$, so it is not crossed by the formulas $\forall x_1, \psi(x_1; x', b_1, \ldots, b_k) \to \phi(x_1, x'; b)$ and $\forall x_1, \psi(x_1; x', b_1, \ldots, b_k) \to \neg\phi(x_1, x'; b)$. If for some (a_1, a') in Δ, $\mathcal{M} \models \phi(a_1, a'; b)$, then $\mathcal{M} \models \forall x_1, \psi(x_1; x', b_1, \ldots, b_k) \to \phi(x_1, x'; b)$ for $x' = a'$, and thus for all x' in the projection of Δ, so $\mathcal{M} \models \psi(x; b)$ for all $x \in \Delta$.

Finally we can count the number of cells of $\mathcal{T}(B)$. For each $\psi \in \mathcal{T}_1$, and each $b_1, \ldots, b_k$, there are, by induction, $\mathcal{O}(|B|^{k((d-1)-d_0)+r})$ cells in $\mathcal{T}'(\{(b_1, \ldots, b_k)\} \times B)$, each inducing a cell of $\mathcal{T}(B)$. Multiplying by the $|B|^k$ possible tuples $(b_1, \ldots, b_k) \in B^k$ and a finite number of formulas ψ, we get the desired bound $\mathcal{O}(|B|^{k(d-d_0)+r})$. □

2.4 Weakly o-Minimal Structures

In any structure $\mathcal{M}$, for any n, there is a formula $\phi(x; y)$ with $|x| = n$ such that the the dual VC density of ϕ is $|x|$, giving a lower bound on the distal density (see [ADH16, Section 1.4]). In this section, we construct an optimal distal cell decomposition for the case $|x| = 1$, and then use Theorem 2.3.1 to construct distal cell decompositions for all Φ, and bound their sizes. In the case where $\mathcal{M}$ is an o-minimal expansion of a group, we start instead with the optimal bound for $|x| = 2$ from [CS18] and obtain a the bound on the size of the sign-invariant stratification in [CEG91], and improves the bounds on [Bar13, Theorem 4.0.9].

Theorem 2.4.1. *If $\Phi(x; y)$ is a finite family of formulas in a weakly o-minimal structure $\mathcal{M}$, then Φ admits a distal cell decomposition for Φ with exponent $2|x| - 1$.*

If $\mathcal{M}$ is an o-minimal expansion of a group and $|x| \geq 2$, then the distal density is at most $2|x| - 2$.

Proof. In any weakly o-minimal structure, if $\Phi(x; y)$ has $|x| = 1$, then there exists a distal cell decomposition $\mathcal{T}$ with $|\mathcal{T}(B)| = \mathcal{O}(|B|)$ with 2 parameters.

Indeed, by weak o-minimality, for any $\varphi(x; y) \in \Phi$ with $|x| = 1$, there is some number N_φ such that the set $\varphi(M; b)$ is a union of at most N_φ convex subsets for any $b \in M^{|y|}$. Let

$N := \max_{\varphi \in \Phi} N_\varphi$. Then for each $\varphi(x; y) \in \Phi$, we can define formulas $\varphi^1(x; y), \dots, \varphi^N(x; y)$ by

$$\begin{aligned}\varphi^n(x; y) := &\exists x_1, x_2, \dots, x_{n-1}, y_1, \dots, y_{n-1}, \\ &\varphi(x; y) \wedge (x_1 < y_1 < x_2 < \cdots < y_{n-1} < x) \wedge \bigwedge_{i=1}^{n-1} (\varphi(x_i; y) \wedge \neg\varphi(y_i; y))\end{aligned}$$

and then

$$\varphi(M; b) = \varphi^1(M; b) \cup \cdots \cup \varphi^N(M; b)$$

for all b, each $\varphi^i(M; b)$ is convex, and $\varphi^i(M; b) < \varphi^{i+1}(M; b)$ for each i, in the sense that for every $x_i \in \varphi^i(M; b)$ and $x_{i+1} \in \varphi^{i+1}(M; b)$, $x_i < x_{i+1}$.

Then for each $\varphi \in \Phi$ we can also define

$$\varphi^i_{\le}(x; y) := \exists x_0(\varphi^i(x_0; y) \wedge x \le x_0),$$

$$\varphi^i_{<}(x; y) := \forall x_0(\varphi^i(x_0; y) \to x < x_0).$$

Note that each $\varphi^i_\square(M; b)$ for $\square \in \{<, \le\}$ is closed downwards. Thus for any finite subset $B \subset M^{|y|}$, the family of sets $\mathcal{F}(B) = \{\varphi^i_\square(M, b) : b \in B, \varphi \in \Phi, 1 \le i \le N, \square \in \{<, \le\}\}$ is linearly ordered under inclusion. Thus the atoms in the boolean algebra $\mathcal{B}$ generated by $\mathcal{F}(B)$ are of the form $X_1 \setminus X_2$ where $X_1, X_2 \in \mathcal{F}(B)$ and X_2 is the unique maximal element of $\mathcal{F}(B)$ properly contained in X_1, or $\mathcal{M} \setminus X_1$ where X_1 is the unique maximal element of $\mathcal{F}(B)$. Thus only one atom of the boolean algebra can be of the form $X_1 \setminus X_2$ for each X_1, and thus the number of such atoms is at most $|\mathcal{F}(B)| + 1$, which is $\mathcal{O}(|B|)$.

Now we construct $\mathcal{T}$. We let Ψ consist of the formulas of the form $\psi(x; y_1, y_2) := \varphi^i_{\square_1}(x; y_1) \wedge \neg\varphi^j_{\square_2}(x; y_2)$ or $\psi(x; y) := \neg\varphi^j_{\square_1}(x; y)$ with $1 \le i \le N, \square \in \{<, \le\}$, and then for each potential cell $\Delta = \psi(M; b_1, b_2)$, let $\mathcal{I}(\Delta)$ just consist of all $b \in M^{|y|}$ such that Δ is crossed by $\varphi_0(M; b)$ for some $\varphi_0 \in \Phi$. Then $\mathcal{T}(B)$ is exactly the set of atoms in the boolean algebra generated by $\mathcal{F}(B)$, so $|\mathcal{T}(B)| = \mathcal{O}(|B|)$. Each cell is not crossed by any set in $\mathcal{F}(B)$, and thus not by any $\varphi(x; B)$, or $\Phi(x; B)$ itself, so this is a valid distal cell decomposition, where every cell is defined using at most 2 parameters from B.

Thus we can use Theorem 2.3.1, setting $d_0 = 1$, $r = 1$, and $k = 2$, to find that any family of formulas $\Phi(x; y)$ has a distal cell decomposition of exponent at most $2(|x|-1)+1 = 2|x|-1$.

If $\mathcal{M}$ is an o-minimal expansion of a group, we can instead set $d_0 = 2$, then we can set $r = 2$, and by [CGS20, Theorem 4.1], for $\Phi(x; y)$ with $|x| = 2$, Φ admits a distal cell decomposition of exponent 2. (In [CGS20], this is only proven for the case where $\mathcal{M}$ is an expansion of a field, but the proof only uses it for definable choice, which o-minimal expansions of groups also have.) Then for $|x| \geq 2$, $\Phi(x; y)$ admits a distal cell decomposition of exponent $2(|x| - 2) + 2 = 2|x| - 2$. $\square$

In the case of the ordered field $\mathbb{R}$, more is known. In that case, the distal cell decomposition produced in the above proof is the stratification in [CEG91]. An earlier version of that paper includes an improved bound for the case where $|x| = 3$, showing that $|\mathcal{T}(B)| = \mathcal{O}(|B|^3\beta(|B|)) = \mathcal{O}(|B|^{3+\varepsilon})$ for all $\varepsilon > 0$, where β is an extremely slowly growing function defined using the inverse of the Ackermann function.[CEG89] The argument uses Davenport-Schinzel sequences, purely combinatorial objects which lend themselves naturally to counting the complexity of cells defined by inequalities of a bounded family of functions. The lengths of Davenport-Schinzel sequences can be bounded in terms of the inverse Ackermann function, giving rise to the $\beta(|B|)$ term. For a general reference on such sequences, see [SA10]. These techniques are extended in [Kol04] to the case $|x| = 4$, where it is shown that $|\mathcal{T}(B)| = \mathcal{O}(|B|^{4+\varepsilon})$ for all $\varepsilon > 0$. These results imply that any finite set of formulas $\Phi(|x|; |y|)$ over $\mathbb{R}$ the ordered field has distal density 3 if $|x| = 3$, and $2|x| - 4$ if $|x| \geq 4$. It would be interesting to see if these bounds hold in any o-minimal structure, again using Davenport-Schinzel sequences. It seems possible that every $\Phi(x; y)$ in an o-minimal structure has distal density $|x|$, or admits a distal cell decomposition of exponent exactly $|x|$, although new tools would be required to prove such claims.

2.4.1 Locally Modular o-minimal Groups

The trichotomy theorem for o-minimal structures classifies them locally into three cases: trivial, ordered vector space over an ordered division ring, and expansion of a real closed field [PS98]. The o-minimal structures that are locally isomorphic to ordered vector spaces are known as the *linear* structures, and can also be classified as those satisfying the *CF property* [LP93]. Any such structure must extend the structure of either an ordered abelian group or an interval in an ordered abelian group. We will show that with the added assumption of local modularity, all finite families of formulas in o-minimal expansions of groups admit optimal distal cell decompositions. This includes the special case of any ordered vector space over an ordered division ring.

Theorem 2.4.2. *Let $\mathcal{M}$ be an o-minimal expansion of an ordered group, with $\mathrm{Th}(\mathcal{M})$ locally modular. Let $\Phi(x;y)$ be a finite set of formulas in the language of $\mathcal{M}$. Then Φ admits a distal cell decomposition of exponent $|x|$.*

To prove this theorem, we will need the following lemma:

Lemma 2.4.3. *Let $\mathcal{M}$ be an $\mathcal{L}$-structure.*

Let $\Phi(x;y)$ be a set of $\mathcal{L}$-formulas such that the negation of each $\varphi \in \Phi$ is a disjunction of other formulas in Φ. Assume that for any nonempty finite $B \subset M^{|y|}$ and $\varphi \in \Phi$, the conjunction $\bigwedge_{b\in B} \varphi(x;b)$ is equivalent to the formula $\varphi(x;b_0)$ for some $b_0 \in B$, or is not realizable. Then Φ admits a distal cell decomposition $\mathcal{T}$ such that for all finite B, the cells of $\mathcal{T}(B)$ are in bijection with the Φ-types $S^{\Phi}(B)$. In particular, the distal density of Φ equals the dual VC density of Φ.

Proof. Let Ψ be the set of all formulas of the form $\psi(x;(y_\varphi)_{\varphi\in\Phi}) := \bigwedge_{\varphi\in\Phi'} \varphi(x;y_\varphi)$, where $\Phi' \subset \Phi$ is arbitrary.

To define the distal cell decomposition $\mathcal{T}$, for each $\psi \in \Psi$, let $\theta_\psi(y; (y_\varphi)_{\varphi\in\Phi})$ denote

$$\bigvee_{\varphi\in\Phi} \exists x, (\varphi(x_1; y) \wedge \psi(x; (y_\varphi)_{\varphi\in\Phi}) \wedge \neg\exists x_2, (\varphi(x_2; y) \wedge \psi(x; (y_\varphi)_{\varphi\in\Phi}).$$

Then for a fixed finite $B \subset M^{|y|}$, and fixed $b_\varphi : \varphi \in \Phi$ in B, let Δ be the cell $\psi(M; (b_\varphi)_{\varphi\in\Phi})$. Then for $b \in B$, we see that $b \in I(\Delta)$ if and only if the cell defined by $\psi(x; (b_\varphi)_{\varphi\in\Phi})$ is crossed by $\varphi(x; b)$ for some $\varphi \in \Phi$.

We now claim that for any finite $B \subset M^{|y|}$, the cells of $\mathcal{T}(B)$ correspond exactly to the Φ-types $S^\Phi(B)$. As each cell Δ of $\mathcal{T}(B)$ is not crossed by $\Phi(B)$, its elements belong to a unique type of $S^\Phi(B)$. We claim that this type will be realized exactly by the elements of Δ. This type is equivalent to a single formula, which will be of the form $\bigwedge_{\varphi\in\Phi} \left(\bigwedge_{b\in B} \square_{\varphi,b}\varphi(x; b)\right)$, where each $\square_{\varphi,b}$ is either $\neg$ or nothing. For each φ, b such that $\square_{\varphi,b}$ is $\neg$, we may simply drop $\neg\varphi(x; b)$ from the conjunction, because $\neg\varphi(x; b)$ is equivalent to the disjunction $\bigvee_{\varphi\in\Phi_\varphi} \varphi(x; b)$ for some subset $\Phi_\varphi \subseteq \Phi$, and as the type is realizable, $\varphi_{i_0}(x; b)$ rather than its negation must already appear in the conjunction for some i_0, and we can replace $\varphi_{i_0}(x; b) \wedge \bigvee_{\varphi\in\Phi_\varphi} \varphi(x; b)$ with simply $\varphi_{i_0}(x; b)$. In this way, inductively, we can continue to remove all of the negated formulas in the conjunction, until we are left with $\bigwedge_{\varphi\in\Phi'} \left(\bigwedge_{b\in B_\varphi} \varphi(x; b)\right)$ where $\Phi' \subseteq \Phi$, and each $B_\varphi \subseteq B$ is nonempty. By our other assumption, as this formula is realizable, it is equivalent to $\bigwedge_{\varphi\in\Phi'} \square_{\varphi,b_\varphi}\varphi(x; b_\varphi)$ where each $b_\varphi \in B$, which in turn is a defining formula for a cell of $\mathcal{T}(B)$, which must be Δ. $\square$

Proof of Theorem 2.4.2. By o-minimality, we can assume the group is abelian. Let $\mathcal{L}_\mathcal{M}$ be the language of $\mathcal{M}$. Corollary 6.3 of [LP93] shows that $\mathcal{M}$ admits quantifier elimination in the language $\mathcal{L}'$, consisting of $+, <$, the set of algebraic points (that is, $\mathrm{acl}(\emptyset)$) as constants, and a unary function symbol for each 0-definable partial endomorphism of $\mathcal{M}$. Recall that a partial endomorphism is defined as a function of type either $f : M \to M$ or $f : (-c, c) \to M$ for some $c \in M$, such that if $a, b, a + b$ are all in the domain, then $f(a + b) = f(a) + f(b)$. The unary symbols representing the partial endomorphisms are assigned the value 0 outside the domain. If f has domain $(-c, c)$, then $c \in \mathrm{acl}(\emptyset)$. Note that by o-minimality, acl = dcl, so each of the constants in this language is in $\mathrm{dcl}(\emptyset)$, so each symbol of this language is $\emptyset$-definable in the original structure $(\mathcal{M}, \mathcal{L}_{\mathcal{M}})$.

Each formula in Φ is equivalent modulo $\mathrm{Th}(\mathcal{M})$ to some formula in $\mathcal{L}'$, so we replace Φ with $\Phi_{\mathcal{L}'}$, a pointwise equivalent finite set of $\mathcal{L}$-formulas. It suffices to find a distal cell decomposition of exponent $|x|$ for $\Phi_{\mathcal{L}'}$. As the interpretation of every symbol of $\mathcal{L}'$ is $\emptyset$-definable in $\mathcal{L}_{\mathcal{M}}$, we can replace each formula of this distal cell decomposition with an equivalent $\mathcal{L}_{\mathcal{M}}$-formula without parameters.

By quantifier elimination in $\mathcal{L}'$, we can find a finite set of atomic $\mathcal{L}'$-formulas Φ_A such that each formula in $\Phi_{\mathcal{L}'}$ is equivalent to a boolean combination of formulas in Φ_A modulo $\mathrm{Th}(\mathcal{M})$. Lemma 2.2.9 tells us that a distal cell decomposition for Φ_A is a distal cell decomposition for $\Phi_{\mathcal{L}'}$, so it suffices to prove the desired result for Φ_A. We then will find another finite set of $\mathcal{L}'$-formulas, Φ', such that each atomic formula in Φ_A is a boolean combination of formulas in Φ', and Φ' satisfies the conditions of the following lemma, providing us with a distal cell decomposition that we can show has the desired exponent. It suffices to find Φ' satisfying the requirements of Lemma 2.4.3 such that any atomic formula in Φ_A is a boolean combination of formulas from Φ', and to show that for any finite Φ and B, $\left|S^{\Phi}(B)\right| \leq \mathcal{O}(|B|^{|x|})$.

We will select Φ' to contain only atomic $\mathcal{L}'$-formulas of the form $f(x)+g(y)+c\square 0$, where f, g are group endomorphisms, c is a term built only out of functions and constants, and $\square \in \{<, =, >\}$. If $\varphi(x; y)$ is of the form $f(x)+g(y)+c = 0$, then for a given B, $\bigwedge_{b \in B} \varphi(x; b)$ is either equivalent to $\varphi(x; b)$ for all $b \in B$ or not realizable. If φ is an inequality, then $\bigwedge_{b \in B} \varphi(x; b)$ is equivalent to $\varphi(x; b_0)$ for some b_0 minimizing or maximizing $g(b)$. Also, for all $\varphi \in \Phi'$, $\neg\varphi(x; y)$ is a disjunction of other formulas in Φ', because $\neg f(x) + g(y) + c = 0$ is equivalent to $f(x) + g(y) + c < 0 \vee f(x) + g(y) + c > 0$, $\neg f(x) + g(y) + c < 0$ is equivalent to $f(x) + g(y) + c = 0 \vee f(x) + g(y) + c > 0$, and $\neg f(x) + g(y) + c > 0$ is equivalent to $f(x) + g(y) + c = 0 \vee f(x) + g(y) + c < 0$.

Now we show that every atomic $\mathcal{L}'$-formula, and thus every formula in Φ_A, can be expressed as a boolean combination of atomic formulas of the form $f(x)+g(y)+c\square 0$ with f and g total (multivariate) definable endomorphisms. Any atomic formula is of the form $f(x;y)\square g(x;y)$, and by subtraction is equivalent to $(f-g)(x;y)\square 0$. Thus it suffices to show that for any $\mathcal{L}'$-term $t(x;y)$ and $\square \in \{<,=,>\}$, the atomic formula $t(x;y)\square 0$ is equivalent to a boolean combination of formulas of the form $f(x)+g(y)+c\square' 0$ with f and g total endomorphisms and $\square' \in \{<,=,>\}$.

We prove this by induction on the number of partial endomorphism symbols in $t(x;y)$ that do not represent total endomorphisms. If that number is 0, then every symbol in the term $t(x;y)$ is a variable, a constant, or represents a total endomorphism. Thus $t(x;y)$ is a composition of affine functions, and is thus itself an affine function, which can be represented as $f(x)+g(y)+c$. Thus $t(x;y)\square 0$ is equivalent to $f(x)+g(y)+c\square 0$. Now let $t(x;y)$ contain $n+1$ partial endomorphism symbols. Let one of them be f, so that $t(x;y)=t_1(f(t_2(x;y)),x,y)$ for some terms t_1,t_2. By [LP93, Lemma 4.3] and local modularity, $\mathcal{L}'$ contains a partial endomorphism symbol g representing a total function such that $f(x)=g(x)$ on the interval $(-c,c)$, with $f(x)=0$ outside of that interval. Thus $t(x;y)\square 0$ is equivalent to

$$\begin{aligned}&(-c<t_2(x;y)\wedge t_2(x;y)<c\wedge t_1(g(t_2(x;y)),x,y)\square 0)\\ &\quad\vee(\neg(-c<t_2(x;y)\wedge t_2(x;y)<c)\wedge t_1(0,x,y)\square 0)\,.\end{aligned}$$

This is equivalent to a boolean combination of the formulas $t_2(x;y)+c>0$, $t_2(x;y)-c<0$, $t_1(g(t_2(x;y)),x,y))\square 0$, and $t_1(0,x,y)\square 0$, each of which has at most n non-total partial endomorphisms, and thus by induction, is a boolean combination of formulas of the desired form.

Now we wish to verify that $\left|S^{\Phi}(B)\right|\leq\mathcal{O}(|B|^{|x|})$. Theorem 6.1 of [ADH16] says that the dual VC density of Φ will be at most $|x|$, which is only enough to show that Φ has distal density $|x|$. However, the proof shows that $\left|S^{\Phi}(B)\right|\leq\mathcal{O}(|B|^{|x|})$. Tracing the logic of that paper, Theorem 6.1 guarantees that a weakly o-minimal theory has the VC1 property, which by Corollary 5.9 implies that Φ has uniform definition of $\Phi(x;B)$ types over finite sets with $|x|$ parameters, which implies that $\left|S^{\Phi}(B)\right| \leq \mathcal{O}(|B|^{|x|})$ (as noted at the end of Section 5.1).

$\square$

2.5 Presburger Arithmetic

Presburger arithmetic is the theory of $\mathbb{Z}$ as an ordered group. As mentioned in Example 2.9 of [CS18], the ordered group $\mathbb{Z}$ admits quantifier elimination in the language $\mathcal{L}_{\text{Pres}} = \{0, 1, +, -, <, \{k\mid\}_{k\in\mathbb{N}}\}$, where for each $k \in \mathbb{N}$ and $x \in \mathbb{Z}$, $\mathbb{Z} \models k \mid x$ when x is divisible by k, so we will work in this language. As this structure is quasi-o-minimal, it is distal, and we will construct an explicit distal cell decomposition with optimal bounds, similar to the distal cell decomposition for o-minimal expansions of locally modular ordered groups in Theorem 2.4.2.

Theorem 2.5.1. *Let G be an ordered abelian group with quantifier elimination in $\mathcal{L}_{\text{Pres}}$. Let $\Phi(x; y)$ be a finite set of formulas in this language. Then Φ has distal density at most $|x|$.*

Proof. Throughout this proof, we will identify $\mathbb{Z}$ with the subgroup of G generated by the constant 1.

As G has quantifier elimination in this language, every $\varphi(x; y) \in \Phi$ is equivalent to a boolean combination of atomic formulas. We will group the atomic formulas into two categories. The first is those of the form $f(x)\square g(y) + c$, where $\square \in \{<, =, >\}$, (f, g) belongs to a finite set F of pairs of $\mathbb{Z}$-linear functions of the form $\sum_{i=1}^{|x|} a_i x_i$ with $a_i \in \mathbb{Z}$, and c belongs to a finite set $C \subseteq \mathbb{Z}$. The second is atomic formulas of the form $k \mid (f(x)+g(y)+c)$ for $k \in \mathbb{N}$, $(f, g) \in F$, and $c \in C$. Furthermore, we may assume that only one symbol of the form $k\mid$ is used. If K is the least common multiple of the finite collection of k such that $k\mid$ appears in one of these atomic formulas, then each $k \mid (f(x)+g(y)+c)$ can be replaced with $K \mid (d \cdot f(x) + d \cdot g(y) + d \cdot c)$, where $d \cdot \sum_{i=1}^{|x|} a_i x_i = \sum_{i=1}^{|x|} (d \cdot a_i) x_i$ and $dk = K$. Note that all of these functions and constants are $\emptyset$-definable.

Then, by Lemma 2.2.9, we may replace $\Phi(x;y)$ with the union of the following two sets of atomic formulas for appropriate choices of F and C:

- Fix C to be a finite subset of $\mathbb{Z}$, F a finite subset of pairs of $\mathbb{Z}$-linear functions of the form $\sum_{i=1}^{|x|} a_i x_i$, and $K \in \mathbb{N}$.
- Let Φ_0 be the set of all $f(x)\Box g(y)+c$ with $(f,g) \in F, c \in C, \Box \in \{<,=,>\}$.
- Let Φ_1 be the set of all $K \mid (f(x)+g(y)+c)$ with $(f,g) \in F, c \in \{0,\ldots,K-1\}$.
- Let $\Phi = \Phi_0 \cup \Phi_1$.

It is straightforward to see that the negation of any formula from Φ_0 is equivalent to the disjunction of two formulas from Φ_0, and a negation of any formula $K|(f(x)+g(y)+c)$ from Φ_1 is equivalent to $\bigvee_{0 \le c' < K, c' \neq c} K|(f(x)+g(y)+c)$, a disjunction of formulas from Φ_1.

To apply Lemma 2.4.3, it suffices to show that for any $\varphi \in \Phi$ and nonempty finite $B \subset M^{|y|}$, $\bigwedge_{b \in B} \varphi(x;b)$ is equivalent to $\varphi(x;b_0)$ for some $b_0 \in B$ or is not realizable. This holds for $\varphi \in \Phi_0$ for reasons discussed in the proof of 2.4.2. For $\varphi \in \Phi_1$, we see that if there exist b_1, b_2 such that $g(b_1) \not\equiv g(b_2) \pmod{K}$, then $\varphi(x;b_1) \wedge \varphi(x;b_2)$ implies $K|(f(x)+g(b_1)+c) \wedge K|(f(x)+g(b_2)+c)$ so $K|\,(g(b_1)-g(b_2))$, a contradiction. Thus this conjunction is not realizable. Otherwise, for any $b_0 \in B$, and any other $b \in B$, $g(b) \equiv g(b_0) \pmod{K}$, so $\bigwedge_{b \in B} \varphi(x;b)$ is equivalent to $\varphi(x;b_0)$.

Now Lemma 2.4.3 gives us a distal cell decomposition $\mathcal{T}$ for Φ, such that for all B, $|T(B)| = \left|S^{\Phi}(B)\right|$. The theory of $\mathbb{Z}$ in $\mathcal{L}_{\text{Pres}}$ is quasi-o-minimal by [BPW00, Example 2], and the same argument will hold for G, because G has quantifier elimination in the same language. The same VC density results apply to quasi-o-minimal theories as to o-minimal theories (see [ADH16, Theorem 6.4]), so $\left|S^{\Phi}(B)\right| \le \mathcal{O}(|B|^{|x|})$. $\square$

2.6 $\mathbb{Q}_p$, the linear reduct

Now we turn our attention to the linear reduct of $\mathbb{Q}_p$, viewed as a structure $\mathcal{M}$ in the language $\mathcal{L}_{\text{aff}} = \{0, +, -, \{c\cdot\}_{c\in\mathbb{Q}_p}, |, \{Q_{m,n}\}_{m,n\in\mathbb{N}\setminus\{0\}}\}$, where $c\cdot$ is a unary function symbol which acts as scalar multiplication by c, $x \mid y$ stands for $v(x) \leq v(y)$, and $\mathcal{M} \models Q_{m,n}(a)$ if and only if $a \in \bigcup_{k\in\mathbb{Z}} p^{km}(1 + p^n\mathbb{Z}_p)$. For each m, n, the set $Q_{m,n}(M) \setminus \{0\}$ is a subgroup of the multiplicative group of $\mathbb{Q}_p$ with finite index. Leenknegt [Lee12, Lee14] introduced this structure (referring to the language as $\mathcal{L}_{\text{aff}}^{\mathbb{Q}_p}$), proved that it is a reduct of Macintyre's standard structure on $\mathbb{Q}_p$, and proved cell decomposition results for it which imply quantifier elimination.

Bobkov [Bob17] shows that every finite set $\Phi(x; y)$ of formulas has dual VC density $\leq |x|$, and this section is devoted to strengthening this by proving the same optimal bound for the distal density:

Theorem 2.6.1. *For any finite set $\Phi(x; y)$ of $\mathcal{L}_{\text{aff}}$-formulas in $\mathbb{Q}_p$, there is a distal cell decomposition $\mathcal{T}$ with $|\mathcal{T}(B)| = \mathcal{O}(|B|^{|x|})$, so Φ has distal density $\leq |x|$.*

It is worth noting that Bobkov used a slightly different version of this language, which included the constant 1, therefore making all definable sets $\emptyset$-definable. Because our distal cell decomposition must be definable without parameters, we will use slightly stronger versions of Leenknegt and Bobkov's basic lemmas, to avoid parameters. The first such result is a cell-decomposition result, proven in [Lee12], but stated most conveniently as [Bob17, Theorem 4.1.5]. To state it, we need to define what a cell is in that context:

Definition 2.6.2. A 0-*cell* is the singleton $\mathbb{Q}_p^0$. A $(k+1)$-*cell* is a subset of $\mathbb{Q}_p^{k+1}$ of the following form:

$$\{(x, t) \in D \times \mathbb{Q}_p \,|\, v(a_1(x)) \,\square_1\, v(t - c(x)) \,\square_2\, v(a_2(x)),\ t - c(x) \in \lambda Q_{m,n}\},$$

where D is a k-cell, a_1, a_2, c are polynomials of degree ≤ 1, called the *defining polynomials*, each of $\Box_1, \Box_2$ is either $<$ or no condition, $m, n \in \mathbb{N}$, and $\lambda \in \mathbb{Q}_p$.

Fact 2.6.3 ([Lee12], see also [Bob17, Theorem 4.1.5])**.** *Any definable subset of $\mathbb{Q}_p^k$ (in the language $\mathcal{L}_{\text{aff}}$) decomposes into a finite disjoint union of k-cells.*

Now we modify these definitions and results to work in an $\emptyset$-definable context:

Definition 2.6.4. A 0-*cell over* $\emptyset$ is just a 0-cell. A $(k+1)$-*cell over* $\emptyset$ is a $(k+1)$-cell $\{(x,t) \in D \times \mathbb{Q}_p | v(a_1(x)) \Box_1 v(t - c(x)) \Box_2 v(a_2(x)), t - c(x) \in \lambda Q_{m,n}\}$ where D is a k-cell over $\emptyset$ and the defining polynomials have constant coefficient 0.

We can now state a $\emptyset$-definable version of the cell decomposition result:

Lemma 2.6.5. *Any $\emptyset$-definable subset of $\mathbb{Q}_p^k$ (in the language $\mathcal{L}_{\text{aff}}$) decomposes into a finite disjoint union of k-cells over $\emptyset$.*

Proof. We trace the proof of the original cell decomposition result in [Lee12]. Lemmas 2.3 and 2.7 establish that finite unions of cells (in the case of finite residue field, equivalent to the "semi-additive sets" of Definition 2.6) are closed under intersections and projections respectively, and Lemma 2.5 (using Lemma 2.4) shows that all quantifier-free definable sets are semi-additive. It suffices to modify each of these four lemmas slightly. In all four lemmas, we modify the assumptions to require that all linear polynomials in the assumptions have constant term 0. In each construction, the polynomials in the results are linear combinations of the polynomials in the assumptions, and thus will also have constant term 0, allowing us to state the results in terms of k-cells *over* $\emptyset$. $\Box$

This tells us that no nonzero constants are definable:

Lemma 2.6.6. *In the structure $\mathcal{M}$ consisting of $\mathbb{Q}_p$ in the language $\mathcal{L}_{\text{aff}}$,* $\operatorname{dcl}(\emptyset) = \{0\}$.

Proof. If $a \in \operatorname{dcl}(\emptyset)$, then $\{a\}$ is $\emptyset$-definable, so it can be decomposed into 1-cells over $\emptyset$. There can only be one cell in the decomposition, $\{a\}$. All of its defining polynomials take in variables from the unique 0-cell, and thus consist only of their constant coefficient, which is 0. Thus the cell must be of the form $\{a\} = \{t \in D \times \mathbb{Q}_p | v(0)\square_1 v(t-0)\square_2 v(0), t-0 \in \lambda Q_{m,n}\}$. The condition $v(0)\square_1 v(t)\square_2 v(0)$ will define one of the following sets: $\emptyset, \{0\}, \mathbb{Q}_p \setminus \{0\}, \mathbb{Q}_p$, and the condition $t \in \lambda Q_{m,n}$ defines $\{0\}$ when $\lambda = 0$, and otherwise, $\lambda Q_{m,n} \subseteq \mathbb{Q}_p \setminus \{0\}$. Thus the whole cell is either $\{0\}$ or $\lambda Q_{m,n}$ which is infinite, so if it is a singleton $\{a\}$, we must have $a = 0$. $\square$

We now check that our cell decomposition for $\emptyset$-definable sets yields $\emptyset$-definable cells:

Lemma 2.6.7. *Any k-cell over $\emptyset$ is $\emptyset$-definable.*

Proof. We prove this by induction on k. The $k = 0$ case is trivial. The $(k+1)$-cell $\{(x,t) \in D \times \mathbb{Q}_p | v(a_1(x))\square_1 v(t - c(x))\square_2 v(a_2(x)), t - c(x) \in \lambda Q_{m,n}\}$ is $\emptyset$-definable if D is, $v(a_1(x))\square_1 v(t - c(x))\square_2 v(a_2(x))$ is, and $t - c(x) \in \lambda Q_{m,n}$ is. We have that D is by the induction hypothesis. For the next condition, it suffices to observe that the defining polynomials are $\emptyset$-definable functions if and only if they have constant coefficient 0, because scalar multiplication is $\emptyset$-definable, but no constant other than 0 is. For the final condition, we see that if $\lambda = 0$, then $t - c(x) \in \lambda Q_{m,n}$ is equivalent to $t - c(x) = 0$, which is $\emptyset$-definable, and if $\lambda \neq 0$, then $t - c(x) \in \lambda Q_{m,n}$ is equivalent to $\lambda^{-1} \cdot (t - c(x)) \in Q_{m,n}$, which is $\emptyset$-definable. $\square$

We now want to generalize the following quantifier-elimination result to the $\emptyset$-definable case:

Lemma ([Bob17, Theorem 4.2.1])**.** *Any $\mathcal{L}_{\text{aff}}$-formula (with parameters) $\phi(x; y)$ where x and y are finite tuples of variables is equivalent in the $\mathcal{L}_{\text{aff}}$-structure $\mathbb{Q}_p$ to a boolean combination of formulas from a collection*

$$\Phi_\phi = \{v(p_i(x) - c_i(y)) < v(p_j(x) - c_j(y))\}_{i,j \in I} \cup \{p_i(x) - c_i(y) \in \lambda Q_{m,n}\}_{i \in I, \lambda \in \Lambda}$$

where $I = \{1, \ldots, |I|\}$ is a finite index set, each p_i is a degree ≤ 1 polynomial with constant *term 0, each c_i is a degree ≤ 1 polynomial, and Λ is a finite set of coset representatives of $Q_{m,n}$ for some $m, n \in \mathbb{N}$.*

Bobkov derives this result from the cell decomposition. If we apply the same logic to the $\emptyset$-definable cell decomposition from Lemma 2.6.5, then all of the polynomials involved have constant term 0, and thus all formulas involved are $\emptyset$-definable:

Lemma 2.6.8. *Any $\mathcal{L}_{\mathrm{aff}}$-formula $\phi(x; y)$ where x and y are finite tuples of variables is equivalent in the $\mathcal{L}_{\mathrm{aff}}$-structure $\mathbb{Q}_p$ to a boolean combination of formulas from a collection*

$$\Phi_\phi = \{v(p_i(x) - c_i(y)) < v(p_j(x) - c_j(y))\}_{i,j \in I} \cup \{p_i(x) - c_i(y) \in \lambda Q_{m,n}\}_{i \in I, \lambda \in \Lambda}$$

where $I = \{1, \ldots, |I|\}$ is a finite index set, each p_i and each c_i is a degree ≤ 1 polynomial with constant term 0 and Λ is a finite set of coset representatives of $Q_{m,n}$ for some $m, n \in \mathbb{N}$.

As a corollary of this lemma and Lemma 2.2.9, we see that we can replace Φ with the set $\bigcup_\phi \Phi_\phi$, and thus assume that Φ takes the form

$$\{v(p_i(x) - c_i(y)) < v(p_j(x) - c_j(y))\}_{i,j \in I} \cup \{p_i(x) - c_i(y) \in \lambda Q_{m,n}\}_{i \in I, \lambda \in \Lambda}.$$

for some fixed $m, n \in \mathbb{N}$.

We now recall some terminology from Bobkov [Bob17].

Definition 2.6.9 ([Bob17], Def. 4.2.3)**.** For the rest of this section, we fix $B \subset M^{|y|}$, and let $T = \{c_i(b) : i \in I, b \in B\}$.

- For $c \in \mathbb{Q}_p$ and $r \in \mathbb{Z}$, we define $B_r(c) := \{x : v(x - c) > r\}$ and refer to it as the *open* ball of radius r around c.
- Let the *subintervals* over a parameter set B be the atoms in the Boolean algebra

generated by the set $\mathcal{B}$ of balls

$$\{B_{v(c_i(b_1)-c_j(b_2))}(c_i(b_1)) : i,j \in I, b_1, b_2 \in B\} \cup \{B_{v(c_j(b)-c_k(b))}(c_i(b)) : i,j,k \in I, b \in B\}$$

- Each subinterval can be expressed as $I(t, \alpha_L, \alpha_U)$ where

$$I(t, \alpha_L, \alpha_U) = B_{\alpha_L}(t) \setminus \bigcup_{t' \in T \cap B_{\alpha_U - 1}(t)} B_{\alpha_U}(t'),$$

  for some $t = c_i(b_0)$ with $i \in I, b_0 \in B$, and $\alpha_L = \alpha_1(b_0, b_1), \alpha_U = \alpha_2(b_0, b_2)$, with α_1, α_2 chosen from a finite set A of $\emptyset$-definable functions $\mathbb{Q}_p^2 \to \Gamma$, including two functions defined, by abuse of notation, as $\pm\infty$.

- The subinterval $I(t, \alpha_L, \alpha_U)$ is said to be *centered* at t.

By this definition, it is not clear that $I(t, \alpha_L, \alpha_U)$ should be uniformly definable from parameters in B, as the set $T \cap B_{\alpha_U - 1}(t)$ could depend on all of B. However, we can eliminate most of the balls from that definition. The ball $B_{\alpha_U-1}(t)$ can be split into p balls of the form $B_{\alpha_U}(t')$ for some $t' \in \mathbb{Q}_p$, call them $B_1, \ldots, B_p$. Let T' be a subset of $T \cap B_{\alpha_U - 1}(t)$ such that for each B_i, if $T \cap B_i \neq \emptyset$, then T' contains only a single representative t_i from B_i. Then

$$\bigcup_{t' \in T \cap B_{\alpha_U - 1}(t)} B_{\alpha_U}(t') = \bigcup_{t' \in T'} B_{\alpha_U}(t'),$$

because each $t' \in T \cap B_{\alpha_U-1}(t)$ belongs to some B_i, so $B_{\alpha_U}(t') = B_i = B_{\alpha_U}(t_i)$. We may assume $|T'|$ to be at most $p-1$, because if all p balls were removed, we could instead define this set as $I(t, \alpha_L, \alpha_U - 1)$. Thus each subinterval can be defined as $I(t, \alpha_L, \alpha_U) = \psi_{\text{sub}}(t, \alpha_L, \alpha_U, \bar{b})$, where ψ_{sub} is one of a finite collection Ψ_{sub} of formulas, and $\bar{b}$ is a tuple of at most $p-1$ elements of B.

Definition 2.6.10 ([Bob17], Def. 4.2.5)**.** For $a \in \mathbb{Q}_p$, define $T{-}\mathrm{val}(a) := v(a-t)$, where a belongs to a subinterval centered at t. By Lemma 4.2.6, [Bob17], this is well-defined, as

$v(a-t)$ is the same for all valid choices of t.

Definition 2.6.11 ([Bob17], Def. 4.2.8)**.** Given a subinterval $I(t, \alpha_L, \alpha_U)$, two points a_1, a_2 in that subinterval are defined to have the same *subinterval type* if one of the following conditions is satisfied:

- $\alpha_L + n \leq T-\text{val}(a_i) \leq \alpha_U - n$ for $i = 1, 2$ and $(a_1 - t)(a_2 - t)^{-1} \in Q_{m,n}$,
- $\neg(\alpha_L + n \leq T-\text{val}(a_i) \leq \alpha_U - n)$ for $i = 1, 2$ and $T-\text{val}(a_1) = T-\text{val}(a_2) \leq v(a_1 - a_2) - n$.

We show that the set of points of each subinterval type is definable over t, α_L, α_U. The subinterval types of the first kind are definable by

$$\psi_{\text{tp}}^{\lambda}(x; t, \alpha_L, \alpha_U) := (\alpha_L + n \leq v(x-t) \leq \alpha_U - n) \wedge (x - t) \in \lambda Q_{m,n}$$

where $\lambda \in \Lambda$. The subinterval types of the second kind are definable by one of

$$\psi_{\text{tp}}^{L,i,q}(x; t, \alpha_L, \alpha_U) := (v(x-t) = \alpha_L + i) \wedge (\alpha_L + i + n \leq v(x - (p^{\alpha_L + i} q + t)))$$

or

$$\psi_{\text{tp}}^{U,i,q}(x; t, \alpha_L, \alpha_U) := (v(x-t) = \alpha_U - i) \wedge (\alpha_U - i + n \leq v(x - (p^{\alpha_U - i} q + t))),$$

where $0 \leq i < n$, and q ranges over a set Q of representatives of the balls of radius n contained in $B_0(0) \setminus B_1(0)$. If we let α be $\alpha_L + i$ or $\alpha_U - i$, this makes $p^{\alpha} q + t$ range over a finite set of representatives of the balls of radius $\alpha + n$ contained in the set $B_{\alpha}(t) \setminus B_{\alpha+1}(t)$ of points a with $v(a-t) = \alpha$. Let Ψ_{tp} be the set of all these formulas: $\{\psi_{\text{tp}}^{\lambda} : \lambda \in \Lambda\} \cup \{\psi_{\text{tp}}^{L,i,q} : 0 \leq i < n, q \in Q\} \cup \{\psi_{\text{tp}}^{U,i,q} : 0 \leq i < n, q \in Q\}$.

2.6.1 Defining the Distal Cell Decomposition

We start by defining $\Psi(x; (y_{0,i} : i \in I), (y_{1,i} : i \in I), (y_{2,i} : i \in I))$ to be the set of all formulas $\psi(x; (y_{0,i} : i \in I), (y_{1,i} : i \in I), (y_{2,i} : i \in I))$ of the form

$$\left(\bigwedge_{i \in I} \left[\psi^i_{\mathrm{sub}}(p_i(x), t_i, \alpha_{L,i}, \alpha_{U,i}, \bar{y}_i) \wedge \psi^i_{\mathrm{tp}}(p_i(x), t_i, \alpha_{L,i}, \alpha_{U,i})\right]\right) \wedge \psi_\sigma(x; t_1, \dots, t_{|I|})$$

where $\psi^i_{\mathrm{sub}} \in \Psi_{\mathrm{sub}}$, $\psi^i_{\mathrm{tp}} \in \Psi_{\mathrm{tp}}$, $\psi_\sigma(x, t_1, \dots, t_{|I|})$ is, for some permutation σ of I,

$$v(p_{\sigma(1)}(x) - t_{\sigma(1)}) > \dots > v(p_{\sigma(|I|)}(x) - t_{\sigma(|I|)}),$$

and we define $t_i, \alpha_{L,i}, \alpha_{U,i}$ so that $t_i = c_j(y_{0,i})$ for some $j \in I$, $\alpha_{L,i} = \alpha_1(y_{0,i}, y_{1,i})$, and $\alpha_{L,i} = \alpha_2(y_{0,i}, y_{2,i})$ for some $\alpha_1, \alpha_2 \in A$.

For each potential cell Δ, we will define $\mathcal{I}(\Delta)$ so that Δ will be included in $\mathcal{T}(B)$ exactly when each set $\psi^i_{\mathrm{sub}}(M, t_i, \alpha_{L,i}, \alpha_{U,i}, \bar{b}_i)$ is actually a subinterval. Then each cell of $\mathcal{T}(B)$ will consist of all elements $a \in M^{|x|}$ such that for all i, $p_i(a)$ belongs to a particular subinterval and has a particular subinterval type, and the set $\{T - \mathrm{val}(p_i(a)) : i \in I\}$ has a particular ordering. These cells are not crossed by $\Phi(x; B)$, as a consequence of the following lemma:

Lemma 2.6.12 ([Bob17, Lemma 4.2.12])**.** *Suppose $d, d' \in \mathbb{Q}_p$ satisfy the following three conditions:*

- *For all $i \in I$, $p_i(d)$ and $p_i(d')$ are in the same subinterval.*
- *For all $i \in I$, $p_i(d)$ and $p_i(d')$ have the same subinterval type.*
- *For all $i, j \in I$, $T - \mathrm{val}(p_i(d)) > T - \mathrm{val}(p_j(d))$ iff $T - \mathrm{val}(p_i(d')) > T - \mathrm{val}(p_j(d'))$.*

  Then d, d' have the same Φ-type over B.

Now we check that we can actually define $\mathcal{I}(\Delta)$ as desired. For some $\psi_{\mathrm{sub}}(x, t, \alpha_L, \alpha_U, \bar{b})$ to be a subinterval, we must check that it actually equals $I(t, \alpha_L, \alpha_U)$, and that that set is not crossed by any other balls in $\mathcal{B}$. If $\bar{b} = (b_1, \ldots, b_{p-1})$, then there are $j_1, \ldots, j_{p-1} \in I$ with this set equal to $B_{\alpha_L}(t) \setminus \bigcup_{k=1}^{p-1} B_{\alpha_U}(c_{j_k}(b_k))$. This is actually $I(t, \alpha_L, \alpha_U)$ as long as there is no $i \in I$, $b \in B$ with $v(c_i(b) - t) = \alpha_U$, but $c_i(b) \notin \bigcup_{k=1}^{p-1} B_{\alpha_U}(c_{j_k}(b_k))$. The only way for this to happen is if $v(c_i(b) - c_{j_k}(b_k)) = \alpha_U$ for all $1 \leq k < p$, so let $\mathcal{I}_1(\Delta)$ be the set of all $b \in B$ where this happens.

For $\Delta = I(t, \alpha_L, \alpha_U)$ to not be a subinterval, it must be crossed by some ball $B_\alpha(t^*) \in \mathcal{B}$. Such a ball crosses $I(t, \alpha_L, \alpha_U)$ if and only if $t^* \in B_{\alpha_L}(t)$, $\alpha_L < \alpha < \alpha_U$, and

$$B_\alpha(t^*) \setminus \bigcup_{t' \in T \cap B_{\alpha_U - 1}(t)} B_{\alpha_U}(t) \neq \emptyset.$$

This last condition follows from the previous two, as

$$\bigcup_{t' \in T \cap B_{\alpha_U - 1}(t)} B_{\alpha_U}(t) \subsetneq B_{\alpha_U - 1}(t),$$

and if $\alpha < \alpha_U$, then either $B_{\alpha_U - 1}(t) \subset B_\alpha(t^*)$ or they are disjoint. The radius α can either be $v(c_j(b) - c_k(b))$, where $t^* = c_i(b)$, for some $i, j, k \in I$, or $v(t' - t^*)$ for some $t' \in T$. Let $\mathcal{I}_2(\Delta)$ be the set of all b such that for some $i, j, k \in I$, $\alpha_L < v(c_j(b) - c_k(b)) < \alpha_U$ and $\alpha_L < v(c_i(b) - t)$. This handles the former case. In the latter case, where $\alpha = v(t' - t^*)$, we see that as $\alpha_L < \alpha$, $t' \in B_{\alpha_L}(t^*) = B_{\alpha_L}(t)$, so $\alpha_L < v(t - t')$. Also, $\min\{v(t - t'), v(t - t^*)\} \leq v(t' - t^*) < \alpha_U$, so either the ball $B_{v(t-t')}(t)$ or $B_{v(t-t^*)}(t)$ has radius between α_L and α_U, and thus crosses Δ. Thus Δ is crossed by a ball of the form $B_{v(t'-t^*)}(t^*)$ if and only if it is crossed by a ball of the form $B_{v(t-t')}(t')$ if and only if there is some $t' \in T$ with $\alpha_L < v(t - t') < \alpha_U$, so we let $\mathcal{I}_3(\Delta)$ be the set of all b such that there exists $i \in I$ with $\alpha_L < v(t - c_i(b)) < \alpha_U$.

Then if we let $\mathcal{I}(\Delta) = \mathcal{I}_1(\Delta) \cup \mathcal{I}_2(\Delta) \cup \mathcal{I}_3(\Delta)$, which is uniformly definable from just the parameters used to define Δ, then Δ is a subinterval if and only if $B \cap \mathcal{I}(\Delta) = \emptyset$, as desired.

2.6.2 Counting the Distal Cell Decomposition

To calculate the distal density of Φ, we will count the number of cells of $\mathcal{T}(B)$ by following Bobkov's estimate of $\left|S^{\Phi}(B)\right|$. Because our cells are defined less in terms of x itself than the values $p_i(x)$, we define a function to shift our problem to study those values directly:

Definition 2.6.13 ([Bob17, Def. 4.3.4])**.** Let $f : \mathbb{Q}_p^{|x|} \to \mathbb{Q}_p^I$ be $(p_i(x))_{i\in I}$. Define the segment set Sg to be the image $f(\mathbb{Q}_p^{|x|})$.

We will need a notation for recording certain coefficients of elements of $\mathbb{Q}_p$:

Definition 2.6.14 ([Bob17, Def. 4.2.9])**.** For $c \in \mathbb{Q}_p$, $\alpha < \beta \in v(\mathbb{Q}_p)$, c can be expressed uniquely as $\sum_{\gamma\in v(\mathbb{Q}_p)} c_\gamma p^\gamma$ with $c_\gamma \in \{0, 1, \ldots, p-1\}$. Then define $c \upharpoonright [\alpha, \beta)$ to be the tuple $(c_\alpha, c_{\alpha+1}, \ldots, c_{\beta-1}) \in \{0, 1, \ldots, p-1\}^{\beta-\alpha}$.

This coefficient function $\upharpoonright$ will be useful in allowing us to reduce the information of $\{a_i : i \in I\} \in \mathbb{Q}_p^I$ to a linearly independent subset together with a finite number of coordinates, using this lemma:

Lemma 2.6.15 ([Bob17, Cor. 4.3.2])**.** *Suppose we have a finite collection of vectors $\{\vec{p}_i\}_{i\in I}$ with each $\vec{p}_i \in \mathbb{Q}_p^{|x|}$. Suppose $J \subseteq I$ and $i \in I$ satisfy $\vec{p}_i \in \operatorname{span}\{\vec{p}_j\}_{j\in J}$, and we have $\vec{c} \in \mathbb{Q}_p, \alpha \in v(\mathbb{Q}_p)$ with $v(\vec{p}_j \cdot \vec{c}) > \alpha$ for all $j \in J$. Then $v(\vec{p}_i \cdot \vec{c}) > \alpha - \gamma$ for some $\gamma \in v(\mathbb{Q}_p), \gamma \geq 0$. Moreover γ can be chosen independently from $J, j, \vec{c}, \alpha$ depending only on $\{\vec{p}_i\}_{i\in I}$.*

As each homogeneous linear polynomial $p_i(x)$ can be written as the dot product $\vec{p}_i \cdot x$ for some $\vec{p}_i \in \mathbb{Q}_p^{|x|}$, let $\gamma \in v(\mathbb{Q}_p)_{\geq 0}$ satisfy the criteria of Lemma 2.6.15 for $\{\vec{p}_i\}_{i\in I}$.

Definition 2.6.16 ([Bob17, Def. 4.3.3])**.** Any $a \in \mathbb{Q}_p$ belongs to a unique subinterval $I(t, \alpha_L, \alpha_U)$. Define $T - \mathrm{fl}(a) := \alpha_L$.

Using this function, we partition Sg into $(2|I|)!$ pieces, corresponding to the possible order types of $\{T - \mathrm{fl}(x_i) : i \in I\} \cup \{T - \mathrm{val}(x_i) : i \in I\}$. We will show that each piece of this partition intersects only $\mathcal{O}(|B|^{|x|})$ cells of $\mathcal{T}(B)$.

Let Sg' be a piece of the partition. Relabel the functions p_i such that

$$T - \mathrm{fl}(a_1) \geq \cdots \geq T - \mathrm{fl}(a_{|I|})$$

for all $(a_i)_{i \in I} \in \mathrm{Sg}'$. Using a greedy algorithm, find $J \subseteq I$ such that $\{\vec{p}_j\}_{j \in J}$, with the new labelling, is linearly independent, and for each $i \in I$, $\vec{p}_i$ is a linear combination of $\{\vec{p}_j\}_{j \in J, j < i}$.

Definition 2.6.17.

- Denote $\{0, \ldots, p-1\}^\gamma$ as Ct.
- Let Tp be the set of all subinterval types. Lemma 4.2.11 from [Bob17] shows that $|\mathrm{Tp}| \leq K$, where K is a constant that does not depend on B.
- Let Sub be the set of all subintervals. Lemma 4.2.4 from [Bob17] tells us that $|\mathrm{Sub}| = O(|B|)$.

Now we can define a function identifying subintervals, subinterval types, and γ many coefficients of the components of each element of Sg':

Definition 2.6.18. Define $g : \mathrm{Sg}' \to \mathrm{Tp}^I \times \mathrm{Sub}^J \times \mathrm{Ct}^{I \setminus J}$ as follows:

Let $a = (a_i)_{i \in I} \in \mathrm{Sg}'$.

For each $i \in I$, record the subinterval type of a_i to form the component in Tp^I.

For each $j \in J$, record the subinterval of a_j to form the component in Sub^J.

For each $i \in I \setminus J$, let $j \in J$ be maximal with $j < i$. Then record $a_i \upharpoonright [T - \mathrm{fl}(a_j) - \gamma, T - \mathrm{fl}(a_j)) \in \mathrm{Ct}$, and list all of these as the component in $\mathrm{Ct}^{I \setminus J}$.

Combine these three components to form $g(a)$.

As $\{\vec{p}_j\}_{j \in J}$ a linearly independent set in the $|x|$-dimensional vector space $\mathbb{Q}_p^{|x|}$, $|J| \leq |x|$, so

$$\left| \mathrm{Sg}' \to \mathrm{Tp}^I \times \mathrm{Sub}^J \times \mathrm{Ct}^{I \setminus J} \right| = \mathcal{O}(K^{|I|} \cdot |B|^{|J|} \cdot p^{\gamma |I \setminus J|}) = \mathcal{O}(|B|^{|J|}),$$

and it suffices to show that if $a, a' \in \mathbb{Q}_p^{|x|}$ are such that $f(a), f(a') \in \mathrm{Sg}'$, and $g(f(a)) = g(f(a'))$, then a, a' are in the same cell of $\mathcal{T}(B)$. That would show that the number of cells intersecting Sg' is at most $\left|\mathrm{Sg}' \to \mathrm{Tp}^I \times \mathrm{Sub}^J \times \mathrm{Ct}^{I \setminus J}\right| = \mathcal{O}(|B|^{|x|})$. Then as the number of pieces in the partition is itself only dependent on I, the total number of cells in $\mathcal{T}(B)$ is also $\mathcal{O}(|B|^{|x|})$ as desired.

If a, a' are such that $f(a), f(a') \in \mathrm{Sg}'$, then immediately we know that $(T - \mathrm{val}(p_i(a)))_{i \in I}$ and $(T - \mathrm{val}(p_i(a')))_{i \in I}$ have the same order type. If also $g(f(a)) = g(f(a'))$, then for each $i \in I$, $p_i(a)$ and $p_i(a')$ have the same subinterval type, so it suffices to show that for each i, $p_i(a)$ and $p_i(a')$ are in the same subinterval. This is clearly true for $i \in J$, but we need to consider the $\mathrm{Ct}^{I \setminus J}$ component of g to show that it is true for $i \in I \setminus J$. Bobkov shows this in Claim 4.3.8 and the subsequent paragraph of [Bob17]. That argument is summarized here:

Fix such an $i \in I \setminus J$, and let $j \in J$ be maximal with $j < i$. By the definition of Sg', $T - \mathrm{fl}(a_i) \leq T - \mathrm{fl}(a_j)$ and $T - \mathrm{fl}(a'_i) \leq T - \mathrm{fl}(a'_j)$, and as a_j, a'_j lie in the same subinterval, $T - \mathrm{fl}(a_j) = T - \mathrm{fl}(a'_j)$. Claim 4.3.8 in [Bob17] shows that $v(a_i - a'_i) > T - \mathrm{fl}(a_j) - \gamma$. As the Ct components of $g(f(a))$ and $g(f(a'))$ are also the same, we know that $a_i \upharpoonright [T - \mathrm{fl}(a_j) - \gamma, T - \mathrm{fl}(a_j)) = a'_i \upharpoonright [T - \mathrm{fl}(a'_j) - \gamma, T - \mathrm{fl}(a'_j))$, but as $[T - \mathrm{fl}(a_j) - \gamma, T - \mathrm{fl}(a_j)) = [T - \mathrm{fl}(a'_j) - \gamma, T - \mathrm{fl}(a'_j))$ and $v(a_i - a'_i) > T - \mathrm{fl}(a_j) - \gamma$, this tells us that even more coefficents of a_i and a'_i agree, so $v(a_i - a'_i) > T - \mathrm{fl}(a_j) \geq \max(T - \mathrm{fl}(a_i), T - \mathrm{fl}(a'_i))$. Assume without loss of generality that $T - \mathrm{fl}(a_i) \leq T - \mathrm{fl}(a'_i)$, and let the subintervals of a_i and a'_i be $I(t, T - \mathrm{fl}(a_i), \alpha_U)$ and $I(t', T - \mathrm{fl}(a'_i), \alpha'_U)$. Then as $v(a_i - a'_i) > T - \mathrm{fl}(a'_i)$ and $v(t' - a'_i) > T - \mathrm{fl}(a'_i)$, the ultrametric inequality gives us $v(a_i - t') > T - \mathrm{fl}(a'_i)$, so $a_i \in B_{T - \mathrm{fl}(a'_i)}(t')$ and $a_i \in B_{T - \mathrm{fl}(a_i)}(t)$, so one ball is contained in the other. By the assumption on the radii, $B_{T - \mathrm{fl}(a'_i)}(t') \subseteq B_{T - \mathrm{fl}(a_i)}(t)$. If the subintervals are distinct, they must be disjoint, in which case $B_{T - \mathrm{fl}(a'_i)}(t') \subseteq B_{T - \mathrm{fl}(a_i)}(t) \setminus I(t, T - \mathrm{fl}(a_i), \alpha_U)$. However, $a_i \in B_{T - \mathrm{fl}(a'_i)}(t') \cap I(t, T - \mathrm{fl}(a_i), \alpha_U)$, contradicting this. Thus the subintervals are the same.

2.6.3 A Conjecture about Locally Modular Geometric Structures

The following proposition, together with Theorem 2.6.1, lends support to a conjecture about distal cell decompositions in locally modular geometric structures. Recall that a structure is *geometric* when the acl operation defines a pregeometry and the structure is uniformly bounded (it eliminates the $\exists^\infty$ quantifier) [HP94].

Proposition 2.6.19. *The structure $\mathcal{M}$ with universe $\mathbb{Q}_p$ in the language $\mathcal{L}_{\text{aff}}$ is a modular geometric structure.*

Proof. To check this, it suffices to check that this structure is uniformly bounded, and that its algebraic closure operation acl gives rise to a modular pregeometry.

First we check uniform boundedness. That is, we wish to show that for all partitioned $\mathcal{L}_{\text{aff}}$-formulas $\varphi(x;y)$ with $|x| = 1$, there is some $n \in \mathbb{N}$ such that for all $b \in M^{|y|}$, either $|\varphi(M;b)| \leq n$ or $\varphi(M;b)$ is infinite.

By Lemma 2.6.5, $\varphi(M, M^{|y|})$ is a disjoint union of $(|y|+1)$-cells of the form $\{(x,y) \in \mathbb{Q}_p \times D | v(a_1(y)) \square_1 v(x - c(y)) \square_2 v(a_2(y)), x - c(y) \in \lambda Q_{m,n}\}$. Let n_φ be the number of cells in that disjoint union. We will show that for all $b \in M^{|y|}$, either $|\varphi(M;b)| \leq n_\varphi$ or $\varphi(M;b)$ is infinite. To do this, we will show that for each cell Δ, defined by the formula $\psi(x;y)$, that for all $b \in M^{|y|}$, either the fiber $\psi(M;b)$ is infinite, or $|\psi(M;b)| \leq 1$. Then for $b \in M^{|y|}$, if the original set $\varphi(M;b)$ is finite, then each fiber $\psi(M;b)$ of the cells are finite, and thus each is at most a singleton. Thus $|\varphi(M;b)|$ is at most the number of cells n_φ.

Now consider a formula $\psi(x;y)$ that defines an $(|y|+1)$-cell, and the fibers of $\psi(M;b)$ for various $b \in M^{|y|}$. The fibers are of the form $\{x | v(a_1(b)) \square_1 v(x - c(b)) \square_2 v(a_2(b)), x - c(b) \in \lambda Q_{m,n}\}$, and we will show that any set of that form is either empty, infinite, or the singleton $\{c(b)\}$.

For simplicity, let us assume $c(b) = 0$. This amounts just to a translation of the set, and will not effect its size. Then assume $a \in \{x | v(a_1(b)) \square_1 v(x) \square_2 v(a_2(b)), x \in \lambda Q_{m,n}\}$, and we will show either that the set is $\{a\}$, or that it is infinite. If $\lambda = 0$, then $\lambda Q_{m,n} = \{0\}$, so we have $a = 0$ and the set is $\{0\}$. Thus we assume $\lambda \neq 0$. As $a \in \lambda Q_{m,n}$, there are some $k \in \mathbb{Z}, z \in \mathbb{Z}_p$ such that $a = \lambda p^{km}(1 + p^n z)$, and $v(a) = v(\lambda) + km + v(1 + p^n z)$. As $n \neq 0$, we have $v(p^n z) = nv(z) \geq n > 0$, so $v(1 + p^n z) = v(1) = 0$ by the ultrametric property, and $v(a) = v(\lambda) + km$. Now for any $z' \in \mathbb{Z}_p$, $v(\lambda p^{km}(1 + p^n z')) = v(a)$, and $\lambda p^{km}(1 + p^n z') \in \lambda Q_{m,n}$, so $\lambda p^{km}(1 + p^n z')$ is also in this set. As $\lambda \neq 0$, these are all distinct elements of the set, which is infinite.

Now we check that acl gives rise to a modular pregeometry. To do this, it suffices to check that acl is just the span operation, equal to acl in the plain vector space language, which also gives rise to a modular pregeometry. If $B \subseteq M, a \in M$, then $a \in \operatorname{acl}(B)$ if and only if there exists a formula $\varphi(x; y)$ with $|x| = 1$ and a tuple $b \in B^{|y|}$ such that $\varphi(M, b)$ is finite and $\mathcal{M} \models \varphi(a, b)$. If we decompose $\varphi(M; M^{|y|})$ into cells, then we see that there must exist a cell (say it is defined by $\psi(x; y)$) such that $a \in \psi(M, b)$. As $\psi(M, b) \subseteq \varphi(M, b)$ is also finite, and $\psi(x; y)$ defines a cell, $\psi(M; b) = \{c(b)\}$ for a defining polynomial c of the cell, which can be assumed to be linear with constant coefficient 0. Thus $a = c(b)$, so a is in the span of B. Clearly also the span of B is contained in $\operatorname{dcl}(B) \subseteq \operatorname{acl}(B)$, so acl = dcl, and both represent the span. □

Conjecture 2.6.20. We conjecture that all distal locally modular geometric structures admit distal cell decompositions of exponent 1. We have already shown this in the o-minimal case with Theorem 2.4.2, and now we have shown this for the linear reduct of $\mathbb{Q}_p$ with Theorem 2.6.1.

2.7 $\mathbb{Q}_p$, the Valued Field

Let $\mathcal{K}$ be a P-minimal field, taken as a structure in Macintyre's language, which consists of the language of rings together with a symbol to define the valuation and a unary relation P_n for each $n \geq 2$, interpreted so that $P_n(x) \iff \exists y, y^n = x$. While the symbol to define the valuation can be chosen either to be a unary predicate defining the valuation ring or a binary relation $|$ interpreted so that $x|y \iff v(x) \leq v(y)$, we will refer directly to the valuation v for legibility. The symbols P_n are included so that this structure has quantifier-elimination [Mac76]. Furthermore, assume that $\mathcal{K}$ has definable Skolem functions. (This assumption is only required to invoke the cell decomposition seen at equation 7.5 from [ADH16]. The existence of this cell decomposition is shown to be equivalent to definable Skolem functions in [Mou09].)

Theorem 2.7.1. *Let Φ be a finite set of formulas of the form $\varphi(x; y)$. Then Φ admits a distal cell decomposition with exponent $3|x| - 2$.*

Proof. This follows from Lemma 2.7.2 below, together with Theorem 2.3.1. $\square$

Lemma 2.7.2. *If $|x| = 1$, then Φ admits a distal cell decomposition $\mathcal{T}$ with 3 parameters and exponent 1.*

In the rest of this section, we prove Lemma 2.7.2.

2.7.1 Simplification of Φ

To construct our distal cell decomposition, we start with a simpler notion of cell decomposition. Each formula $\varphi(x; y)$ with $|x| = 1$, and thus every $\varphi \in \Phi$, has a cell decomposition in the sense that $\varphi(x; y)$ is equivalent to the disjoint disjunction of the formulas $\varphi_i(x; y) : 1 \leq i \leq N$, each of the form

$$v(f(y))\square_1 v(x - c(y))\square_2 v(g(y)) \wedge P_n(\lambda(x - c(y)))$$

for some $n, N > 0$, where $\square_1$ is $<$ or no condition, $\square_2$ is $\leq$ or no condition, f, g, c are $\emptyset$-definable functions, and $\lambda \in \Lambda$, a finite set of representatives of the cosets of $P_n^\times$. By Hensel's Lemma, we can choose $\Lambda \subset \mathbb{Z} \subseteq \mathrm{dcl}(\emptyset)$, so that each cell is $\emptyset$-definable [Mac76]. Let F be the set of all functions appearing as f, g in these formulas, and C the set of all functions appearing as c (See equation 7.5, [ADH16]).

Now we define $\Phi_{F,C,\Lambda}(x;y)$ as the set of formulas $\{v(f(y)) < v(x - c(y)) : f \in F, c \in C\} \cup \{P_n(\lambda(x-c(y))) : c \in C, \lambda \in \Lambda\}$. It is easy to check that every formula of Φ is a boolean combination of formulas in $\Phi_{F,C,\Lambda}$, so a distal cell decomposition for $\Phi_{F,C,\Lambda}$ will also be a distal cell decomposition for Φ. Thus we may assume that Φ is already of the form $\Phi_{F,C,\Lambda}$. For additional ease of notation, we also assume F contains the constant function $f_0 : y \mapsto 0$.

2.7.2 Subintervals and subinterval types

Let $B_r(c)$ denote again the open ball centered at c with radius r: $B_r(c) = \{x \in K : v(x-c) > r\}$. Fix a finite set $B \subset M^{|y|}$, and let $\mathcal{B}$ be a set of balls, similar to those referred to in [ADH16], Section 7.2 as "special balls defined over B", which we express as $\mathcal{B} := \mathcal{B}_F \cup \mathcal{B}_C$, where

$$\mathcal{B}_F = \{B_{v(f(b))}(c(b)) : b \in B, f \in F, c \in C\}$$

and

$$\mathcal{B}_C = \{B_{v(c_1(b_1)-c_2(b_2))}(c_1(b_1)) : b_1, b_2 \in B, c_1, c_2 \in C\}.$$

Clearly $|\mathcal{B}_F| = \mathcal{O}(|B|)$. It is less clear that $|\mathcal{B}_C| = \mathcal{O}(|B|)$, but this is a consequence of [ADH16, Lemma 7.3]. Thus $|\mathcal{B}| = \mathcal{O}(|B|)$.

Definition 2.7.3. We now define *subintervals* and surrounding notation, analogously to Definition 2.6.9, but with a different notion of subinterval types.

- Define a *subinterval* as an atom in the boolean algebra generated by $\mathcal{B}$.
- Each subinterval can be expressed as $I(t, \alpha_L, \alpha_U)$ where

  $$I(t, \alpha_L, \alpha_U) = B_{\alpha_L}(t) \setminus \bigcup_{t' \in T \cap B_{\alpha_U - 1}(t)} B_{\alpha_U}(t'),$$

  for some $t = c_i(b_0)$ with $i \in I, b_0 \in B$, and $\alpha_L = \alpha_1(b_0, b_1), \alpha_U = \alpha_2(b_0, b_2)$, with α_1, α_2 chosen from a finite set A of $\emptyset$-definable functions $\mathbb{Q}_p^2 \to \Gamma$, including two functions defined, by abuse of notation, as $\pm\infty$.

- The subinterval $I(t, \alpha_L, \alpha_U)$ is said to be *centered* at t.

- For $a \in \mathbb{Q}_p$, define $T - \mathrm{val}(a) := v(a - t)$, where a belongs to a subinterval centered at t. As in Definition 2.6.10, this is well-defined.

- Given a subinterval $I(t, \alpha_L, \alpha_U)$, two points a_1, a_2 in that subinterval are defined to have the same *subinterval type* if one of the following conditions is satisfied:

  1. $\alpha_L + 2v(n) < T{-}\mathrm{val}(a_i) < \alpha_U - 2v(n)$ for $i = 1, 2$, and $(a_1 - t)(a_2 - t)^{-1} \in P_n^{\times}$
  2. $\neg(\alpha_L + 2v(n) < T{-}\mathrm{val}(a_i) < \alpha_U - 2v(n))$ for $i = 1, 2$, and

     $T{-}\mathrm{val}(a_1) = T{-}\mathrm{val}(a_2) < v(a_1 - a_2) - 2v(n)$.

We will construct a distal cell decomposition $\mathcal{T}(B)$ where each cell consists of all points in a fixed subinterval with a fixed subinterval type. There are several requirements to check for this:

1. The sets of points in a fixed subinterval with a fixed subinterval type are uniformly definable from three parameters in B.

2. If two points lie in the same subinterval and have the same subinterval type, then they have the same Φ-type over B.

3. K has $\mathcal{O}(|B|)$ subintervals, and each divides into a constant number of subinterval types.

The first and second requirements will verify that this is a valid distal cell decomposition. The third will verify that $|\mathcal{T}(B)| \leq \mathcal{O}(|B|)$, and thus that $\mathcal{T}$ has exponent 1. The first will guarantee that $\mathcal{T}$ uses only three parameters.

First we check the first requirement. We see that the triple (t, α_L, α_U) can always be defined from a triple $(b_0, b_1, b_2) \in B^3$, so it suffices to show that each cell (subinterval type) in the subinterval $I(t, \alpha_L, \alpha_U)$ can be defined from (t, α_L, α_U) and no other parameters in B. Note that while in Section 2.6, we showed that the subintervals are uniformly definable, and the same argument would hold here, the defining formulas there may need more than three parameters, so we give a different argument.

A subinterval type of the first kind can be defined from t, α_L, α_U by $\psi_\lambda(t, \alpha_L, \alpha_U) := \alpha_L + 2v(n) < v(x-t) < \alpha_U - 2v(n) \wedge P_n(\lambda(x-t))$. A subinterval type of the second kind is just a ball, of the form $B_{r+2v(n)}(q)$, where either $r = \alpha_L + i$ with $0 < i \leq 2v(n)$, or $r = \alpha_U - i$, with $0 \leq i \leq 2v(n)$, and q satisfies $T - \mathrm{val}(q) = r$, which is implied by $v(t-q) = r$. For a fixed t, α_L, α_U, there are a constant number of choices for r, and q can be chosen to be $p^r(q_0) + t$, where q_0 is chosen from a set Q of representatives for open balls of radius $2v(n)$ such that $v(q_0) = 0$.

Given a potential cell Δ which represents a subinterval type within the set $I(t, \alpha_L, \alpha_U)$, we want to define $\mathcal{I}(\Delta)$ so that $\mathcal{I}(\Delta) \cap B = \emptyset$ if and only if there actualy is a subinterval $I(t, \alpha_L, \alpha_U)$. There is such an interval if and only if there are no balls in $\mathcal{B}$ strictly containing $B_{\alpha_U}(t)$ and strictly contained in $B_{\alpha_L}(t)$. A ball $B_{v(f(b)}(c(b)) \in \mathcal{B}_F$ for some $b \in B, f \in F, c \in C$ lies between those two balls if and only if $\alpha_L < v(f(b)) < \alpha_U$ and $v(f(b)) < v(t - c(b))$, so define

$$\theta_{f,c}(y; t, \alpha_L, \alpha_U) := \alpha_L < v(f(y)) < \alpha_U \wedge v(f(b)) < v(t - c(b)).$$

A ball $B_{v(c_1(b_1)-c_2(b_2))}(c_1(b_1)) \in \mathcal{B}_C$ for some $b_1, b_2 \in B, c_1, c_2 \in C$ lies between those two balls if and only if $\alpha_L < v(c_1(b_1) - c_2(b_2)) < \alpha_U$ and $v(c_1(b_1) - c_2(b_2)) < v(t - c_1(b_1))$. If this is true, then $B_{v(c_1(b_1)-c_2(b_2))}(c_1(b_1)) = B_{v(t-c_2(b_2))}(t)$, so it is enough to check if there is a ball $B_{v(t-c(b))}(t)$ that lies between those two balls. That happens if and only if $\alpha_L < v(t-c(b)) < \alpha_U$, so define

$$\theta_c(y; t, \alpha_L, \alpha_U) := \alpha_L < v(t - c(y)) < \alpha_U.$$

Then $\mathcal{I}(\Delta)$ is defined by the formula

$$\bigvee_{c \in C} \left(\theta_c(y; t, \alpha_L, \alpha_U) \vee \left(\bigvee_{f \in F} \theta_{f,c}(y; t, \alpha_L, \alpha_U) \right) \right)$$

as desired.

Now we will check the third requirement. Ordering the balls of $\mathcal{B}$ by inclusion forms a poset, whose Hasse diagram can be interpreted as a graph. By the ultrametric property, any two intersecting balls are comparable in this ordering, which rules out cycles in the graph. As the number of vertices is $|\mathcal{B}| = \mathcal{O}(|B|)$ and the graph is acyclic, the number of edges is also $\mathcal{O}(|B|)$. There are also $\mathcal{O}(|B|)$ subintervals, because there is (almost) a surjection from edges of the graph to subintervals: given an edge between B_1 and B_2, assuming without loss of generality that $B_2 \subsetneq B_1$, we can assign it to the subinterval $I(t, \alpha_L, \alpha_U)$, where $t \in B_2$, α_L is the radius of B_1, and α_U is the radius of B_2. This omits the subintervals with outer ball K, and the subintervals representing minimal balls in $\mathcal{B}$, but there are $\mathcal{O}(|B|)$ of those as well.

Now we will check that each subinterval breaks into only a constant number of subinterval types. Fix a subinterval $I(t, \alpha_L, \alpha_U)$. Then the subinterval types of the first kind correspond with cosets of $P_n^\times$, of which there are n (or $n+1$ if one takes into account the fact that 0 is not in the multiplicative group at all). As in Section 2.6, or [Bob17, Lemma 4.2.11], there will also be a constant number of subinterval types of the second kind. We have seen that these can be defined as $B_{r+2v(n)}(q)$. For our fixed (t, α_L, α_U), r must be either $\alpha_L + i$ with $0 < i \leq 2v(n)$ or $\alpha_U - i$ with $0 \leq i \leq 2v(n)$, which leaves only finitely many choices. For a fixed r, q must be of the form $p^r(q_0) + t$, where q_0 is chosen from a fixed finite set, so there are $|Q|$ choices of q.

Now we check the second requirement. Let $\varphi \in \Phi, b \in B$. Then $\varphi(x; b)$ is either of the form $v(f(b)) < v(x - c(b))$ for $f \in F, c \in C$ or $P_n(\lambda(x - c(b)))$ for $c \in C, \lambda \in \Lambda$.

If $\varphi(x; b)$ is $v(f(b)) < v(x - c(b))$, then the set of points satisfying $\varphi(x; b)$ is a ball in

$\mathcal{B}$, so a subinterval, as an atom in the boolean algebra generated by $\mathcal{B}$, is not crossed by that ball, or the formula $v(f(b)) < v(x - c(b))$. Thus each cell of $\mathcal{T}(B)$, being a subset of a subinterval, is not crossed by $\varphi(x; b)$.

Now it suffices to check that each cell is not crossed by $\varphi(x; b)$, where $\varphi(x; b)$ is $P_n(\lambda(x - c(b)))$ for $c \in C, \lambda \in \Lambda$. To do this, we will need the following lemma:

Lemma 2.7.4 (7.4 in [ADH16]). *Suppose $n > 1$, and let $x, y, a \in K$ with $v(y - x) > 2v(n) + v(y - a)$. Then $(x - a)(y - a)^{-1} \in P_n^\times$.*

We will show that any two points a_1, a_2 in a given subinterval $I(t, \alpha_L, \alpha_U)$ with a given subinterval type satisfy $(a_1 - c(b))(a_2 - c(b))^{-1} \in P_n^\times$. This shows that $\mathcal{K} \models P^n(\lambda(a_1 - c(b))) \iff P^n(\lambda(a_2 - c(b)))$, so the cell defined by points in that subinterval with that subinterval type is not crossed by $\varphi(x; b)$.

We will do casework on the two kinds of subinterval types, but for both we use the fact that the definition of $I(t, \alpha_L, \alpha_U)$ implies that either $v(t - c(b)) \leq \alpha_L$, or $v(t - c(b)) \geq \alpha_U$.

In the first kind of subinterval type, we have $(a_1 - t)(a_2 - t)^{-1} \in P_n^\times$ by definition, so it suffices to show, without loss of generality, that $(t - a_1)(c(b) - a_1)^{-1} \in P_n^\times$. Lemma 2.7.4 shows that this follows from $v(t - c(b)) > 2v(n) + v(t - a_1)$. As $T - \mathrm{val}(a_1) = v(t - a_1)$, this is equivalent to $v(t - c(b)) \geq \alpha_U$. By the construction of $I(t, \alpha_L, \alpha_U)$, this is one of two cases, and we are left with the case $v(t - c(b)) \leq \alpha_L$. In that case, $v(t - c(b)) + 2v(n) < v(t - a_1)$. Thus $(a_1 - c(b))(t - c(b))^{-1} \in P_n^\times$, and similarly, $(a_2 - c(b))(t - c(b))^{-1} \in P_n^\times$, so we get $(a_1 - c(b))(a_2 - c(b))^{-1} \in P_n^\times$.

In the second kind of subinterval type, we have $v(a_1 - t) = v(a_2 - t) < v(a_1 - a_2) - 2v(n)$. If $v(t - c(b)) \geq \alpha_U$, then as $a_1 \in I(t, \alpha_L, \alpha_U)$, we have $\alpha_L < v(a_1 - t) \leq \alpha_U$, we have $v(a_1 - c(b)) = v(a_1 - t)$ by the ultrametric property. Thus $v(a_1 - c(b)) + 2v(n) < v(a_1 - a_2)$, so by Lemma 2.7.4, $(a_1 - c(b))(a_2 - c(b))^{-1} \in P_n^\times$. In the other case, $v(t - c(b)) \leq \alpha_L < v(a_1 - t)$, so the lemma tells us that $v(a_1 - c(b)) = v(t - c(b)) < v(a_1 - t) - 2v(n)$, so by the lemma, $v(a_1 - c(b))(a_1 - t)^{-1} \in P_n^\times$, and also $v(a_2 - c(b))(a_2 - t)^{-1} \in P_n^\times$, so as also

$v(a_1 - t) + 2v(n) < v(a_1 - v_2)$, so $(a_1 - t)(a_2 - t)^{-1} \in P_n^\times$, so we can combine all these facts to get $(a_1 - c(b))(a_2 - c(b))^{-1} \in P_n^\times$.

2.8 Zarankiewicz's Problem

In this section, we introduce background on Zarankiewicz's problem, and the bounds known for the case of distal-definable bipartite graphs in general. We then combine these general bounds with the bounds on distal cell decompositions throughout in this chapter, arriving at concrete combinatorial corollaries for the distal structures we have discussed.

2.8.1 Background

First we will want to define the notion of a bigraph. A bigraph consists of a pair of sets X, Y and a relation $E \subset X \times Y$ such that E is a bipartite graph with parts X and Y. We say that such a bigraph *contains* a $K_{s,u}$ if there is a subset $A \subset X$ with $|A| = s$ and a subset $B \subset Y$ with $|B| = t$ such that E restricted to $A \times B$ is a complete bipartite graph (isomorphic to $K_{s,u}$).

Zarankiewicz's problem asks to bound asymptotically in m and n the number of edges in the largest bipartite graph on $m \times n$ omitting the subgraph $K_{s,t}$. Better bounds are known when we fix a particular infinite bigraph E omitting some $K_{s,t}$, and bound the size of the largest subgraph with parts of size m, n respectively. If P, Q are subsets of the parts of E, then we write $E(P, Q)$ to denote the set of edges between P, Q, so we concern ourselves with bounding $|E(P, Q)|$ in terms of $|P|$ and $|Q|$. This applies easily to problems in incidence geometry - if Γ is a family of curves on $\mathbb{R}^n$, we may consider an incidence graph on parts $\mathbb{R}^n$ and Γ defined by placing an edge between (p, γ) exactly when $p \in \gamma$. When these curves are algebraic of bounded degree, Bézout's theorem bounds the size of a complete bipartite subgraph $K_{s,t}$ in this incidence graph, and then we are interested in the number of edges (incidences) between a finite set of points and a finite set of curves. For a general reference on incidence geometry, see [She22].

We will concern ourselves with the case where the bigraph is definable in a distal structure. In the incidence example, this happens when the curves in Γ are uniformly definable in some distal structure on $\mathbb{R}$. In [CGS20], the authors set an upper bound for Zarankiewicz's problem in bigraphs definable in a distal structure, using distal cell decompositions as the foundation of their approach. The resulting bound depends essentially on the distal density of the definable graph - this is our primary motivation for defining distal density and distal exponents in this chapter.

The approach of [CGS20] follows a classic divide-and-conquer argument used in [Mat02, Section 4.5] to prove the Szemerédi-Trotter theorem, which states that if we let Γ be the set of lines in $\mathbb{R}^2$, then

$$|E(P,Q)| = \mathcal{O}\left(|P|^{2/3}|Q|^{2/3} + |P| + |Q|\right).$$

This is proven using cuttings:

Definition 2.8.1. Let $\mathcal{F}$ be a finite family of subsets of a set X with $|\mathcal{F}| = n$. Given a real $1 < r < n$, we say that a family $\mathcal{C}$ of subsets of X is a $\frac{1}{r}$-*cutting for* $\mathcal{F}$ when $\mathcal{C}$ forms a cover of X and each set $C \in \mathcal{C}$ is crossed by at most $\frac{n}{r}$ elements of $\mathcal{F}$.

Cuttings differ from abstract cell decompositions in that a limited amount of crossing is allowed, but they are still related. In [Mat02, Section 6.5], a bound ([Mat02, Lemma 4.5.3]) is given on the size of an $\frac{1}{r}$-cutting into triangles with respect to any finite set of lines. For a given set of points and a given set of lines, a particular value of r is chosen, an $\frac{1}{r}$-cutting is found, and then for each triangle in the cutting, the set of incidences between points in the triangle and lines that cross the triangle is bounded. These bounds are summed, and after considering some exceptional cases, this proves Szemerédi-Trotter.

In [CGS20], meanwhile, the authors find uniformly definable cuttings for each definable relation, starting with a distal cell decomposition. The size of the cutting given by this cutting lemma scales directly with the size of the given distal cell decomposition, so the bounds on distal cell decompositions throughout this chapter also function as bounds on the sizes of cuttings.

Fact 2.8.2 (Distal Cutting Lemma: [CGS20, Theorem 3.2])**.** *Let $\phi(x;y)$ be a formula admiting a distal cell decomposition of exponent d. Then for any natural n and any real $1 < r < n$, there exists $t = \mathcal{O}(r^d)$ such that for any finite $H \subseteq M^{|y|}$ of size n, there are uniformly definable sets $X_1, \ldots, X_t \subseteq M^{|x|}$ which form an $\frac{1}{r}$-cutting for $\{\phi(x;h) : h \in H\}$.*

The proof of this also follows the proof of the cutting lemma for lines in [Mat02, Sections 4.6 and 6.5], which in turn uses the random sampling technique of Clarkson and Shor.[CS89].

From this cutting lemma, a similar divide-and-conquer argument works. Given a formula $\phi(x;y)$ on a distal structure M defining a bigraph E on $M^{|x|} \times M^{|y|}$, for any finite subset $H \subseteq M^{|y|}$, the authors of [CGS20] use a distal cell decomposition and the distal cutting lemma to find a suitable cutting for $\{\phi(x;h) : h \in H\}$. They then, in summary, use other tools to bound the incidences between the points in each cell of the cutting and formulas $\phi(x;h)$ which cross it, and combine these bounds to find a final result, quoted here in our terminology:

Fact 2.8.3 ([CGS20, Theorem 5.7])**.** *Let $\mathcal{M}$ be a structure and $d, t \in \mathbb{N}_{\geq 2}$. Assume that $E(x,y) \subseteq M^{|x|} \times M^{|y|}$ is a definable relation given by an instance of a formula $\theta(x,y;z) \in \mathcal{L}$, such that the formula $\theta'(x;y,z) := \theta(x,y;z)$ has a distal cell decomposition of exponent t, and such that the VC density of $\theta''(x,z;y) := \theta(x,y;z)$ is at most d. Then for any $k \in \mathbb{N}$ there is a constant $\alpha = \alpha(\theta, k)$ satisfying the following.*

For any finite $P \subseteq M^{|x|}, Q \subseteq M^{|y|}$, $|P| = m, |Q| = n$, if $E(P,Q)$ is $K_{k,k}$-free, then we have:

$$|E(P,Q)| \leq \alpha \left(m^{\frac{(t-1)d}{td-1}} n^{\frac{t(d-1)}{td-1}} + m + n \right).$$

While d, t are assumed to be integral in their theorem statement, they could be replaced with any real $d, t \in \mathbb{R}_{\geq 2}$ and their proof would work unchanged. If θ' has distal density t, then it is not known if θ must have a distal cell decomposition of exponent precisely t. However, we can still get nearly the same bound, as for all $\varepsilon > 0$, θ' has a distal cell decomposition with exponent $t + \varepsilon$. As $\lim_{\varepsilon \to 0} \frac{(t+\varepsilon-1)d}{(t+\varepsilon)d-1} = \frac{(t-1)d}{td-1}$, and $\frac{(t+\varepsilon)(d-1)}{(t+\varepsilon)d-1} \leq \frac{t(d-1)}{td-1}$, the theorem still holds for θ' with distal density t, except with the final bound replaced by

$$|E(P,Q)| \leq \alpha \left(m^{\frac{(t-1)d}{td-1}+\varepsilon} n^{\frac{t(d-1)}{td-1}} + m + n \right)$$

for arbitrary $\varepsilon > 0$ and $\alpha = \alpha(\theta, k, \varepsilon)$.

Contrast this result to an analogous result for semi-algebraic sets, using polynomial partitioning for the divide-and-conquer argument instead of cuttings:

Fact 2.8.4 ([Wal20, Corollary 1.7])**.** *Let P be a set of m points and let $\mathcal{V}$ be a set of n constant-degree algebraic varieties, both in $\mathbb{R}^d$, such that the incidence graph of $P \times \mathcal{V}$ does not contain $K_{s,t}$. Then for every $\varepsilon > 0$, we have*

$$I(P, \mathcal{V}) = \mathcal{O}_{d,s,t,\varepsilon} \left(m^{\frac{(d-1)s}{ds-1}+\varepsilon} n^{\frac{d(s-1)}{ds-1}} + m + n \right).$$

The initial version of this result, [FPS17, Theorem 1.2], had an extra factor of m^{ε} in the first term. The m^{ε} was removed first in special cases, such as in [BS16, Theorem 1.5], with a more involved application of polynomial partitioning, eventually leading to [Wal20].

Remark 1. The special case of $d = s = 2$ is proven in [FPS17, Theorem 1.1], without the extra factor of m^{ε}, using the cutting lemma strategy generalized by [CGS20]. This method would imply the rest of Fact 2.8.4 given a distal cell decomposition of exponent $|x|$ for each finite set $\Phi(x; y)$ of formulas in the language of ordered rings over $\mathbb{R}$.

As a last remark before examining these combinatorial applications in specific structures, we mention some other combinatorial applications of distal cell decompositions which may be improved using specific bounds like those in this chapter. While the papers are different in strategy and scope, both [Bas09, Theorem 2.6] and [CS18, Theorem 1.9] apply techniques that we now recognize as distal cell decompositions and distal cutting lemmas Ramsey-theoretically, showing that sets definable in distal structures satisfy a property that [CS18] dubs the *strong Erdős-Hajnal property.* The constants in this asymptotic bound are improved by providing better bounds on exponents of distal cell decompositions.

2.8.2 New Results in Specific Structures

In this subsection, we collect the results from earlier in the chapter and combine them with the Zarankiewicz bounds of [CGS20] as cited above.

We begin by just applying Fact 2.8.3 with known distal exponent and VC density bounds, listing the exponents in the resulting Zarankiewicz bounds in a table.

Corollary 2.8.5. *Let $\mathcal{M}$ be a structure from the left column of the following table and let $E \subseteq M^a \times M^b$ be a definable bigraph. Then for any $k \in \mathbb{N}$, there is a constant $\alpha = \alpha(\theta, k)$ such that for the corresponding values of q and r in this table, and any finite $P \subseteq M^a, Q \subseteq M^b$, $|P| = m, |Q| = n$, if $E(P, Q)$ is $K_{k,k}$-free, then $|E(P, Q)| \leq \alpha\left(m^q n^r + m + n\right).$*

$\mathcal{M}$	q	r
o-minimal expansions of groups	$\frac{(2a-3)b}{(2a-2)b-1}$	$\frac{(2a-2)(b-1)}{(2a-2)b-1}$
weakly o-minimal structures	$\frac{(2a-2)b}{(2a-1)b-1}$	$\frac{(2a-1)(b-1)}{(2a-1)b-1}$
ordered vector spaces over ordered division rings	$\frac{(a-1)b}{ab-1}$	$\frac{a(b-1)}{ab-1}$
Presburger arithmetic	$\frac{(a-1)b}{ab-1}$	$\frac{a(b-1)}{ab-1}$
$\mathbb{Q}_p$ the valued field	$\frac{(3a-3)(2b-1)}{(3a-2)(2b-1)-1}$	$\frac{(3a-2)(2b-2)}{(3a-2)(2b-1)-1}$
$\mathbb{Q}_p$ in the linear reduct	$\frac{(a-1)b}{ab-1}$	$\frac{a(b-1)}{ab-1}$

Table 2.2: Zarankiewicz Bounds for Definable Graphs in Distal Structures

Proof. The bounds on VC densities and exponents of distal cell decompositions are listed in Theorem 2.1.1. The VC densities come from the literature cited in that theorem, as does the exponent for the distal cell decomposition in the case of o-minimal expansions of groups with $a = 2$ from [CGS20], but the rest of the distal cell decomposition bounds are new to this chapter. $\square$

In some applications to Zarankiewicz's problem, the omitted bipartite graph $K_{s,u}$ may give a better bound on the relevant VC density than is known for general formulas. The following lemma bounds the VC density for formulas defining relations which do not contain a $K_{s,u}$:

Lemma 2.8.6. *Let $\mathcal{M}$ be a first-order structure, and $\varphi(x; y)$ be a formula such that the bigraph with edge relation $\varphi(M^{|x|}; M^{|y|})$ does not contain $K_{s,u}$. Then* $\mathrm{vc}(\varphi) \leq s$.

Proof. An equivalent way (see [ADH16]) of defining $\pi_\varphi(n)$ is as $\max_{A\subset M^{|x|},|A|=n} |\varphi \cap A|$, where $\varphi \cap A$ is shorthand for $\{A \cap \varphi(M^{|x|}, b) : b \in M^{|y|}\}$.

Given $A \subset M^{|x|}$, find $B \subset M^{|y|}$ such that for each subset $A_0 \in \varphi \cap A$, there is exactly one $b \in B$ such that $A_0 = A \cap \varphi(M^{|x|}, b)$. Thus $|B| = |\varphi \cap A|$.

The number of subsets of A in $\varphi \cap A$ of size less than B is trivially bounded by $\sum_{i=0}^{s-1} \binom{|A|}{i} = \mathcal{O}(|A|^{s-1})$. Thus there are $\mathcal{O}(|A|^{s-1})$ elements $b \in B$ for which $\left|\varphi(M^{|x|}, b) \cap A\right| < s$. However, by assumption, for each subset $A_B \subseteq A$ of size B, there are most $t-1$ elements b of B with $\mathcal{M} \models \varphi(a, b)$ for all $a \in A_s$. Thus there are at most $(t-1)\binom{|A|}{s} = \mathcal{O}(|A|^s)$ elements $b \in B$ for which $\left|\varphi(M^{|x|}, b) \cap A\right| \geq s$, and in general, $|B| = \mathcal{O}(|A|^s)$, so $\pi_\varphi(n) = \mathcal{O}(n^s)$, and $\mathrm{vc}(\varphi) \leq s$. $\square$

Combining this lemma with Theorem 2.8.3 gives us the following Zarankiewicz bound for bigraphs defined in distal structures, making use only of the omitted complete bipartite subgraph for the VC density bound.

Corollary 2.8.7. *Let $\mathcal{M}$ be a structure and $t \in \mathbb{R}_{\geq 2}$. Assume that $E(x, y) \subseteq M^{|x|} \times M^{|y|}$ is a definable relation given by an instance of a formula $\theta(x, y; z) \in \mathcal{L}$, such that the formula $\theta'(x; y, z) := \theta(x, y; z)$ has a distal cell decomposition of exponent t, and the graph $E(x, y)$ does not contain $K_{s,u}$. Then there is a constant $\alpha = \alpha(\theta, s, u)$ satisfying the following.*

For any finite $P \subseteq M^{|x|}, Q \subseteq M^{|y|}$, $|P| = m, |Q| = n$, we have:

$$|E(P, Q)| \leq \alpha \left(m^{\frac{(t-1)s}{ts-1}} n^{\frac{t(s-1)}{ts-1}} + m + n \right).$$

This Corollary recalls one version of Theorem 2.6 of [CS20], which provides the same bound on $|E(P, Q)|$ from a slightly different assumption on t, and either the same condition of $\varphi(x; y)$ omitting $K_{s,u}$ for some u, or $\varphi(x; y)$ omitting $K_{u,u}$ for some u and having dual VC density at most s.

To phrase this corollary in terms of distal density, we must add a small error term again.

If instead t is the distal density of θ', then for all $\varepsilon \in \mathbb{R}_{>0}$, we get the bound

$$|E(P,Q)| \leq \alpha \left(m^{\frac{(t-1)s}{ts-1}+\varepsilon} n^{\frac{t(s-1)}{ts-1}} + m + n \right),$$

where α depends also on ε.

To illustrate the generality of Corollary 2.8.7, we will apply it to some specific structures. Let us first apply it to $\mathcal{M} = \mathbb{R}_{\text{exp}} = \langle \mathbb{R}; 0, 1, +, *, <, e^x \rangle$. This structure is an expansion of a field, and o-minimal by [Wil96], allowing us to apply the distal exponent bounds from Theorem 2.4.1. We define an *exponential polynomial* to be a function $\mathbb{R}^n \to \mathbb{R}$ in $\mathbb{Z}[x_1, \ldots, x_n, e^{x_1}, \ldots, e^{x_n}]$ as in [BKW10], and an *exponential-polynomial inequality* to be an inequality of exponential polynomials. As any exponential polynomial function over $\mathbb{R}$ is definable in this structure, a boolean combination of exponential-polynomial inequalities or equations will be as well. Combining all of this with Corollary 2.8.7 gives the following result:

Corollary 2.8.8. *Assume that $E(x,y) \subseteq \mathbb{R}^{|x|} \times \mathbb{R}^{|y|}$ is a relation given by a boolean combination of exponential-polynomial (in)equalities, and the graph $E(x,y)$ does not contain $K_{s,u}$. Then there is a constant $\alpha = \alpha(\theta, s, u)$ satisfying the following.z*

For any finite $P \subseteq \mathbb{R}^{|x|}, Q \subseteq \mathbb{R}^{|y|}$, $|P| = m, |Q| = n$, we have:

$$|E(P,Q)| \leq \alpha \left(m^{\frac{(2|x|-2)s}{(2|x|-1)s-1}} n^{\frac{(2|x|-1)(s-1)}{(2|x|-1)s-1}+\varepsilon} + m + n \right).$$

Let us also apply Corollary 2.8.7 to subanalytic sets over $\mathbb{Z}_p$, defined as in [DD88]:

Definition 2.8.9.

- A set $S \subseteq \mathbb{Z}_p^n$ is semianalytic if for every $x \in S$, there is an open neighborhood U of x such that $U \cap S$ can be defined by a boolean combination of inequalities of analytic functions.
- A set $S \subseteq \mathbb{Z}_p^n$ is subanalytic if for every $x \in S$, there is an open neighborhood U of x and a semianalytic set S' in $U \times \mathbb{Z}_p^N$ for some N such that $U \cap S = \pi(S')$, where

$\pi : U \times \mathbb{Z}_p^N \to U$ is the projection map.

For any n, the subanalytic subsets of $\mathbb{Z}_p^n$ are exactly the quantifier-free definable subsets in a structure $\mathcal{R}_{\text{an}}$, which is a substructure of the structure $\mathcal{K}_{\text{an}}$, consisting of $\mathbb{Q}_p$ with its analytic structure, as described in [DHM99]. As per Theorem A'/B from [DHM99], this structure is P-minimal with definable Skolem functions, we can apply the distal exponent bounds from Theorem 2.7.1, giving us this corollary:

Corollary 2.8.10. *Assume that $E(x, y) \subseteq \mathbb{Z}_p^{|x|} \times \mathbb{Z}_p^{|y|}$ is a subanalytic relation, and the graph $E(x, y)$ does not contain $K_{s,u}$. Then there is a constant $\alpha = \alpha(\theta, s, u)$ satisfying the following.*

For any finite $P \subseteq \mathbb{Z}_p^{|x|}, Q \subseteq \mathbb{Z}_p^{|y|}$, $|P| = m, |Q| = n$, we have:

$$|E(P, Q)| \leq \alpha \left(m^{\frac{(3|x|-3)s}{(3|x|-2)s-1}} n^{\frac{(3|x|-2)(s-1)}{(3|x|-2)s-1}+\varepsilon} + m + n \right).$$

CHAPTER 3

Fuzzy VC Combinatorics and Distality in Continuous Logic

The fields of model theory, machine learning, and combinatorics each have generalizations of VC-dimension for fuzzy and real-valued versions of set systems. These different dimensions define a unique notion of a VC-class for both fuzzy sets and real-valued functions. In this chapter, we study these VC-classes, obtaining generalizations of certain combinatorial results from the discrete case. These include appropriate generalizations of ε-nets, the fractional Helly property and the (p, q)-theorem.

We then apply these results to continuous logic. We prove that NIP for metric structures is equivalent to an appropriate generalization of honest definitions, which we use to study externally definable predicates and the Shelah expansion. We then examine distal metric structures, providing several equivalent characterizations, in terms of indiscernible sequences, distal types, strong honest definitions, and distal cell decompositions.

3.1 Introduction

Distal structures were first studied as a way to characterise non-stable behavior in NIP theories, and defined in terms of indiscernible sequences[Sim13]. They include some important non-stable NIP structures, such as weakly o-minimal structures and the p-adics. Subsequently, distality was re-defined combinatorially, in terms of strong honest definitions or distal cell decompositions, generalizing o-minimal cell decompositions, and providing the most general model-theoretic setting for semialgebraic incidence combinatorics [CS15] [CGS20] [CS18].

Continuous logic replaces the standard first-order structures of model theory with metric structures, and formulas with continous functions to the real interval $[0,1]$[BBH08]. This makes it the natural setting to study analytic objects such as probability algebras, Banach spaces, and C^*-algebras. It also has natural connections to topological dynamics, as Polish groups are exactly the automorphism groups of metric structures, and new research has linked stability and NIP to dynamics in this way[Mel10][BT16][Iba16]. Stability, NIP, n-dependence, and some other dividing lines of neostability theory have already been defined for metric structures, with applications such as continuous n-dependent or stable regularity [BU10] [Ben09] [CT20] [CCP24]. Meanwhile, distal metric structures have only been mentioned in the context of hyperimaginaries[KP22].

In this chapter, we aim to lay the groundwork for studying distality in continuous logic. We set up the basic theory of distal metric structures, proposing continuous versions of several definitions of distality, and proving them equivalent. Along the way, we prove results relevant to all NIP metric structures, including versions of honest definitions and uniform definability of types over finite sets (UDTFS), generalizing results from [CS13] and [CS15] in the discrete case. Chapters 4 and 5 will provide examples of distal metric structures and consider distal regularity (as developed in [CS18] and simplified in [Sim16]) in the context of continuous logic, providing further characterizations of distal metric structures in terms of Keisler measures.

In order to understand distality in continuous logic, we must first better understand NIP, and the various fuzzy and real-valued generalizations of VC-dimension. In Section 3.2, we use these to prove real-valued versions of some classic combinatorial theorems of VC-classes. Classically, a set system on a set X is a set or family of subsets of X, which can be thought of in terms of their characteristic functions $X \to \{0,1\}$. A fuzzy set system replaces the characteristic function $X \to \{0,1\}$ with a characteristic function $X \to \{0,1,*\}$, where $*$

denotes an indeterminate truth value. The most important examples of fuzzy set systems come from classes of functions $X \to [0,1]$. Given any family $\mathcal{F}$ of such functions, and any $0 \leq r < s \leq 1$, we can define a fuzzy set system by replacing each function $f : X \to [0,1]$ with the function $f_{r,s}$ such that $f(t) = 0$ for $t \leq r$, $f(t) = *$ for $r < t < s$, and $f(t) = 1$ for $t \geq s$. These fuzzy set systems arising from real-valued functions are central to Ben Yaacov's development of NIP in continuous logic, including a proof that randomizations of NIP structures are NIP[Ben09].

In Section 3.2, we review the different notions of VC-dimension for fuzzy set systems and real-valued function systems, such as fuzzy VC-dimension, Rademacher complexity, and covering numbers, and compare these, checking that all of these give rise to the same definition of a VC class of functions. We then show that VC classes of fuzzy sets admit ε-nets, while VC classes of functions admit ε-approximations. We then use a combination of these techniques to show a fractional Helly theorem (Theorem 3.2.29) and a real-valued (p,q)-theorem (Theorem 3.2.31), which we will later apply to the model theoretic context to get uniform (strong) honest definitions.

In Section 3.4, we apply the results of Section 3.2 to NIP metric structures, using background on continuous logic provided in Section 3.3. Just as in classical logic, where an NIP structure is one where every definable class of sets has finite VC-dimension, a metric structure is NIP (as defined in [Ben09]) when every definable class of functions is a VC class in any of the equivalent senses of Section 3.2. We find several more equivalent definitions of NIP, summarized in Theorem 3.4.14. In particular, NIP metric structures are characterized by the following version of honest definitions:

Theorem 3.1.1. *Let A be a closed subset of M^x where $M \preceq \mathcal{U}$ and $(M, A) \preceq (M', A')$. Let $\phi(x; y)$ be a definable predicate. Then there exists a definable predicate $\psi(x; z)$, which we call a* uniform honest definition *for $\phi(x; y)$, such that for every $b \in M^y$, there exists $d \in A^z$ such that*

- *for all $a \in A$, $\phi(a; b) = \psi(a; d)$*
- *for all $a' \in A'$, $\phi(a'; b) \leq \psi(a'; d)$.*

We then define the Shelah expansion of a metric structure by externally definable predicates, and use honest definitions to show that the Shelah expansion of an NIP metric structure is NIP, just as in the classical case developed in [CS13].

With these techniques for studying NIP metric structures, we turn our attention to distal metric structures in Section 3.5. Distal metric structures were briefly mentioned in [KP22], defined by applying the definition of distal indiscernible sequences to the continuous logic context. In this section, we flesh out the theory of distal metric structures, starting with that indiscernible sequence definition, and proving several equivalent characterizations:

Theorem 3.1.2 (Theorem 3.5.5). *If a metric theory T is NIP, then the following are equivalent:*

1. *T is distal.*
2. *Every global type is distal.*
3. *Every formula admits strong honest definitions (see Definition 3.5.8).*
4. *Every formula admits an ε-distal cell decomposition for each $\varepsilon > 0$ (see Definition 3.5.17).*

This generalizes characterizations of distality from [Sim15], [CS15], and [CGS20] to work with metric structures.

3.2 Fuzzy Combinatorics

In this section, we will generalize some combinatorial facts about set systems and relations of finite VC-dimension to fuzzy set systems and fuzzy relations.

The VC-dimension of fuzzy set systems was introduced for model theory purposes in [Ben09], and for machine learning purposes in [AHH22]. A fuzzy subset S of a set X, denoted $S \sqsubseteq X$, is formalized as pair (S_+, S_-) of disjoint subsets of X, where S_+ is the set of elements such that $x \in S$, S_- is the set of elements such that $x \notin S$, but for $x \in X \setminus (S_+ \cup S_-)$, the truth value of $x \in S$ is undefined. (These can also be modeled as partial functions to $\{0, 1\}$ on X, or as in [AHH22], functions to $\{0, *, 1\}$.) A fuzzy set system on X is a set of fuzzy subsets of X, and a fuzzy relation between X and Y is a fuzzy subset of $X \times Y$. A fuzzy relation $R \sqsubseteq X \times Y$ can produce two fuzzy set systems: R^Y is the fuzzy set system on X given by $\{(\{x : (x, y) \in R_+\}, \{x : (x, y) \in R_-\}) : y \in Y\}$, and R_X is the similarly-defined fuzzy set system on Y. Each fuzzy relation $R \sqsubseteq X \times Y$ has a corresponding dual fuzzy relation, $R^* \sqsubseteq Y \times X$, given by $(y, x) \in R^*_+ \iff (y, x) \in R_+$ and $(y, x) \in R^*_- \iff (y, x) \in R_-$. Any fuzzy set system $\mathcal{F}$ can also be thought of as a fuzzy relation $X \times \mathcal{F}$, given by $(\{(x, S) : x \in S_+\}, \{(x, S) : x \in S_-\})$, and thus we can define a dual fuzzy set system, $\mathcal{F}^*$, which is the fuzzy set system on $\mathcal{F}$ induced by the dual of that fuzzy relation.

Sometimes, for combinatorial results, it is more useful to think of a fuzzy subset of X as a pair of nested subsets, as $S_+ \subseteq X \setminus S_-$, where we think of the inner subset as the elements that are definitely in S, and the outer subset as the elements that could possibly be in S. If $\mathcal{F}$ is a fuzzy set system on X, then we can define the *inner* and *outer* set systems by $\mathcal{F}_i = \{S_+ : S \in \mathcal{F}\}$ and $\mathcal{F}_o = \{X \setminus S_- : S \in \mathcal{F}\}$. We can translate many of the combinatorial theorems known for non-fuzzy set systems by showing that if the assumptions of the theorem hold for $\mathcal{F}_i$, then the results will hold for $\mathcal{F}_o$.

Definition 3.2.1. Let $\mathcal{F}$ be a fuzzy set system on X and $Y \subseteq X$. We will define the basic notions of shattering and the shatter functions associated to $\mathcal{F}$.

- Let $\mathcal{F} \cap Y$ be the set of all subsets $Z \subseteq Y$ such that there exists $S \in \mathcal{F}$ with $S_+ \cap Y = Z$ and $S_- \cap Y = Y \setminus Z$.

- Let $\pi_{\mathcal{F}}(n) = \max_{Y\subseteq X:|Y|=n} |\mathcal{F} \cap Y|$. We call $\pi_{\mathcal{F}}$ the *shatter function* of $\mathcal{F}$.
- Say that $\mathcal{F}$ *shatters* Y when $\mathcal{F} \cap Y = \mathcal{P}(Y)$, or equivalently, $|\mathcal{F} \cap Y| = 2^{|Y|}$.
- Define the *dual shatter function*, $\pi^*_{\mathcal{F}}$, to be $\pi_{\mathcal{F}^*}$.

We now have the nomenclature to define the VC-dimension of a fuzzy set system, and VC classes.

Definition 3.2.2. Let $\mathcal{F}$ be a fuzzy set system on X, and $d \in \mathbb{N}$. We say that $\mathcal{F}$ has VC-dimension at least d when $\pi_{\mathcal{F}}(n) = 2^n$ for all $n \leq d$. The *VC-dimension* of $\mathcal{F}$, denoted $\mathrm{vc}(\mathcal{F})$, is then the largest such d if there is one, and is otherwise ∞. We say that $\mathcal{F}$ is a *VC class* when $\mathcal{F}$ has finite VC-dimension.

We define the *dual VC-dimension* $\mathrm{vc}^*(\mathcal{F})$ to be the VC-dimension of $\mathcal{F}^*$.

Note that this notion of dimension differs by 1 (in the finite case) from the notion of VC-index discussed in [Ben09]. This more closely matches the convention adopted in the combinatorics literature that will be cited later.

The following lemma shows that we do not need to define dual-VC classes, as they are the same as VC classes.

Fact 3.2.3 ([Ben09, Fact 2.14])**.** *If $R \sqsubseteq X \times Y$ is a fuzzy relation, then R_X is a VC class if and only if R^Y is. Thus we can simply speak of VC-relations without specifying whether we are referring to R_X or R^Y having finite VC-dimension.*

In order to understand the shatter function, we note that the Sauer-Shelah lemma translates easily to the fuzzy context:

Fact 3.2.4 ([Ben09])**.** *If $\mathcal{F}$ is a fuzzy set system on X with VC-dimension at most d, then for all n, $\pi_{\mathcal{F}}(n) \leq p_d(n)$, where $p_d(n) = \sum_{k\leq d} n^k = O(n^d)$.*

Unfortunately, this polynomial bound does not suffice to translate all probabilistic arguments using the shatter function, as for a fuzzy set system $\mathcal{F}$ on X and a subset $Y \subseteq X$, the number of actual possible fuzzy subset intersections $(S_+ \cap Y, S_- \cap Y)$ for $S \in \mathcal{F}$ could be much larger. In some cases, counting a *strong disambiguation* (as described in [AHH22]) will be more helpful:

Definition 3.2.5. If $S \sqsubseteq X$, say that a subset $S' \subseteq X$ *strongly disambiguates* S when $S_+ \subseteq S'$ and $S' \cap S_- = \emptyset$. Say that a set system $\mathcal{F}'$ on a set X *strongly disambiguates* a fuzzy set system $\mathcal{F}$ on X when for every fuzzy set $S \in \mathcal{F}$, there is some $S' \in \mathcal{F}'$ refining S.

The following lemma is an immediate consequence of [AHH22, Theorem 13], and can be thought of as a version of Sauer-Shelah for strong disambiguations, though its bound is slightly worse than polynomial.

Lemma 3.2.6. *Let $\mathcal{F}$ be a fuzzy set system on a finite set X of VC index at most d. Then there exists a non-fuzzy set system $\mathcal{F}'$ strongly disambiguating $\mathcal{F}$, with $|\mathcal{F}'| \leq |X|^{O(d \log(|X|))}$.*

We now look at fuzzy set systems derived from classes of real-valued functions. If $Q \subseteq [0,1]^X$, and $0 \leq r < s \leq 1$, then Q gives rise to the fuzzy set system $Q_{r,s}$ consisting of the fuzzy sets $q_{r,s} = (q_{\leq r}, q_{\geq s})$ for $q \in Q$, where $q_{\leq r} = \{x : q(x) \leq r\}$ and $q_{\geq s} = \{x : q(x) \geq s\}$. Then the inner set system of $Q_{r,s}$ is $Q_{\leq r} := \{\{x : q(x) \leq r\} : q \in Q\}$, and the outer is $Q_{<s} := \{\{x : q(x) < s\} : q \in Q\}$. If instead of a set of functions $Q \subseteq [0,1]^X$, we have a function $Q : X \times Y \to [0,1]$, we can define a fuzzy relation $Q_{r,s}$ on X and Y.

Definition 3.2.7. Let $Q \subseteq [0,1]^X$ be a collection of functions. We say that Q is a *VC-class* when for any $0 \leq r < s \leq 1$, the fuzzy set system $Q_{r,s}$ has finite VC-dimension.

If instead Q is a function $Q : X \times Y \to [0,1]$, we say that Q is a *VC-function* when for any $0 \leq r < s \leq 1$, the fuzzy relation $Q_{r,s}$ has finite VC-dimension.

3.2.1 Rademacher/Gaussian Complexity and ε-Approximations

In order to express another equivalent definition of VC classes of functions, we need to introduce the concepts of Rademacher/Gaussian complexity and mean width. This definition of a VC class will then allow us to retrieve a version of the VC Theorem, guaranteeing the existence of ε-approximations to VC classes.

Definition 3.2.8. Let $A \subseteq \mathbb{R}^n$. Let σ be a randomly chosen vector in $\mathbb{R}^n$. Define the *mean width* of A, $w(A, \sigma)$, to be $\mathbb{E}_\sigma[\sup_{a \in A} \sigma \cdot a]$.

If σ is chosen uniformly from $\{+1, -1\}$, then we call $w(A, \sigma)$ the *Rademacher mean width*, denoted $w_R(A) = w(A, \sigma)$.

If $\sigma = (\sigma_1, \ldots, \sigma_n)$, where the σ_is are independent Gaussian variables with distribution $N(0, 1)$, then we call $w(A, \sigma)$ the *Gaussian mean width*, denoted $w_G(A) = w(A, \sigma)$.

The following fact allows us to translate between statements using Rademacher and Gaussian variables:

Fact 3.2.9 ([Wai19, Exercise 5.5])**.** *For any $A \subseteq [0, 1]^n$,*

$$w_R(A) \leq \sqrt{\frac{\pi}{2}} w_G(A) \leq 2\sqrt{\log n} w_R(A)$$

We can now apply these definitions to function classes.

Definition 3.2.10. Let $Q \subseteq [0, 1]^X$ be a function class, and let $\bar{x} = (x_1, \ldots, x_n) \in X^n$. Define $Q(\bar{x}) = \{(q(x_1), \ldots, q(x_n)) : q \in Q\}$.

Then define the Rademacher mean width $r_Q(n)$ to be $\sup_{\bar{x} \in X} w_R(Q(\bar{x}))$, and the Gaussian mean width $g_Q(n)$ to be $\sup_{\bar{x} \in X} w_G(Q(\bar{x}))$.

If μ is a probability measure on X, then define the Rademacher complexity $r_{Q,\mu}(n)$ to be $\frac{1}{n}\mathbb{E}_{\mu^n}[w_R(Q(\bar{x}))]$, and the Gaussian complexity $g_{Q,\mu}(n)$ to be $\frac{1}{n}\mathbb{E}_{\mu^n}[w_G(Q(\bar{x}))]$. (Note the normalization factor $\frac{1}{n}$ - this is more useful for probability applications.)

It is easy to see that for all choices of Q, n, μ, we have $\frac{r_Q(n)}{n} \leq r_{Q,\mu}(n)$ and $\frac{g_Q(n)}{n} \leq g_{Q,\mu}(n)$. We now have the language to connect these notions to VC classes:

Lemma 3.2.11. *Let X be a set, and $Q \subseteq [0,1]^X$. The following are equivalent:*

1. *Q is a VC class.*
2. $\lim_{n\to\infty} \frac{g_Q(n)}{n} = 0$
3. $\lim_{n\to\infty} \frac{r_Q(n)}{n} = 0$

As a consequence, if Q is a VC class, and μ a probability measure on X, then

$$\lim_{n\to\infty} r_{Q,\mu}(n) = \lim_{n\to\infty} g_{Q,\mu}(n) = 0.$$

Proof. The equivalence between (i) and (ii) is given by [Ben09, Theorem 2.11], and the equivalence between (ii) and (iii) is evident from Fact 3.2.9. □

Definition 3.2.12. For a function $q \in [0,1]^X$ and $(x_1, \ldots, x_n) \in X^n$, define

$$\mathrm{Av}(x_1, \ldots, x_n; q) = \frac{1}{n}\sum_{i=1}^{n} q(x_i).$$

For a function class $Q \subseteq [0,1]^X$, a probability measure μ on X, and $\varepsilon > 0$, say that a tuple $(x_1, \ldots, x_n)$ is a ε-*approximation* for Q with respect to μ when for every $q \in Q$,

$$|\mathrm{Av}(x_1, \ldots, x_n; q) - \mathbb{E}_\mu[q]| \leq \varepsilon.$$

Fact 3.2.13 ([Wai19, Theorem 4.10])**.** *Let Q be a class of functions from X to $[0,1]$. Then for any finitely-supported probability measure μ on X, and any $\delta > 0$, we have*

$$\mu^n\left(\sup_{q\in Q} |\mathrm{Av}(x_1, \ldots, x_n; q) - \mathbb{E}_\mu[q]| > 2r_{Q,\mu}(n) + \delta\right) \leq \exp\left(-\frac{n\delta^2}{2}\right).$$

We can use this fact as a version of the VC theorem for function classes:

Theorem 3.2.14. *If $n \in \mathbb{N}$ is such that $n > 0$ and $\frac{r_Q(n)}{n} < \varepsilon$, then for any finitely-supported probability measure μ on X, then there exists an ε-approximation for Q in the support of μ of size at most n.*

In particular, if Q is a VC class and μ a finitely-supported probability measure, then for every $\varepsilon > 0$, there exists a ε-approximation for Q in the support of μ, of size at most n, where $n = n(\varepsilon, r_Q)$.

Proof. Fix $0 < \delta < \varepsilon - \frac{r_Q(n)}{n}$. Then the probability that a randomly selected tuple $(x_1, \ldots, x_n)$ is not an ε-approximation is

$$\mu^n \left(\sup_{q \in Q} |\mathrm{Av}(x_1, \ldots, x_n; q) - \mathbb{E}_{\mu_1}[q]| > \varepsilon \right) \leq \exp\left(-\frac{n\delta^2}{2} \right) < 1.$$

If Q is a VC class, such a ε can always be selected for large enough n, as $\lim_{n\to\infty} \frac{r_Q(n)}{n} = 0$. $\square$

3.2.2 Covering Numbers and ε-Approximations

In this section, we follow the covering number approach of [ABC97] to bound the sizes of ε-approximations, in a measure-theoretic generality suitable for Keisler measures, as in [Sim15, Section 7.5].

Definition 3.2.15. For $\bar{x} \in X^n$, let $\mathcal{N}(Q, \bar{x}, \varepsilon)$ be the l_∞-distance covering number of the set $Q(\bar{x})$ - that is, the minimum size of a set $A \subseteq [0,1]^n$ such that for all $q \in Q(\bar{x})$, there is $a \in A$ with $d(a, q) \leq \varepsilon$, with d denoting the l_∞ distance.

Let $\mathcal{N}_{Q,\varepsilon}(n) = \sup_{\bar{x} \in X^n} \mathcal{N}(Q, \bar{x}, \varepsilon)$.

To bound the covering number, we will use variations on the VC-dimension:

Definition 3.2.16. Let $Q \subseteq [0,1]^X$ be a class of functions, and $\varepsilon > 0$.

Let the ε-VC-dimension of Q, $\mathrm{vc}_\varepsilon(Q)$, be the supremum of the VC-dimensions $\mathrm{vc}(Q_{r,r+\varepsilon})$ where $r \in [0, 1-\varepsilon]$.

Define the *fat-shattering* dimension of Q, denoted $\mathrm{fs}_\varepsilon(Q)$, to be the maximal cardinality (or ∞ if there is no maximum) of a finite set $A \subseteq X$ such that there is a function $f : A \to [0,1]$ such that $(Q - f)_{-\varepsilon,\varepsilon}$ shatters A.

Ben Yaacov [Ben09] has shown that Q is a VC-class if and only if $\mathrm{vc}_\varepsilon(Q)$ is finite for all $\varepsilon > 0$. The fat-shattering dimension also corresponds (up to constants) to the idea of "determining a d-dimensional ε-box" in [Ben09], where it is also shown that Q is a VC-class if and only if $\mathrm{fs}_\varepsilon(Q)$ is finite for all $\varepsilon > 0$. The following fact relates the two dimensions more concretely:

Fact 3.2.17 ([ABC97, Lemma 2.2])**.** *Let $Q \subseteq [0,1]^X, \varepsilon > 0$. Then*

$$\mathrm{vc}_{2\varepsilon}(Q) \leq \mathrm{fs}_\varepsilon(Q) \leq \left(2\lceil\frac{1}{\varepsilon}\rceil - 1\right)\mathrm{vc}_\varepsilon(Q).$$

The fat-shattering dimension is useful for the following lemma. (The version given here is stated in the proof of the cited lemma.)

Fact 3.2.18 ([ABC97, Lemma 3.5])**.** *Let $\mathrm{fs}_{\varepsilon/4}(Q) \leq d$. Then*

$$\mathcal{N}_{Q,\varepsilon}(n) \leq 2\left(\frac{4n}{\varepsilon^2}\right)^{d\log(2en/d\varepsilon)} = n^{O_{d,\varepsilon}(\log n)}.$$

We can deduce from this and Fact 3.2.17 that the bound of $\mathbb{N}_{Q,\varepsilon}(n) = n^{O_{d,\varepsilon}(\log n)}$ also holds when $\mathrm{vc}_{\varepsilon/4}(Q) \leq d$, although with a different constant.

We can also bound the VC-dimension from the covering numbers.

Lemma 3.2.19. *Let $Q \subseteq [0,1]^X, 0 \leq r < s \leq 1, 0 < \varepsilon < \frac{s-r}{2}$. Then*

$$\pi_{Q_{r,s}}(n) \leq \mathcal{N}_{Q,\varepsilon}(n).$$

Proof. Let $A \subseteq X$ be such that $|A| = n$ and $|Q_{r,s} \cap A|$ is maximized, so $|Q_{r,s} \cap A| = \pi_{Q_{r,s}}(n)$. Then for each subset $B \subseteq A$ in $|Q_{r,s} \cap A|$, there is some $q_B \in Q$ with $q_B(a) \leq r$ for $a \in B$ and $q_B(a) \geq s$ for $a \in A \setminus B$. The points $(q_B(a) : a \in A)$ for $B \in Q_{r,s} \cap A$ thus each have ℓ_∞-distance at least $s - r$ from each other. Thus no two of them can lie in the same ε-ball in that metric, and the covering number must be at least $\pi_{Q_{r,s}}(n)$. $\square$

In particular, any sub-exponential bound on the covering number for each ε implies that Q is a VC class of functions.

Alon et al. use the covering number bound to prove the existence of ε-approximations using the following fact:

Fact 3.2.20 ([ABC97, Lemma 3.4])**.** *Let $\varepsilon > 0, n \geq \frac{2}{\varepsilon^2}, Q \subseteq [0,1]^X$, and let $\bar{x} = (x_1, \dots, x_n)$ be a tuple of i.i.d. random variables with values in X. Then subject to measurability constraints which are satisfied if the probability distribution of each x_i is finitely supported,*

$$\mathbb{P}\left[\sup_{q \in Q} (\mathrm{Av}(\bar{x}, q) - \mathbb{E}[q(x_1)]) > \varepsilon\right] \leq 12n\mathcal{N}_{Q,\varepsilon/6}(2n) \exp\left\{-\frac{\varepsilon^2 n}{36}\right\}.$$

Combining the previous two facts gives a bound on the minimum size of ε-approximations for Q with respect to any finitely-supported probability measure μ:

Fact 3.2.21 ([ABC97, Theorem 3.6])**.** *Let $Q \subseteq [0,1]^X$ satisfy $\mathrm{fs}_{\varepsilon/24}(Q) \leq d$. Then if μ is a finitely-supported probability measure on X, for all $\varepsilon, \delta > 0$, if $\bar{x} = (x_1, \dots, x_n)$ consists of i.i.d. random variables with distribution given by μ, we have*

$$\mathbb{P}\left[\sup_{q \in Q} (\mathrm{Av}(\bar{x}, q) - \mathbb{E}[q(x_1)]) > \varepsilon\right] \leq \delta$$

for

$$n = O\left(\frac{1}{\varepsilon^2}\left(d \ln^2 \frac{d}{\varepsilon} + \ln \frac{1}{\delta}\right)\right).$$

In Chapter 4, we will derive version of Facts 3.2.20 and 3.2.21 for generically stable

Keisler measures in continuous logic, bounding the sizes of ε-approximations for definable predicates with respect to a fixed generically stable Keisler measure.

3.2.3 Transversals and ε-nets

While ε-approximations lend themselves naturally to real-valued function classes, there is another way of approximating set systems with respect to measures that more naturally generalizes to fuzzy set systems: ε-nets. In this subsection, we will use a fuzzy set system generalization of ε-nets to prove fuzzy versions of a bound on transversal numbers and to prove a fractional Helly property and (p, q)-theorem for fuzzy set systems. This generalizes the classical combinatorial results for set systems described in [Mat02, Chapter 10].

Definition 3.2.22. Let $\mathcal{F}$ be a fuzzy set system on X, μ a probability measure on X and $\varepsilon > 0$. An ε-net for $\mathcal{F}$ with respect to μ is a subset $A \subseteq X$ such that for every $(S_+, S_-) \in \mathcal{F}$ such that $\mu(S_+) \geq \varepsilon$, $A \not\subseteq S_-$.

In order to construct ε-nets out of ε-approximations, we will need to define a construction that crops a function class down to a particular interval. Let $f_{r,s} : [0, 1] \to [0, 1]$ be the piecewise linear function given by

$$f_{r,s}(x) = \begin{cases} r & x \leq r \\ x & r \leq x \leq s \\ s & x \geq s \end{cases}.$$

Now let $Q^{r,s} = \{f_{r,s} \circ q : q \in Q\}$. If Q is a VC-class, then $Q^{r,s}$ will be one as well, and in fact, for any $r' < s'$, the VC-dimension of $(Q^{r,s})_{r',s'}$ will be at most the VC-dimension of $Q_{r',s'}$, and for all n, $g_{Q^{r,s}}(n) \leq g_Q(n)$.

Lemma 3.2.23. *For any $\varepsilon > 0$, $0 \leq r < s \leq 1$, if $Q \subset [0, 1]^X$ is a class of functions, μ is a probability measure on X, and $\bar{A} = (a_1, \ldots, a_n)$ is a δ-approximation for $Q^{r,s}$, where*

$\delta < (s-r)\varepsilon$, then $A = \{a_1, \ldots, a_n\}$ is a ε-net for $Q_{r,s}$ with respect to μ.

Proof. Fix ε, g, Q, and μ, let $\delta < (s-r)\varepsilon$, and let $\bar{A}$ be a δ-approximation for $Q^{r,s}$. Now let $q \in Q$, and assume that $\mu(q_{\leq r}) = \mu((f_{r,s} \circ q)_{\leq r}) \geq \varepsilon$. Then $\mathbb{E}_\mu[f_{r,s} \circ q] \leq s - (s-r)\varepsilon$, and accordingly $\mathrm{Av}(a_1, \ldots, a_n; q) \leq s - (s-r)\varepsilon + \delta < s$, so there exists at least one a_i with $q(a_i) < s$. □

Theorem 3.2.24. *For any $\varepsilon > 0$, $0 \leq r < s \leq 1$ and $g : \mathbb{N} \to [0, \infty)$ such that $g(n) = o(1)$, there is $N = N((s-r)\varepsilon, g)$ such that if $Q \subset [0,1]^X$ is a class of functions such that $\frac{r_Q(n)}{n} \leq g(n)$ for all n, and μ is a finitely-supported probability measure on X, there is an ε-net A for $Q_{r,s}$ with respect to μ with $|A| \leq N$.*

Proof. Let N, δ be such that $g(N) < \delta < (s-r)\varepsilon$. Using Theorem 3.2.14, we can find a δ-approximation $\bar{A}$ for $Q^{r,s}$, which by Lemma 3.2.23 is a ε-net for $Q_{r,s}$. □

This statement is easy to deduce from the VC-Theorem, but it only applies to fuzzy set systems derived from classes of functions. With a direct probabilistic argument, adapted from the classical proof by Haussler and Welzl ([Mat02, Theorem 10.2.4]), we can bound the size on ε-nets for any VC fuzzy set system based only on ε and the VC-dimension, up to some measurability assumptions. In Chapter 4, we will prove that this also holds in the context of generically stable Keisler measures.

Theorem 3.2.25. *For any $\varepsilon > 0$ and $d \in \mathbb{N}$, there is $N = O(d\varepsilon^{-1} \log \varepsilon^{-1})$ such that if $\mathcal{F}$ is a fuzzy set system on X with VC-dimension at most d, and μ is a finitely-supported probability measure on X, there is an ε-net A for $\mathcal{F}$ with respect to μ with $|A| \leq N$.*

If μ is not necessarily finitely-supported, then the result still holds, assuming the following *events are measurable:*

$$S_\pm : S \in \mathcal{F}$$

$$E_0(x_1, \ldots, x_N) = \bigcup_{S\in\mathcal{F},\mu(S_+)\geq\varepsilon} \bigcap_{i=1}^{N} [x_i \in S_-]$$

$$E_1(x_1, \ldots, x_N, y_1, \ldots, y_N) = \bigcup_{S\in\mathcal{F},\mu(S_+)\geq\varepsilon} \left(\bigcap_{i=1}^{N} [x_i \in S_-]\right) \cap \left(\bigcup_{I\subseteq\{1,\ldots,N\},|I|\geq\lceil\frac{N\varepsilon}{2}\rfloor} \bigcap_{i\in I} [y_i \in S_+]\right)$$

These will be measurable, for instance, if we assume that $\mathcal{F}$ is a countable set system of measurable fuzzy sets, or that μ is a Borel probability measure on a topological space X where for each $S \in \mathcal{F}$, S_+ and S_- are both open.

Proof. This proof generalizes the argument by Haussler and Welzl used in [Mat02, Theorem 10.2.4].

Let $N = Cd\varepsilon \log(\varepsilon^{-1})$, with C to be determined later. Let $\bar{A} = (a_1, \ldots, a_N)$ be a tuple of independently selected variables with values in X and distribution μ. Then let E_0 be the event that $\{a_1, \ldots, a_N\}$ is not a ε-net. We wish to show that for large enough C, $\mathbb{P}[E_0] < 1$, so there must exist a ε-net of size N. We can express $E_0 = \bigcup_{S\in\mathcal{F},\mu(S_+)\geq\varepsilon} \bigcap_{i=1}^{N} [a_i \in S_-]$. If either $\mathcal{F}$ or the support of μ is countable, then this is clearly measurable, and if each S_- is open, then this is open.

Let $\bar{B} = (b_1, \ldots, b_N)$ be a second tuple of random variables, independent of $\bar{A}$ with the same distribution. Let E_1 be the event that there exists $S \in \mathcal{F}$ such that $\mu(S_+) \geq \varepsilon$, for each $1 \leq i \leq N$, $a_i \in S_-$ *and* there are at least $k = \lceil\frac{N\varepsilon}{2}\rfloor$ values of i such that $b_i \in S_+$. We will show that $\mathbb{P}[E_1] \geq \frac{1}{2}\mathbb{P}[E_0]$, and then we will show that $\mathbb{P}[E_1] < \frac{1}{2}$. We can see that E_1 is measurable for the same reasons as E_0 is.

To show that $\mathbb{P}[E_1] \geq \frac{1}{2}\mathbb{P}[E_0]$, we will fix $\bar{A}$, select B conditioned on $\bar{A}$, and show that $\mathbb{P}[E_1|\bar{A}] \geq \frac{1}{2}\mathbb{P}[E_0|\bar{A}]$. If $\{a_1, \ldots, a_N\}$ is not a ε-net, then $\mathbb{P}[E_0|\bar{A}] = 0$, and as $E_1 \subseteq E_0$,

$\mathbb{P}[E_1|\bar{A}] = 0$. Assume $\{a_1, \ldots, a_N\}$ is an ε-net. Then if I_i for $1 \leq i \leq N$ are the indicator random variables for $b_i \in S_+$, and $I = I_1 + \cdots + I_N$, we have that $\mathbb{P}[E_1|\bar{A}] = \mathbb{P}[I \geq k]$. The I_is are i.i.d. random variables, equalling 1 with probability $\mu(S_+) \geq \varepsilon$. By a standard Chernoff tail bound for binomial distributions, we have that $\mathbb{P}[X \geq k] \geq \frac{1}{2} = \frac{1}{2}\mathbb{P}[E_0|\bar{A}]$. Thus in general, $\mathbb{P}[E_1] \geq \frac{1}{2}\mathbb{P}[E_0]$.

To show that $\mathbb{P}[E_1] < \frac{1}{2}$, we will instead condition on the value of the multiset $D = \{a_1, \ldots, a_N, b_1, \ldots, b_N\}$. Select $\bar{A}$ and $\bar{B}$ by permuting D uniformly at random. All events will be measurable as this probability space is finite. For any fixed fuzzy set $S \sqsubseteq X$, let E_S be the conditional event that $\bar{A} \subseteq S_-$ and there are at least k values of i such that $b_i \in S_+$, given the choice of multiset D. We find that if S' is a strong disambiguation of $S \cap D$, then $E_S \subseteq E_{S'}$, so if $\mathcal{F}'$ is a strong disambiguation of $\mathcal{F}$ restricted to D, we have that

$$E_1|D = \bigcup_{S \in \mathcal{F} : \mu(S_+) \geq \varepsilon} E_S \subseteq \bigcup_{S \in \mathcal{F}} E_S \subseteq \bigcup_{S' \in \mathcal{F}'} E_{S'}.$$

Now we apply Lemma 3.2.6, and find a strong disambiguation $\mathcal{F}'$ with $|\mathcal{F}'| = (2N)^{O(d \log(2N))}$, or as we will prefer later, there is C' such that $|\mathcal{F}'| \leq (2N)^{C'(d \log(2N))}$. We find that for each $S' \in \mathcal{F}'$, if $|D \cap S'| < k$, then $\mathbb{P}[E_{S'}] = 0$, and that if $|D \cap S'| \geq k$, then $\mathbb{P}[E_{S'}]$ is the probability that when a set of N elements of D is selected at random, the set is disjoint with S'. This is at most

$$\frac{\binom{2N - |D \cap S'|}{N}}{\binom{2N}{N}} \leq \frac{\binom{2N-k}{N}}{\binom{2N}{N}} \leq \left(1 - \frac{k}{2N}\right)^N \leq e^{-(k/2N)N} = \varepsilon^{Cd/4}.$$

Now we bound the probability of the union, letting C' be the constant of :

$$\begin{aligned}
\mathbb{P}[E_1|D] &\leq \sum_{S' \in \mathcal{F}'} \mathbb{P}[E_{S'}] \\
&\leq |\mathbb{F}'|\varepsilon^{-Cd/4} \\
&\leq (2N)^{C'(d\log(2N))}\varepsilon^{Cd/4} \\
&= \left((2Cd\varepsilon^{-1}\log\varepsilon^{-1})^{C'(\log\left(2Cd\varepsilon^{-1}\log\varepsilon^{-1}\right))}\varepsilon^{C/4}\right)^{D}
\end{aligned}$$

While this expression is somewhat complicated, it is still clear that an increasing quasipolynomial function of C times a decreasing exponential of C will limit to 0, so for large enough C, we find that $\mathbb{P}[E_1|D] < \frac{1}{2}$. $\square$

We apply this first to transversal numbers. We will only apply these to actual discrete set systems, so the definitions here are the same as in [Mat02].

Definition 3.2.26. Let $\mathcal{F}$ be a set system on a set X. A *transversal* of $\mathcal{F}$ is a set $T \subseteq X$ such that for all $S \in \mathcal{F}$, $T \cap S \neq \emptyset$. The *transversal number* of $\mathcal{F}$, $\tau(\mathcal{F})$ is the minimum size of a finite transversal $T \subseteq X$, if it exists.

A *fractional transversal* of $\mathcal{F}$ is a finitely-supported function $t : X \to [0, 1]$ such that for all $S \in \mathcal{F}$, $\sum_{s \in S} t(s) \geq 1$. The *fractional transversal number* of $\mathcal{F}$, $\tau^*(\mathcal{F})$ is the minimum size of a fractional transversal t, if it exists, with the size of t being defined as $\sum_{x \in X} t(x)$.

We can now use Theorem 3.2.25 on the existence of ε-nets of fuzzy set systems to bound the transversal number of the outer set system in terms of the fractional transversal nmuber of the inner set system.

Theorem 3.2.27. *Let $d \in \mathbb{N}$, and let $t > 0$. There is $T = T(t, d)$ such that if $\mathcal{F}$ is a finite fuzzy set system on X with VC-dimension at most d, and $\tau^*(\mathcal{F}_i) \leq t$, then $\tau(\mathcal{F}_o) \leq T$.*

Proof. As $\mathcal{F}$ is finite, we may assume that there is an optimal fractional transversal f :

$X \to [0,1]$ for $\mathcal{F}_i$ of finite support. This f leads to a probability measure μ on X defined by $\mu(\{x\}) = \frac{f(x)}{\tau^*(\mathcal{F}_i)}$ for all $x \in X$, which itself has finite support.

Now we claim that any $\frac{1}{t}-$net for the fuzzy set system $\mathcal{F}$ is a transversal. If indeed a set $A \subseteq X$ is a $\frac{1}{t}-$net, then for any $S \in \mathcal{F}$ such that $\mu(S_+) \geq \frac{1}{\tau^*(\mathcal{F}_i)}$, we also have $\mu(S_+) \geq \frac{1}{t}$, and thus $A \not\subseteq S_-$ by the $\frac{1}{t}-$net property. As for every $S \in \mathcal{F}$, we have $\mu(S_+) = \frac{\sum_{x \in S_+} f(x)}{\tau^*(\mathcal{F}_i)} \geq \frac{1}{\tau^*(\mathcal{F}_i)}$ by the assumption that f is a fractional transversal, A must be a transversal for $\mathcal{F}_o$.

Thus we can simply let T be large enough that there must be a $\frac{1}{t}$-net of size at most T. By Theorem 3.2.25, we can choose T depending only on d and t. $\square$

We now use ε-nets for fuzzy relations to give a bound on a fuzzy fractional Helly number. This generalizes the results of [Mat04], using the following definition of a fractional Helly number for a fuzzy relation:

Definition 3.2.28. We say that a fuzzy relation $R \sqsubseteq X \times Y$ has *fractional Helly number k* when for every $\alpha > 0$, there is a $\beta > 0$ such that if $b_1, \ldots, b_n \in Y$ are such that $\bigcap_{i \in I} R_+^{b_i} \neq \emptyset$ for at least $\alpha\binom{n}{k}$ sets $I \in \binom{[n]}{k}$, then there is $J \subseteq [n]$ with $|J| \geq \beta n$ such that $\bigcap_{j \in J}(X \setminus R_-^{b_j}) \neq \emptyset$.

We can bound the fractional Helly number of a fuzzy relation by its dual VC-density, that is, the exponent of growth of the dual shatter function (the shatter function of the fuzzy set system S_X on Y).

Theorem 3.2.29. *[Generalizing [Mat04]] Let $R \sqsubseteq X \times Y$ be a fuzzy set system with $\pi_{R_X}(n) = o(n^k)$. Then R has fractional Helly number k.*

Proof. This proof follows Matousek's probabilistic argument closely, but it is important to keep track of when an element S of the set system is replaced with S_+ or S_-.

Let $\alpha > 0$. Fix m such that $\pi_{R_X}(m) < \frac{\alpha}{4}\binom{m}{k}$, and set $\beta = \frac{1}{2m}$. If $n \leq 2m^2 = \frac{m}{\beta}$, then for any $b_1, \ldots, b_n \in Y$, all that is required to find a set $J \subseteq [n]$ with $|J| \geq \beta n$ such that

$\bigcap_{j\in J}(X\setminus R_-^{b_j})\neq\emptyset$ is a singleton $J=\{b_j\}$ with $R_+^{b_j}\neq\emptyset$. Thus it suffices to show that for $n\geq 2m^2=\frac{m}{\beta}$, if $b_1,\dots,b_n\in Y$ are such that $\bigcap_{i\in I}R_+^{b_i}\neq\emptyset$ for at least $\alpha\binom{n}{k}$ sets $I\in\binom{[n]}{k}$, then there is $J\subseteq[n]$ with $|J|\geq\beta n$ such that $\bigcap_{j\in J}(X\setminus R_-^{b_j})\neq\emptyset$.

For contradiction, suppose that $b_1,\dots,b_n\in Y$ satisfy these assumptions, but $\bigcap_{j\in J}(X\setminus R_-^{b_j})=\emptyset$ for each J with $|J|\geq\beta n$. We say that a pair (J,I) with $J\in\binom{[n]}{m}$, $I\in\binom{J}{k}$ is *good* when there is $a\in X$ with $a\in R_+^i$ for each $i\in I$ and $a\in R_-^j$ for each $j\in J\setminus I$. For any given J, the set of Is such that (J,I) is good is exactly $R_X\cap J$, and by definition, $|R_X\cap J|\leq\pi_{R_X}(m)$, As by assumption, $\pi_{R_X}(m)<\frac{\alpha}{4}\binom{m}{k}$, the probability that (J,I) is good with a randomly chosen I is less than $\frac{\alpha}{4}$.

We now contradict this bound and show that the probability that a randomly chosen (J,I) is good is at least $\frac{\alpha}{4}$. Start by choosing $I\in\binom{[n]}{k}$. By assumption, the probability that there is $a\in X$ with $a\in R_+^i$ for each $i\in I$ is at least α. For each such i, fix an a, and we will show that when we choose $J\setminus I\in\binom{[n]\setminus I}{m-k}$ at random, $a\in R_-^j$ for each $j\in J\setminus I$ with probability at least $\frac{1}{4}$. By assumption, $a\notin R_-^b$ for less than βn values of $b\in\{b_1,\dots,b_n\}$, so the probability that $a\in R_-^j$ for some j is at least

$$\frac{\left(\binom{\lceil(1-\beta)n\rceil}{m-k}\right)}{\binom{n-k}{m-k}}\geq\prod_{i=0}^{m-k-1}\frac{(1-\beta)n-i}{n-i}\geq\prod_{i=0}^{m-1}\frac{(1-\beta)n-m}{n-m}\geq\left(\frac{(1-\beta)n-m}{n-m}\right)^m.$$

Recalling that $m\leq\beta n$ and $\beta=\frac{1}{2m}$, we see that this is

$$\left(1-\frac{\beta n}{n-m}\right)^m\geq(1-2\beta)^m=\left(1-\frac{1}{m}\right)^m\geq\frac{1}{4}.$$

□

We now recall the (p,q) property, a property of classical set systems. We will use VC-dimension of fuzzy set systems to prove a (p,q)-theorem generalizing that of [AK92].

Definition 3.2.30. Let $\mathcal{F}$ be a set system on a set X. Then $\mathcal{F}$ has the (p,q) *property* when for any $S_1, \dots, S_p \in \mathcal{F}$, there are $i_1, \dots, i_q$ such that $\bigcap_{j=1}^q S_{i_j} \neq \emptyset$.

If $p = q$, then the (p,p) property just states that any p elements of a set system have nonempty intersection. We can now adapt the classical proof of the (p,q)-theorem, starting with the bound on the fractional transversal number.

Theorem 3.2.31. *[Generalizes [AK92]] Let $p \geq q \geq d+1$ and $0 \leq r < s \leq 1$. Let $\mathcal{F}$ be a finite fuzzy set system with $\mathrm{vc}^*(\mathcal{F}) \leq d$. If $\mathcal{F}_i$ has the (p,q)-property, then $\tau^*(\mathcal{F}_o) \leq N$, where $N = N(p,q,d)$.*

Proof. We first note that $\tau^*(\mathcal{F}_o) = \nu^*(\mathcal{F}_o)$ when $\mathcal{F}$ is finite, so it suffices to bound $\nu^*(\mathcal{F}_o)$. Now let $f : \mathcal{F} \to [0,1]$ be such that $S_- \mapsto f(S)$ is an optimal fractional packing for $\mathcal{F}_o$, which takes rational values, as $\mathcal{F}$ is finite. (See [Mat02, Chapter 10].)

Let D be a common denominator so that $m(S) := Df(S)$ is always an integer. We now define a new fuzzy relation by letting Y be the set of pairs $\{(S,i) : S \in \mathcal{F}, 1 \leq i \leq m(S)\}$, and defining $R_m \subseteq X \times Y$ by $(R_m)_+ = \{(a,S,i) : a \in S_+\}$ and $(R_m)_- = \{(a,S,i) : a \in S_-\}$. Then the inner set system $(R_m)_i^Y$ has the (p',q)-property, where $p' = p(d-1)+1$. Let $N = |Y_m| = D\nu^*(\mathcal{F}_o)$.

We claim there exists some $a \in X$ such that $a \notin (R_m^{(S,i)})_-$ for at least βN pairs (S,i) for some β depending only on p and d. By the fractional Helly theorem, as this class also has VC-codensity at most d, it suffices to find $\alpha = \alpha(p,d) > 0$ such that if for at least $\alpha\binom{N}{k}$ sets $I \in \binom{[N]}{k}$, there is some $a \in R_m^{y_i}$ for each $i \in I$. Every set of p' sets in this collection contains at least one set of $(d+1)$ sets with nonempty intersection, and each such set of $(d+1)$ sets is contained in $\binom{N-d+1}{p-d+1}$ sets of p sets. Thus the number of intersecting sets of $(d+1)$ sets from this collection is at least

$$\frac{\binom{N}{p}}{\binom{N-d+1}{p-d+1}} \geq \alpha\binom{N}{d+1}$$

for some $\alpha = \alpha(p,d) > 0$.

Now since we have $a \in X$ such that $R_m^{(S,i)}$ is not false for at least βN pairs (S,i), we have that

$$1 \geq \sum_{S \in \mathcal{F}; a \in S_-} f(S) \geq \sum_{S \in \mathcal{F}; a \in S_-} \frac{m(S)}{D} \geq \frac{1}{D} \beta N = \beta \nu^*(\mathcal{F}_o)$$

so $\nu^*(\mathcal{F}_o) \leq \frac{1}{\beta}$. □

This (p, q) theorem can now be combined with the earlier bound relating the transversal and fractional transversal numbers (Theorem 3.2.27). In this process, we end up looking at three nested set systems, using the properties of the innermost to bound the fractional transversal number of the middle set system, and then using that to bound the transversal number of the outermost set system. To simplify this presentation, we will only give this corollary in the case where the three nested set systems come from the same set of functions, which is exactly the setup we will need for model-theoretic applications:

Corollary 3.2.32. *For all $0 \leq r < t < s \leq 1$, $d_1, d_2 \in \mathbb{N}$, and $p \geq q \geq d_1 + 1$, there exists $N = N(d_1, d_2, p, q) \in \mathbb{N}$ such that if $Q \subseteq [0,1]^X$ is a finite function class such that $\mathrm{vc}^*(Q_{r,t}) \leq d_1$ and $\mathrm{vc}(Q_{t,s}) \leq d_2$, then for all finite Q, if the set system $Q_{\leq r}$ has the (p, q)-property, then $\tau(Q_{<s}) \leq N$.*

Proof. We will first apply Theorem 3.2.31 to the set system $Q_{\leq r}$ to bound $\tau^*(Q_{<t})$, then enlarge the sets slightly without increasing the fractional transversal number, bounding $\tau^*(Q_{\leq t})$, and finally apply Theorem 3.2.27 to bound $\tau(Q_{<s})$.

Fix $p \geq q \geq d_1 + 1$. We will also have $q \geq \mathrm{vc}^*(Q_{r,t}) + 1$. Applying Theorem 3.2.31 now gives us an N_0 not depending on Q such that $\tau^*(Q_{<t}) \leq N_0$. As adding to the sets in this set system cannot increase the fractional transversal number, we find that $\tau^*(Q_{\leq t}) \leq \tau^*(Q_{<t}) \leq N_0$.

We now look at the fuzzy set system $Q_{t,s}$. Thus we know that $\tau^*(Q_{\leq t}) \leq N_0$, and it suffices to find N such that $\tau(Q_{<s}) \leq N$. As $\mathrm{vc}(Q_{t,s}) \leq d_2$, Theorem 3.2.27 gives us an $N = N(N_0, d_2)$ such that $\tau((Q_0^{t,s})_{<1}) \leq N$. □

3.3 Model-Theoretic Preliminaries and Notation

We refer to [BBH08] for an introduction to metric structures and continuous logic, although we will need a few additional pieces of notation and background, provided in this section. Throughout this chapter, let T be a theory in continuous logic, using the language $\mathcal{L}$. We fix a monster model $\mathcal{U} \vDash T$, and will use M to denote a submodel of $\mathcal{U}$, small in the sense that $\mathcal{U}$ is $|M|^+$-saturated.

In continuous logic, it is natural to deal with variable tuples of countably infinite length. As if x, y are infinite tuples, $|x|$ equals $|x, y|$, we shall just refer to the relevant cartesian products of a set M as M^x and $M^x \times M^y$, rather than $M^{|x|}$ or $M^{|x,y|}$.

In classical model theory, we frequently use the notation $\phi(M; b)$ to indicate the subset of M^x defined by the formula $\phi(x; y)$ using the parameter $b \in M^y$. As this chapter will deal with metric structures, where the definable predicate $\phi(x; y)$ can take on any value in $[0, 1]$, $\phi(M; b)$ will be defined as the subset of M^x on which $\phi(x; b) = 0$. For other $r \in [0, 1]$, we will use the notations $\phi_{\leq r}(M; b)$ and $\phi_{\geq r}(M; b)$ to denote the sets where $\phi(x; b) \leq r$ and $\phi(x; b) \geq r$. Given any condition (an inequality or equality of definable predicates), we will use notation such as $[\phi(x) \geq r]$ to denote the subset of a type space $S_x(A)$ where that condition is true.

3.3.1 Pairs

In classical model theory, we frequently add a predicate to pick out a specific subset of a structure, thus making that set definable in the expansion. In continuous logic, a closed subset of a metric structure is considered definable when its distance predicate is definable. [BBH08, Def 9.16] These definable sets are exactly the sets that can be quantified over when constructing definable predicates. Thus to pick out a particular subset, we restrict our attention to closed subsets, and add a predicate for the distance to that closed subset.

Definition 3.3.1. If M is a metric $\mathcal{L}$-structure, and $A \subseteq M^x$ is closed, then let (M, A)

be the expansion of M to the language $\mathcal{L}_P$, adding a relation symbol P interpreted as $P(x) = \mathrm{dist}(x, A)$.

This is a valid metric structure, because $\mathrm{dist}(x, A)$ is bounded and 1-Lipschitz.

Per [BBH08, Theorem 9.12], there are axioms indicating that a predicate is the distance predicate of a closed set, so any structure (N, B) elementary equivalent to (M, A) will be an expansion of some N elementarily equivalent to M by a distance predicate for a closed set $B \subseteq N^x$. Sometimes if $y = (x_1, \ldots, x_n)$ or $y = (x_1, x_2, \ldots)$, we will use $P(y)$ to denote a definable predicate indicating that $x_i \in A$ for each i. If $y = (x_1, \ldots, x_n)$, this can straightforwardly be $P(y) = \max_{i=1}^n P(x_i)$, but if $y = (x_1, x_2, \ldots)$, we may use $P(y) = \sum_{i \in \mathbb{N}} 2^{-i} P(x_i)$, and we will still have $P(\bar{a}) = 0$ if and only if $P(a_i) = 0$ for all i.

If we wish to define two definable subsets, we will say that (M, A, B) is the expansion adding a distance predicate P to A and a distance predicate Q to B.

3.3.2 Coding Tricks

Lemma 3.3.2. *Let $\phi_1(x; y), \ldots, \phi_n(x; y)$ be a series of definable predicates, and $A \subseteq \mathcal{U}^y$ be such that $|A| \geq 2$. Then there is a single definable predicate $\phi(x; y_1, y_2, \ldots, y_k)$ such that for every $1 \leq i \leq n$ and $a \in A$, there is some $\bar{a} \in A^k$ such that $\phi_i(x; a) = \phi(x; \bar{a})$ for all x.*

Proof. Let $a_1, a_2 \in A$ be distinct. Then let $k = 2n + 1$ and let

$$\phi(x; y_1, \ldots, y_{2^n}) = \sum_{i=1}^{n} \frac{d(y_{2i-1}, y_{2i})}{d(a_1, a_2)} \phi_i(x, y_k).$$

Then for any $1 \leq i \leq n$ and $a \in A$, we can let $b_k = a$, and choose $b_1, \ldots, b_{2n} \in \{a_1, a_2\}$ so that $b_{2j-1} = b_{2j}$ if and only if $j \neq i$. Then $\phi(x; b_1, \ldots, b_k) = \phi_i(x; a)$. $\square$

3.3.3 Other Facts

The following application of the compactness theorem for metric structures will come up in a few proofs later on in the chapter.

Lemma 3.3.3. *Let $A \subseteq \mathcal{U}$. Let $p(x)$ be a partial A-type, let $q(x)$ be a partial $\mathcal{U}$-type, and let $\phi(x)$ be a $\mathcal{U}$-definable predicate. Then if $p(x) \cup q(x)$ implies $\phi(x) = 0$, there is an A-definable predicate $\theta(x)$ such that $p(x)$ implies $\theta(x) = 0$, and $q(x)$ implies $\phi(x) \leq \theta(x)$.*

Proof. This is a combination of compactness and the proof of [BBH08, Prop. 7.14].

Write $p(x) = \{\psi(x) = 0 : \psi \in \Psi\}$. For every $n \in \mathbb{N}$, we see that $\{\psi(x) \leq \delta : \delta > 0, \psi \in \Psi\} \cup q(x) \cup \{\phi(x) \geq 2^{-n}\}$ is inconsistent, so by compactness, there is a subtype $p_n(x) \subseteq p(x)$ of the form $\{\psi(x) \leq \delta_n : \psi \in \Psi_n\}$ for some δ and some finite $\Psi_n \subseteq \Psi$ such that $p_n(x) \cup q(x) \cup \{\phi(x) \geq 2^{-n}\}$ is inconsistent. Thus if $\theta_n(x) = \max_{\psi \in \Psi_n} \psi(x)$, we see that $p(x)$ implies $\theta_n(x) = 0$, and $\theta_n(x) \leq \delta_n$ implies $\phi(x) < 2^{-n}$. Thus also $p(x)$ implies $\sum_{n \in \mathbb{N}} 2^{-n}\theta_n(x) = 0$, and for all n, $q(x)$ and $\sum_{n \in \mathbb{N}} 2^{-n}\theta_n(x) \leq 2^{-n}\delta_n$ implies $\phi(x) < 2^{-n}$.

Thus by [BBH08, Prop. 2.10], there is an increasing continuous function $\alpha : [0, 1] \to [0, 1]$ such that on the subspace of $S_x(\mathcal{U})$ realizing $q(x)$, we have $\phi(x) \leq \alpha\left(\sum_{n \in \mathbb{N}} 2^{-n}\theta_n(x)\right)$. Thus we may define $\theta(x) = \alpha\left(\sum_{n \in \mathbb{N}} 2^{-n}\theta_n(x)\right)$, and we find $q(x)$ implies $\phi(x) \leq \theta(x)$. □

Lemma 3.3.4. *Let $A \subseteq \mathcal{U}$, let $p(x) \subseteq S_x(A)$ be a partial type, and let $\phi(x)$ be a $\mathcal{U}$-definable predicate such that for every global type $q(x) \in S_x(\mathcal{U})$ extending $p(x)$, $q|_A$ implies $|\phi(x) = r_p|$ for some r_p. Then there is an A'-formula $\psi(x)$ such that $p(x)$ implies $\psi(x) = \phi(x)$.*

Proof. The restriction of parameters map $S_x(\mathcal{U}) \to S_x(A)$ is a continuous surjection of compact Hausdorff spaces, and is thus a quotient map. The set $[p(x)]$ in either space is closed in $S_x(A)$, and we have assumed that $\phi(x)$, restricted to $[p(x)] \subseteq S_x(\mathcal{U})$, lifts to a function from $[p(x)] \subseteq S_x(A)$ to $\mathbb{R}$, which is continuous by the quotient property. This continuous function extends to all of $S_x(A)$ by Tietze's extension theorem, and that continuous function is an A-definable predicate, $\psi(x)$. □

The following fact about partitions of unity (see [Rud87, Theorem 2.13]) will come up repeatedly in this chapter:

Fact 3.3.5. *Let K be a compact Hausdorff space, and let $U_1, \dots, U_n$ be open sets that cover K. Then there are functions $u_1, \dots, u_n : K \to [0,1]$ such that*

- *for all $x \in K$ and for all i, $0 \leq u_i(x) \leq 1$*
- *for all $x \in K$, $u_1(x) + \cdots + u_n(x) = 1$*
- *for all i, the support of u_i is contained in U_i for each i.*

We will also use the notion of a *forced limit* from [BU10], in order to carefully define a predicate as a limit of formulas that may not necessarily converge uniformly.

Definition 3.3.6. Let $(a_n : n < \omega)$ be a sequence in $[0,1]$. Define the sequence $(a_{\mathcal{F}\mathrm{lim},n} : n < \omega)$ recursively:

$$a_{\mathcal{F}\mathrm{lim},0} = a_0$$

$$a_{\mathcal{F}\mathrm{lim},n+1} = \begin{cases} a_{\mathcal{F}\mathrm{lim},n} + 2^{-n-1} & a_{\mathcal{F}\mathrm{lim},n} + 2^{-n-1} \leq a_{n+1} \\ a_{n+1} & a_{\mathcal{F}\mathrm{lim},n} - 2^{-n-1} \leq a_{n+1} \leq a_{\mathcal{F}\mathrm{lim},n} + 2^{-n-1} \\ a_{\mathcal{F}\mathrm{lim},n} - 2^{-n-1} & a_{\mathcal{F}\mathrm{lim},n} - 2^{-n-1} \geq a_{n+1} \end{cases},$$

and define the *forced limit* $\mathcal{F}\mathrm{lim}_{n\to\infty} a_n = \lim_{n\to\infty} a_{\mathcal{F}\mathrm{lim},n}$.

The authors of [BU10] make some observations about their construction:

Fact 3.3.7 ([BU10, Lemma 3.7])**.**

- *The function $\mathcal{F}\mathrm{lim} : [0,1]^\omega \to [0,1]$ is continuous*
- *If $(a_n : n < \omega)$ is a sequence such that $|a_n - a_{n+1}| \leq 2^{-n}$ for all n, then $\mathcal{F}\mathrm{lim} a_n = \lim a_n$*
- *If $a_n \to b$ fast enough that $|a_n - b| \leq 2^{-n}$ for all n, then $\mathcal{F}\mathrm{lim} a_n = b$.*

We wish to make one more observation (our technical reason for using this explicit construction):

Lemma 3.3.8. *If $(a_n : n < \omega)$ is such that $b - 2^{-n} \leq a_n$ for all n, then $b \leq \mathcal{F}\mathrm{lim} a_n$.*

Proof. We just need to show inductively that $b - 2^{-n} \leq a_{\mathcal{F}\mathrm{lim},n}$, as then the limit of this sequence must be at least b.

By definition, $b - 2^{-0} \leq a_0 = a_{\mathcal{F}\mathrm{lim},0}$.

Then assume $b - 2^{-n} \leq a_{\mathcal{F}\mathrm{lim},n}$. In the three cases of the definition of $a_{\mathcal{F}\mathrm{lim},n+1}$, either $a_{\mathcal{F}\mathrm{lim},n+1} \geq a_{n+1}$ or $a_{\mathcal{F}\mathrm{lim},n+1} = a_{\mathcal{F}\mathrm{lim},n} + 2^{-n-1}$. In the first case, we have $b - 2^{-n-1} \leq a_{n+1} \leq a_{\mathcal{F}\mathrm{lim},n+1}$, and in the second, we have $a_{\mathcal{F}\mathrm{lim},n} + 2^{-n-1} \geq b - 2^{-n} + 2^{-n-1} = b - 2^{-n-1}$. □

As $\mathcal{F}\mathrm{lim}$ is continuous, it can be used as an infinitary connective on definable predicates. That is, if $(\phi_n : n < \omega)$ is a sequence of definable predicates, $\mathcal{F}\mathrm{lim}\phi_n$, defined by pointwise forced limits, will be as well.

3.4 NIP and Honest Definitions

The following definition of NIP for metric structures comes from [Ben09]:

Definition 3.4.1 (IP and NIP)**.** We say a formula $\phi(x; y)$ is *independent* or has *IP* when there exists an indiscernible sequence $(a_i : i \in \omega)$, some tuple b, and some $0 \leq r < s \leq 1$ such that for all even i, $\vDash \phi(a_i; b) \leq r$ and for all odd i, $\vDash \phi(a_i; b) \geq s$.

We say that T is/has *NIP* when no formula $\phi(x; y)$ has IP.

This indiscernible definition is equivalent to a definition in terms of fuzzy VC-theory, by [Ben09, Lemma 5.4].

Fact 3.4.2. *The following are equivalent:*

- *The formula $\phi(x; y)$ is NIP*

- *For all models $\mathcal{M} \vDash T$, the function $\phi(x;y) : M^x \times M^y \to [0,1]$ is a VC-function.*

We can also give a geometric description of NIP formulas. If $\phi(x;y)$ is a formula, we can view the set of ϕ-types $S_\phi(B)$ over some parameter set B as a subset of $[0,1]^B$, defining it as

$$S_\phi(B) = \{(\phi(a;b) : b \in B) : a \in M^x\}.$$

Lemma 3.4.3. *A formula $\phi(x;y)$ has IP if and only if there exists an infinite parameter set $B \subseteq M$ in some $M \vDash T$ such that the closed convex hull of $S_\phi(B)$ has nonempty interior in the ℓ_∞ metric.*

Proof. Suppose that B is infinite and the convex hull of $S_\phi(B)$ has nonempty interior. Then for some $\varepsilon > 0$, there is some open ε-ball in the ℓ_∞ metric contained in the closed convex hull of $S_\phi(B)$, so the closed convex hull of the function class $(\phi(x;b) : b \in B)$ on M^x has infinite $\frac{\varepsilon}{2}$-fat-shattering dimension. The not-necessarily-closed convex hull will also have infinite δ-fat-shattering dimension for every $\delta < \frac{\varepsilon}{2}$. Thus by [Men02, Theorem 1.5], which places a bound on the δ-fat-shattering dimension of a convex hull in terms of the $\frac{\delta}{4}$-fat-shattering dimension of the larger class, we see that the $\frac{\varepsilon}{8}$-fat-shattering dimension of $(\phi(x;b) : b \in B)$ is infinite, so $\phi(x;y)$ has IP.

Suppose that $\phi(x;y)$ has IP. Then there is some $\varepsilon > 0$ such that the ε-fat-shattering dimension of $\phi(x;y)$ is infinite in some model $M \vDash T$. This means that the partial type on variables $(x_\sigma : \sigma \in \{0,1\}^{\mathbb{N}}; y_n : n \in \mathbb{N})$ consisting of $\phi(x_\sigma; y_n) + 2\varepsilon \leq \phi(x_\tau; y_n)$ for each $n \in \mathbb{N}$ and σ, τ that are equal except for the nth coordinate where $\sigma(n) = 0$ and $\tau(n) = 1$ is consistent, so we can find some $(a_\sigma : \sigma \in \{0,1\}^{\mathbb{N}}; b_n : n \in \mathbb{N})$ realizing this type. Then if $B = \{b_n : n \in \mathbb{N}\}$, the convex hull of $(\mathrm{tp}_\phi(a_\sigma; B) : \sigma \in \{0,1\}^{\mathbb{N}})$ will contain a ε-ball. $\square$

This section is dedicated to defining a continuous logic version of a third equivalent definition of NIP, honest definitions, and proving its equivalence to the others.

Definition 3.4.4. Let A be a closed subset of M^x where $M \preceq \mathcal{U}$ and $(M, A) \preceq (M', A')$. Let $\phi(x; b)$ be an M-predicate, and let $\psi(x; d)$ be an A'-predicate. We say that $\psi(x; d)$ is an *honest definition* for $\phi(x; b)$ over A when

- for all $a \in A$, $\phi(a; b) = \psi(a; d)$
- for all $a' \in A'$, $\phi(a'; b) \leq \psi(a'; d)$.

If the same predicate $\psi(y; z)$ works for any choice of M, A, b with $|A| \geq 2$, then we call $\psi(x; z)$ an honest definition for $\phi(x; y)$. Also, because we are only concerned with honest definitions with parameters in $A \subseteq M^x$, we assume that $z = (x_1, \dots, x_n)$ or $z = (x_1, x_2, \dots)$. In either case, we abuse notation slightly and use A^z to refer to A^n or $A^{\mathbb{N}}$ in those respective cases.

For all $\phi(x; y)$ and $\psi(y; z)$, we also define a predicate

$$\mathrm{HD}_{\phi,\psi,P,Q}(y; z) = \max\left(\sup_{x:P(x)} |\phi(x; y) - \psi(x; z)|, \sup_{x:Q(x)} \phi(x; y) \dot{-} \psi(x; z)\right).$$

Then for $d \in A'^z$, $(M', A, A') \vDash \mathrm{HD}_{\phi,\psi,P,Q}(b; d)$ if and only if $\psi(x; d)$ is an honest definition for $\phi(x; b)$. We will abuse notation later to write $\mathrm{HD}_{\phi,\psi,A,A'}(b; d)$ for the value of $\mathrm{HD}_{\phi,\psi,P,Q}(b; d)$ in (M', A, A').

In classical logic, all predicates are formulas, and only take values 0 and 1 corresponding to true and false. Then our definition of $\psi(x; d)$ being an honest definition for $\phi(x; b)$ over A corresponds to

$$\phi(A; b) \subseteq \psi(A'; d) \subseteq \phi(A'; b),$$

which is how honest definitions are presented in [Sim15, Theorem 3.13].

Because the property of $\psi(x; d)$ being an honest definition for $\phi(x; b)$ is encapsulated in $\mathrm{HD}_{\phi,\psi,A,A'}(b; d)$, we see that it does not depend on the choice of (M', A'), as long as $d \in A'^z$. On our way to honest definitions, it will sometimes be easier to work with $\psi(x; d)$ such that $\mathrm{HD}_{\phi,\psi,A,A'}(b;d)$ is small, but not necessarily zero. In fact, finding such $\psi(x;d)$ with $\mathrm{HD}_{\phi,\psi,A,A'}(b;d)$ arbitrarily small implies the existence of an honest definition.

Lemma 3.4.5. *Let A be a closed subset of M^x where $M \preceq \mathcal{U}$, and let $\phi(x;b)$ be an M-predicate. Let $(M,A) \preceq (M',A')$, and let $d \in A'^z$. Let $(\psi_n(x;z) : n \in \mathbb{N})$ be a sequence of definable predicates with $\mathrm{HD}_{\phi,\psi_n,A,A'}(b;d) \leq 2^{-n}$ for each n. Then $\mathcal{F}\lim \psi_n(x;d)$ is an honest definition for $\phi(x;b)$ over A.*

If instead we have a sequence $(\psi_n(x;z_n) : n \in \mathbb{N})$ with different $d_n \in A'^{z_n}$ such that $\mathrm{HD}_{\phi,\psi_n,A,A'}(b;d_n) \leq 2^{-n}$ for each n, then $\mathcal{F}\lim \psi_n(x;d)$ is an honest definition for $\phi(x;b)$ over A, where d is a concatenation of all the tuples d_n.

Proof. Let $\psi(x;z) = \mathcal{F}\lim \psi_n(x;z)$. If $a \in A$, we have $|\psi_n(a;d) - \phi(a;b)| \leq 2^{-n}$, so by Fact 3.3.7, $\mathcal{F}\lim \psi_n(a;d) = \phi(a;b)$. If $a' \in A'$, we have $\phi(a';b) - 2^{-n} \leq \psi_n(a';d)$, so by Lemma 3.3.8, $\phi(a';b) \leq \mathcal{F}\lim \psi_n(a';d)$. Thus $\psi(x;d)$ is an honest definition of $\phi(x;b)$ over A.

If each $\psi_n(x;z_n)$ uses on different parameters d_n, and we let z be a concatenation of all variable tuples z_n, with d a concatenation of all the parameters d_n, then for each n, we can think of ψ_n as a predicate $\psi_n(x;z)$ with $\psi_n(x;d) = \psi_n(x;d_n)$. Then we have $|\psi_n(a;d) - \phi(a;b)| \leq 2^{-n}$, so defining $\psi(x;z) = \mathcal{F}\lim \psi_n(x;z)$ we still get that $\psi(x;d)$ is an honest definition of $\phi(x;b)$ over A. $\square$

Theorem 3.4.6. *Assume T is NIP. Let $\mathcal{M} \models T$, $A \subseteq M^x$ closed, $\phi(x)$ a definable predicate with parameters in M. Then $\phi(x)$ admits an honest definition over A.*

Proof. Let (M',A') be a $|M|^+$-saturated elementary extension.

We use the set $S_A \subseteq S_x(\mathcal{U})$ of types approximately realizable in A, and the fact that S_A is compact. We will replace this with the set of *approximately realized* types, as in [Ben10b, Def. 3.1]. In Fact 3.3, it is established that the set of all such types is in fact closed, and thus compact. Let $p \in S_A$, and let $\phi(p)$ be the unique value of $\phi(a)$ for $a \vDash p$. We will first show that $p|_{A'}(x)$ and $\{P(x) = 0\}$ implies $\phi(x) = \phi(p)$.

Fix $\varepsilon > 0$. We will try to build a Morley sequence $(a_i : i \in \omega)$ for p over A in A' that contradicts NIP, by satisfying these properties:

- $P(a_i) = 0$
- $a_i \models p|_{Aa_{<i}}$
- $\models |\phi(a_{i+1}) - \phi(a_i)| \geq \frac{\varepsilon}{2}$

If we can build such a sequence, it will be indiscernible over A, and will thus violate NIP. Thus for some i, the partial type $p|_{Aa_{\leq i}} \cup \{P(x) = 0\} \cup \{|\phi(x) - \phi(a_i)| \geq \frac{\varepsilon}{2}\}$ is not consistent. We see that p must not contain the formula $|\phi(x) - \phi(a_i)| \geq \frac{\varepsilon}{2}$, or else this would be a subset of $p \cup \{P(x) = 0\}$, which is consistent as p is approximately realizable in A. Thus $|\phi(p) - \phi(a_i)| < \frac{\varepsilon}{2}$, and thus the partial type $p|_{Aa_{\leq i}} \cup \{P(x) = 0\} \cup \{|\phi(x) - \phi(p)| \geq \varepsilon\}$ is not consistent. As this means $p|_{A'} \cup \{P(x) = 0\}$ implies $|\phi(x) - \phi(p)| < \varepsilon$ for every $\varepsilon > 0$, we see that $p|_{A'} \cup \{P(x) = 0\}$ implies $\phi(x) = \phi(p)$. By Lemma 3.3.4, there is an A'-definable $\mathcal{L}_P$ predicate $\psi_P(x; d_1)$ in the pair language such that $S_A(x)$ and $P(x) = 0$ imply $\phi(x) = \psi_P(x; d_1)$. Thus by replacing each instance of the predicate $P(x)$ in $\psi_P(x; z)$ with 0 gives an A definable $\mathcal{L}$-predicate $\psi_0(x; z)$ with $P(x) = 0$ implying $\psi_0(x; z) = \psi_P(x; z)$.

This means that $S_A(x)$ and $P(x) = 0$ imply $\phi(x) = \psi_0(x; d_1)$, so if we let $\theta_0(x) = \phi(x) \dot{-} \psi_0(x; d_1)$, we see that $S_A(x)$ and $P(x) = 0$ imply $\theta_0(x) = 0$, so there is some A'-definable $\theta(x; d_2)$ with $S_A(x)$ implying $\theta(x; d_2) = 0$ and $P(x)$ implying $\theta(x; d_2) \geq \theta_0(x)$. Thus letting $\psi(x; d) = \psi_0(x; d_1) + \theta(x; d_2)$, we see that for $a \in A$, as $S_A(a)$ holds and $P(a) = 0$, $\psi(a; d) = \psi_0(a; d_1) + \theta(a; d_2) = \phi(a)$, and for $a \in A'$, as $P(a') = 0$, we have $\psi(a; d) \geq \psi_0(a; d_1) + (\phi(a) \dot{-} \psi_0(a; d_1)) \geq \phi(a)$. $\square$

In order to uniformize honest definitions, we will work with series of approximations to honest definitions over finite sets.

Lemma 3.4.7. *Let A be a closed subset of M^x where $M \preceq \mathcal{U}$, and let $\phi(x; b)$ be an M-predicate. Fix $(M, A) \preceq (M', A')$ to be $|M|^+$-saturated, $\varepsilon > 0$, and a definable predicate*

$\psi(x; z)$.

If there exists $d \in A'^z$ such that $\mathrm{HD}_{\phi,\psi,A,A'}(b; d) < \varepsilon$, then for all finite $A_0 \subseteq A$, we have there is a tuple $d_{A_0} \in A^z$ such that $\mathrm{HD}_{\phi,\psi,A_0,A}(b; d_{A_0}) < \varepsilon$.

Conversely, if for all finite $A_0 \subseteq A$, there is a tuple $d_{A_0} \in A^z$ with $\mathrm{HD}_{\phi,\psi,A_0,A}(b; d_{A_0}) \leq \varepsilon$, then there exists $d \in A'^z$ such that $\mathrm{HD}_{\phi,\psi,A,A'}(b; d) \leq \varepsilon$.

Proof. First, we observe that for finite A_0, $\mathrm{HD}_{\phi,\psi,A_0,A}(b; z)$ is equivalent to the predicate

$$\max\left(\max_{a_0 \in A_0} |\phi(a_0; y) - \psi(a_0; z)|, \sup_{x:P(x)} \phi(x; y) \dot{-} \psi(x; z)\right)$$

which is expressible using only the predicate P. Thus by elementarity, for any $d \in A^z$, $\mathrm{HD}_{\phi,\psi,A_0,A}(b; d) = \mathrm{HD}_{\phi,\psi,A_0,A'}(b; d)$, and similarly,

$$\inf_{z \in A^z} \mathrm{HD}_{\phi,\psi,A_0,A}(b; d) = \inf_{z \in A'^z} \mathrm{HD}_{\phi,\psi,A_0,A'}(b; z).$$

Now fix $\varepsilon > 0$. First, assume that $d \in A'^z$ is such that $\mathrm{HD}_{\phi,\psi,A,A'}(b; d) < \varepsilon$. Then because inf corresponds to $\exists$ for open conditions, $\inf_{z \in A'^z} \mathrm{HD}_{\phi,\psi,A,A'}(b; z) < \varepsilon$, and for every finite $A_0 \subseteq A$, $\inf_{z \in A'^z} \mathrm{HD}_{\phi,\psi,A_0,A'}(b; z) < \varepsilon$. Then by elementarity, $\inf_{z \in A^z} \mathrm{HD}_{\phi,\psi,A_0,A}(b; z) < \varepsilon$, so there is some $d_{A_0} \in A^z$ such that $\mathrm{HD}_{\phi,\psi,A_0,A}(b; d_{A_0}) < \varepsilon$.

Now, assume that for all finite $A_0 \subseteq A$, there is a tuple $d_{A_0} \in A^z$ with $\mathrm{HD}_{\phi,\psi,A_0,A}(b; d_{A_0}) \leq \varepsilon$.

Let (M', A') be a $|M|^+$-saturated elementary extension of (M, A). We claim that for some fixed $d \in A'^z$, $\mathrm{HD}_{\phi,\psi,A,A'}(b; d) \leq \varepsilon$ if and only if $\mathrm{HD}_{\phi,\psi,A_0,A'}(b; d) \leq \varepsilon$ for each $A_0 \subseteq A$. This is because both inequalities are equivalent to stating that $\phi(a'; b) \leq \psi(a'; d) + \varepsilon$ for all $a' \in A'$, as well as stating that $|\phi(a; b) - \psi(a; d)| \leq \varepsilon$ for all $a \in A$. Thus to find $d \in A'^z$ with $\mathrm{HD}_{\phi,\psi,A,A'}(b; d) \leq \varepsilon$, it suffices to show that the partial type

$$p(z) = \{\mathrm{HD}_{\phi,\psi,A_0,A'}(b; z) \leq \varepsilon : A_0 \subseteq A\}$$

is consistent, which it is, as any finite subtype is implied by the condition $\mathrm{HD}_{\phi,\psi,A_0,A'}(b;z) \leq \varepsilon$ with A_0 finite, which is equivalent to $\mathrm{HD}_{\phi,\psi,A_0,A}(b;z) \leq \varepsilon$, which is realized by some $d_{A_0} \in A^z$. □

Now we work towards uniformizing Honest Definitions, using the characterization over finite sets, so that we can use the same formula ψ for all sets A.

Lemma 3.4.8. *Assume T is NIP.*

Let $\phi(x;y)$ be a formula, $\varepsilon > 0$, and assign to each predicate $\psi(x;z)$ a number $q_\psi \in \mathbb{N}$. Then there are finitely many formulas $\psi_0, \ldots, \psi_{n-1}$ such that:

For any $\mathcal{M} \models T$, $A \subseteq M$ closed, $b \in M^y$, there exists $j < n$ such that for any $A_0 \subseteq A$ of size $|A_0| \leq q_{\psi_j}$, $\inf_{z:P(z)} \mathrm{HD}_{\phi,\psi,A_0,A}(b;z) < \varepsilon$.

Proof. We work in the extended language $\mathcal{L} \cup \{P(x), c_b\}$, where c_b is a tuple of constants of the same cardinality as z.

For each $\psi(x;z)$, let

$$\Theta_\psi = \sup_{x_0,\ldots,x_{q_\psi - 1} \in P} \left(\inf_{z:P(z)} \left(\max \left(\max_{i<q_\psi} |\phi(x_i;c_b) - \psi(x_i;z)|, \sup_{x:P(z)} \phi(x;c_b) \dot{-} \psi(x;z) \right) \right) \right).$$

This formula is defined so that in an expansion (M, A, b) of a model $M \vDash T$,

$$\Theta_\psi = \sup_{A_0 \subseteq A : |A_0| \leq q_\psi} \inf_{z:P(z)} \mathrm{HD}_{\phi,\psi,A_0,A}(b;z).$$

For each model (M, A, b) of this extended language, by Theorem 3.4.6, there is an honest definition $\psi(x;d)$ of $\phi(x;b)$ over A, so $\mathrm{HD}_{\phi,\psi,A,A'}(b;d) = 0 \leq \frac{\varepsilon}{2}$. Thus by Lemma 3.4.7, for all finite $A_0 \subseteq A$, $\inf_{z \in A^z} \mathrm{HD}_{\phi,\psi,A_0,A}(b;z) < \frac{\varepsilon}{2}$. Thus the supremum over all such A_0 is at most $\frac{\varepsilon}{2}$, and $(M, A, b) \vDash \Theta_\psi \leq \frac{\varepsilon}{2}$.

As at least one of the open conditions $\{\Theta_\psi < \varepsilon : \psi(x;z)\}$ holds in every model, this set covers the (zero-variable) type space. Thus by compactness, there is a finite collection $\psi_0, \ldots, \psi_{n-1}$ such that one of the open conditions $\Theta_{\psi_j} < \varepsilon$ is true in each model (M, A, b). Unpacking the definition of Θ_ψ, this yields the result. $\square$

We can now apply Corollary 3.2.32 to finish uniformizing Honest Definitions:

Theorem 3.4.9. *Assume T is NIP. Every definable predicate $\phi(x; y)$ admits an honest definition $\psi(x; z)$.*

Proof. For each $\varepsilon > 0$, we will find $\psi(x; z)$ such that for every A, b, and any finite $A_0 \subseteq A$, there is some $d \in A^z$ such that $\mathrm{HD}_{\phi,\psi,A_0,A}(b; d) < \varepsilon$. Then by Lemma 3.4.7, for every A, b and saturated extension $(M, A) \preceq (M', A')$, there will be $d \in A'^z$ with $\mathrm{HD}_{\phi,\psi,A_0,A}(b; d) \leq \varepsilon$. If for each n, we choose $\psi_n(x; z_n)$ that works for $\varepsilon = 2^{-n}$, then by Lemma 3.4.5, the forced limit $\mathcal{F}\lim \psi_n(x; z)$ will be a uniform honest definition.

Fix $\varepsilon > 0$. If there are finitely many predicates $\psi_0, \ldots, \psi_{n-1}$ such that for each expanded structure (M, A, b, A_0), one suffices, we can use the standard coding tricks (see Lemma 3.3.2) to find a single ψ that can code all of these, provided $|A| \geq 2$.

Having made all these reductions, we now find candidate predicates using Lemma 3.4.8. Given a partitioned predicate $\psi(x; z)$, let $q_\psi = \mathrm{vc}^*_{\frac{\varepsilon}{2}, \frac{3\varepsilon}{4}}(|\phi(x, y) - \psi(x; z)|) + 1$, where we view $|\phi(x, y) - \psi(x; z)|$ as partitioned between variables (x, y) and z. Now let $\psi'_0, \ldots, \psi'_{n-1}$ be the predicates given by Lemma 3.4.8 such that for any $\mathcal{M} \models T$, $A \subseteq M$ closed, $b \in M^y$, there exists $j < n$ such that for any $A_0 \subseteq A$ of size $\leq q_{\psi'_j}$, $\inf_{z:P(z)} \mathrm{HD}_{\phi,\psi'_j,A_0,A}(b; z) < \frac{\varepsilon}{2}$.

Now fix M, A, b, A_0. We know that for some $j < n$, and for all $A'_0 \subseteq A_0$ of size $\leq q_{\psi'_j}$, there is some $d \in A^z$ such that $\mathrm{HD}_{\phi,\psi'_j,A'_0,A}(b; d) < \frac{\varepsilon}{2}$.

Let $D = \{d \in A^z : \forall a \in A, \phi(a; b) < \psi'_j(a; d) + \frac{\varepsilon}{2}\}$. Let Q be the finite function class on D consisting of the functions $\{|\phi(a_0; b) - \psi'_j(a_0; z)| : a_0 \in A_0\}$. Then for any r, s, we have $\mathrm{vc}^*(Q_{r,s}) \leq \mathrm{vc}^*_{r,s}(|\phi(x, y) - \psi'_j(x; z)|)$, as Q consists of fewer functions on a restricted domain. In particular, $q_{\psi'_j} \geq \mathrm{vc}^*(Q_{\frac{\varepsilon}{2}, \frac{3\varepsilon}{4}}) + 1$, and $Q_{\leq \frac{\varepsilon}{2}}$ has the $(q_{\psi'_j}, q_{\psi'_j})$ property, so by Corollary 3.2.32, there is some N depending only on $\mathrm{vc}^*_{\frac{\varepsilon}{2},\frac{3\varepsilon}{4}}(|\phi(x,y)-\psi'_j(x;z)|)$ and $\mathrm{vc}^*_{\frac{3\varepsilon}{4},\varepsilon}(|\phi(x,y)-\psi'_j(x;z)|)$ such that $\tau(Q_{<\varepsilon}) \leq N$. That is, there exist $d_1,\dots,d_N \in D$ such that for each $a \in A_0$, there is some d_i with $|\phi(a;b)-\psi'_j(a;d_i)| < \varepsilon$.

Now we let $\psi_j(x;z_1,\dots,z_N) = \min_{1\leq i\leq N}\psi'_j(x;z_i)$, remembering that N depends only on ϕ,ψ'_j. It suffices to show that $\mathrm{HD}_{\phi,\psi_j,A_0,A}(b;d_1,\dots,d_N) < \varepsilon$. We see that $\psi_j(x;d_1,\dots,d_N)$ satisfies for all $a \in A$, $\phi(a;b) < \psi_j(a;d_1,\dots,d_N)+\varepsilon$, as for each i, $\phi(a;b) < \psi'_j(a;d_i)+\frac{\varepsilon}{2}$, so we have taken a minimum of functions that are all sufficiently large. Also, for each $a \in A_0$, there exists some d_i with $\psi'_j(a;d_i) < \phi(a;b)+\varepsilon$, so taking the minimum $\psi_j(a;d_1,\dots,d_N) < \phi(a;b)+\varepsilon$, and $|\phi(x;b)-\psi_j(a;d_1,\dots,d_N)| < \varepsilon$. □

We now get a version of uniform definability of types over finite sets (UDTFS).

Definition 3.4.10. Let $\phi(x;y)$ be a definable predicate. Then we say $\phi(x;y)$ has UDTFS when there is a definable predicate $\psi(x;z)$ (where z consists of k copies of x, where k is possibly infinite) such that for any finite $A \subseteq \mathcal{U}^x$ with $|A| \geq 2$, and any $b \in \mathcal{U}^y$, there is d in A^k such that $\phi(a;b) = \psi(a;d)$ for all $a \in A$.

Corollary 3.4.11 (UDTFS)**.** *Assuming T is NIP, every definable predicate $\phi(x;y)$ has UDTFS.*

Proof. Simply let $\psi(x;y)$ be an honest definition of $\phi(x;z)$. □

UDTFS also provides polynomial bounds on covering numbers.

Lemma 3.4.12. *Let $\phi(x;y)$ be a formula such that $\phi(x;y)$ has UDTFS, with uniform definition $\psi(x;z)$. Let $\varepsilon > 0$, and let $\psi_\varepsilon(x;z)$ be a formula depending only on a finite number k of the variables of z such that $\vDash \sup_x \sup_z |\psi(x;z)-\psi_\varepsilon(x;z)| \leq \varepsilon$. Then $\mathcal{N}_{\phi(x;y),\varepsilon}(n) = O(n^k)$. (In fact, $\mathcal{N}_{\phi(x;y),\varepsilon}(n) \leq n^k$ for $n \geq 2$.)*

Proof. Recall that $\mathcal{N}_{\phi(x;y),\varepsilon}(n)$ is the supremum of the ε-covering numbers in the ℓ_∞-metric of the sets $\phi(\bar{a};y) = \{(\phi(a_i;b) : 1 \leq i \leq n) : b \in \mathcal{U}^y\}$ for $\bar{a} = (a_1,\dots,a_n) \in (\mathcal{U}^x)^n$.

Fix $\bar{a}$, and let $A = \{a_1, \ldots, a_n\}$. If $|A| = 0$, then $n = 0$, and this is trivial. If $|A| = 1$, then there is a ε-cover of size at most $\mathcal{N}_{\phi(x;y),\varepsilon}(1)$, a constant.

Now assume $|A| \geq 2$. By UDTFS, the set $\phi(\bar{a}; y)$ equals the set $\psi(\bar{a}; z)$. Let $z_0 \subseteq z$ be the finite tuple with $|z_0| = k$ on which ψ_ε depends. Let $\pi : A^z \to A^{z_0}$ be the restriction map, and let $D \subseteq A^z$ be such that π is bijective on D. Thus $|D| = |A^{z_0}| \leq n^k$. Then $\{(\psi_\varepsilon(a_i; d) : 1 \leq i \leq n) : d \in D\}$ is a ε-cover for $\psi(\bar{a}; z)$, as for every $d \in \mathcal{U}^y$, there is some $d' \in D$ with $\pi(d) = \pi(d')$, and thus $\psi_\varepsilon(d) = \psi_\varepsilon(d')$, so in turn, for all $a \in A$, $|\psi(a; d) - \psi_\varepsilon(a; d')| \leq \varepsilon$. Thus $(\psi_\varepsilon(a_i; d') : 1 \leq i \leq n)$ is within ε of $(\psi(a_i; d) : 1 \leq i \leq n)$ in the ℓ_∞-metric. $\square$

We now tie UDTFS back into a characterization of NIP.

Lemma 3.4.13. *Let $\phi(x; y)$ be a definable predicate, and assume that $\phi(x; y)$ has UDTFS. Then $\phi(x; y)$ is NIP.*

Proof. By Lemma 3.2.19, the polynomial bound given by Lemma 3.4.12 on the covering number shows that $\phi(x; y)$ is a VC-class of functions. $\square$

This gives us several equivalent characterizations of NIP:

Theorem 3.4.14. *The following are equivalent:*

- *T is NIP*
- *every definable predicate $\phi(x; y)$ admits an honest definition $\psi(x; z)$*
- *T has UDTFS.*

It remains to be checked whether a given formula or predicate $\phi(x; y)$ being NIP guarantees uniformity of honest definitions and UDTFS, although this was recently established for discrete logic in [EK21].

3.4.1 The Shelah Expansion

We now propose definitions of externally definable predicates and the Shelah expansion in continuous logic. We confirm that it preserves NIP, as in classical logic, using a generalization of the honest definitions proof from [CS13].

Definition 3.4.15 (External definability)**.** Let M be a metric $\mathcal{L}$-structure. We say that a function $f : M^x \to [0,1]$ is an *externally definable predicate* when there is some elementary extension $M \preceq N$, some definable predicate $\phi(x; y)$, and some $b \in N^y$ such that $f(a) = \phi(a; b)$ for all $a \in M^x$.

If $\phi(x; y)$ can be chosen to be a formula rather than just a definable predicate, we say that f is *externally formula-definable.*

Definition 3.4.16 (The Shelah Expansion)**.** Let M be a metric $\mathcal{L}$-structure, with $M \preceq N$ a $|M|^+$-saturated elementary extension. We define M^{Sh}, the *Shelah expansion* of M, to be the metric structure consisting of the same underlying metric space (M, d), together with a predicate symbol $P_{\phi,b}(x)$ for each $\mathcal{L}$-formula $\phi(x; y)$ and $b \in N^y$, interpreted so that $P_{\phi,b}(a) = \phi(a; b)$ for all $a \in M$. The formula $P_{\phi,b}$ is assigned a Lipschitz constant C such that $\phi(x; y)$ is provably C-Lipschitz. Denote this language $\mathcal{L}^{\mathrm{Sh}}$.

Lemma 3.4.17. *Fix $M \preceq N$ with N $|M|^+$-saturated. Then the predicates $M^x \to [0,1]$ given by quantifier-free formulas $\phi(x)$ in $\mathcal{L}^{\mathrm{Sh}}$ are exactly the externally formula-definable predicates on M, and the quantifier-free $\mathcal{L}^{\mathrm{Sh}}$-definable predicates on M^{Sh} are precisely the externally definable predicates on M.*

Proof. By definition, any externally (formula-)definable predicate $f : M^x \to [0,1]$ is given by $\phi(x; b)$ for some formula/definable predicate $\phi(x; y)$ and some $b \in N'^y$ where $M \preceq N'$. For any b' in any extension of M, $\phi(x; b')$ defines f if and only if b realizes the partial type $p(y) = \{\phi(a; y) = f(a) : a \in M^x\}$. This partial type is realized by b, so by saturation, it is realized by some $b' \in N$, so f is externally (formula-)definable with parameters in N. Thus the choice of N does not matter, and it suffices to consider parameters in a fixed N.

Thus if f is externally formula-definable, we may choose a formula $\phi(x;y)$ and $b \in N^y$ such that $f(x) = \phi(x;b) = P_{\phi,b}(x)$ on M, so f is given by a formula in $\mathcal{L}^{\mathrm{Sh}}$.

Conversely, it is clear that for any formula $\phi(x;y)$ and any $b \in N^y$, the basic $\mathcal{L}^{\mathrm{Sh}}$-formula $P_{\phi,b}(x)$ is externally formula-definable by $\phi(x;b)$. Any continuous connectives (not quantifiers) we apply to these predicate symbols will preserve external formula-definability, if we apply them to the defining formulas, so by induction, all quantifier-free $\mathcal{L}^{\mathrm{Sh}}$-formulas are $\mathcal{L}$-externally formula-definable.

The externally definable predicates are exactly the uniform limits of externally formula-definable predicates, as the uniform limit of $(\phi_n(x;b_n) : n < \omega)$ can be externally defined with $\lim_n \phi_n(x;b_0,b_1,\dots)$, with $b_0b_1\dots$ a tuple over N. Thus they are exactly the uniform limits of quantifier-free $\mathcal{L}^{\mathrm{Sh}}$-formulas, which are the quantifier-free $\mathcal{L}^{\mathrm{Sh}}$-definable predicates. $\square$

For the remainder of this section, we assume T is NIP, and fix $M \preceq N$ $|M|^+$-saturated.

Lemma 3.4.18. *Let $f : M^x \to [0,1]$ be externally definable. Then there is a definable predicate $\phi(x;b)$ with $b \in \mathcal{U}^y$ such that $\phi(a;b) = f(a)$ for all $a \in M^x$ and for every M-definable predicate $\theta(x;c)$ with $\theta(a;c) \leq f(a)$ for all $a \in M^x$, we also have $\mathcal{U} \vDash \theta(x;c) \leq \phi(x;b)$.*

Proof. Let $\psi(x;d)$ be an external definition of f, with $M \preceq N$ and $d \in N^z$. Then consider the pair (N,M), and apply Theorem 3.4.9. There is some elementary extension $(N,M) \preceq (N',M')$ and an honest definition $\phi(x;b)$ of $\psi(x;d)$ over M with $b \in M'^y$. This means that for $a \in M^x$, $\phi(a;b) = \psi(a;d) = f(a)$, and $(N',M') \vDash \sup_{x\in P} \psi(a;d) \dot{-} \phi(a;b) = 0$. Now let $\theta(x;c)$ with $c \in M^w$ be such that $\theta(a;c) \leq f(a) = \psi(a;d)$ for all $a \in M^x$ Then $(N,M) \vDash \sup_{x\in P} \theta(x;c) \dot{-} \psi(a;d) = 0$, so the same condition holds in (N',M'), and thus $(N',M') \vDash \sup_{x\in P} \theta(x;c) \dot{-} \phi(a;b) = 0$, so $M' \vDash \sup_x \theta(x;c) \dot{-} \phi(a;b) = 0$, and by elementarity, $\mathcal{U} \vDash \sup_x \theta(x;c) \dot{-} \phi(a;b) = 0$. $\square$

We now generalize Shelah's expansion theorem to continuous logic, using honest definitions as in the proof in the discrete case given in [CS13].

Theorem 3.4.19. *The structure M^{Sh} admits quantifier elimination.*

Proof. By [BBH08, Lemma 13.5], it suffices to show that if $\phi(x;y)$ is a quantifier-free $\mathcal{L}^{\mathrm{Sh}}$-formula, then $\inf_x \phi(x;y)$ is approximable by quantifier-free formulas, that is, is a quantifier-free $\mathcal{L}^{\mathrm{Sh}}$-definable predicate. By Lemma 3.4.17, that means that it is enough to show that if $f(x,y) : M^{xy} \to [0,1]$ is externally formula-definable, then $\inf_{x\in M} f(x,y)$ is also externally definable.

Let $f(x,y)$ be externally formula-definable. In particular, there is some constant C such that $f(x,y)$ is C-Lipschitz, and $f(x,y)$ is externally definable. By Lemma 3.4.18, we may assume that $f(x;y)$ is given by a $\mathcal{L}$-predicate $\phi(x,y;d)$ with $d \in \mathcal{U}^z$, such that for every $\mathcal{L}(M)$-definable predicate $\theta(x,y;c)$ with $\theta(a,b;c) \leq f(a,b)$ for all $a,b \in M^{xy}$, we also have $\mathcal{U} \vDash \theta(x,y;c) \leq \phi(x,y;d)$. We claim that $\inf_{x\in M} f(x;y)$ is externally definable by $\inf_x \phi(x,y;d)$. Clearly for any $b \in M^y$,

$$\inf_x \phi(x,b;d) = \inf_{x\in\mathcal{U}} \phi(x,b;d) \leq \inf_{x\in M} \phi(x,b;d) = \inf_{x\in M} f(x,b),$$

so it suffices to show that for $b \in M^y$, $\inf_{x\in M} f(x,b) \leq \inf_{x\in\mathcal{U}} \phi(x,b;d)$.

Let $\zeta(x,y)$ by the $\mathcal{L}(M)$-formula $\inf_{x\in M} f(x;b) - Cd(y,b)$, noting that $\inf_{x\in M} f(x;b)$ is just a constant. Then for all $(a',b') \in M^{xy}$, we find that by the Lipschitz property of f,

$$f(a',b') \geq f(a',b) - Cd(b',b) \geq \inf_{x\in M} f(x;b) - Cd(b',b) = \zeta(a',b').$$

Thus by assumption on ϕ, $\mathcal{U} \vDash \zeta(x,y) \leq \phi(x,y;d)$, so $\mathcal{U} \vDash \inf_x \zeta(x,b) \leq \inf_x \phi(x,b;d)$. However, ζ has no dependence on x, so

$$\inf_{x\in\mathcal{U}} \zeta(x,b) = \inf_{x\in M} f(x;b) - Cd(b,b) = \inf_{x\in M} f(x;b),$$

and thus $\inf_{x \in M} f(x;b) \leq \inf_{x \in \mathcal{U}} \phi(x,b;d)$. □

Corollary 3.4.20. *The predicates $M^x \to [0,1]$ given by formulas $\phi(x)$ in $\mathcal{L}^{\mathrm{Sh}}$ are exactly the externally formula-definable predicates on M, and the $\mathcal{L}^{\mathrm{Sh}}$-definable predicates on M^{Sh} are precisely the externally definable predicates on M.*

Proof. By 3.4.19, we can drop the "quantifier-free" descriptions from Lemma 3.4.17. □

Corollary 3.4.21. *The structure M^{Sh} is NIP.*

Proof. Any definable predicate over M^{Sh} corresponds to an externally definable predicate $\phi(x;b)$ over M, which is dependent. □

3.5 Definitions of Distality

Let T be a theory in continuous logic. We will present several possible definitions of distality, and determine which of them are equivalent.

The first definition, in terms of indiscernible sequences, is unchanged from discrete logic.

Definition 3.5.1 (Distality)**.** Let I be an indiscernible sequence. Then we say that I is *distal* when for any indiscernible sequence $I_1 + I_2$ with the same EM-type as I, where I_1 and I_2 are dense and without endpoints, if $I_1 + d + I_2$ is also indiscernible and $I_1 + I_2$ is indiscernible over a set B, then $I_1 + d + I_2$ is also indiscernible over B.

We say T is *distal* when every indiscernible sequence in a model of T is distal.

This definition also appears in a limited continuous context in [KP22]. Note that we could equivalently add parameters to this definition. If $I + d + J$ is indiscernible over A with $I + J$ indiscernible over AB, then if $I + d + J$ is not indiscernible over AB, there must be finite tuples $a \subseteq A$, $b \subseteq B$ such that $I + d + J$ is not indiscernible over ab. If we let $I_a = (ia : i \in I)$ and $J_a = (ja : j \in J)$, then $I_a + da + J_a$ will be indiscernible over $\emptyset$ but not over b, and $I_a + J_a$ will be indiscernible over b, contradicting distality.

First we check that this definition of distality implies NIP.

Theorem 3.5.2. *If a metric theory T is distal, then T is NIP.*

Proof. Assume T is not NIP. Let $(a_i : i \in \omega)$ be an indiscernible sequence, b a tuple, $\phi(x; y)$ a formula, and $0 \leq r < s \leq 1$ such that $\models \phi(a_i; b) \leq r$ when i is even and $\models \phi(a_i; b) \geq s$ when i is odd.

We claim that there are sequences I, J of order type $\mathbb{Q}$ and some d such that $I + d + J$ is indiscernible, $I + J$ is indiscernible over b, but for all $i \in I + J$, $\models \phi(i; b) \leq r$ while $\models \phi(d; b) \geq s$. If so, this will contradict distality. Such an $I + d + J$ is exactly a realization of the following partial type Σ in variables

$$X = X_I \cup \{x_d\} \cup X_J = \{x_{iq} : q \in \mathbb{Q}\} \cup \{x_d\} \cup \{x_{jq} : q \in \mathbb{Q}\},$$

where $X^n_<$ is the set of increasing n-tuples of X, and $(X_I \cup X_J)^n_<$ is defined similarly:

$$\begin{aligned} &T \cup \{|\psi(\bar{x}) - \psi(\bar{x}')| = 0 : \psi \in \mathcal{L}; \bar{x}, \bar{x}' \in X^n_<\} \\ &\quad \cup \{|\psi(\bar{x}, b) - \psi(\bar{x}', b)| \leq \frac{1}{m} : \psi \in \mathcal{L}; \bar{x}, \bar{x}' \in (X_I \cup X_J)^n_<; m \in \mathbb{N}\} \\ &\quad \cup \{\phi(x, b) \leq r : x \in X_I \cup X_J\} \\ &\quad \cup \{\phi(x_d, b) \geq s\}. \end{aligned}$$

It suffices to show that Σ is consistent. Let $\Sigma_0 \subset \Sigma$ be finite, and let $\bar{x} \in (X_I)^n_<$, $\bar{x}' \in (X_J)^n_<$, and x_d include all the variables of X appearing in Σ_0. Then we will find a finite subsequence of $(a_i : i \in \omega)$ realizing Σ_0. It will automatically be $\emptyset$-indiscernible, and we will interpret $\bar{x}, \bar{x}'$ with even elements of the sequence, and x_d with an odd element, so we need only make sure that a finite set of conditions of the form $|\psi(x_1, \dots, x_n, b) - \psi(x'_1, \dots, x'_n, b)| \leq \frac{1}{m}$ are satisfied.

To do this, we find an infinite subsequence of $(a_{2i} : i \in \omega)$ such that for all ψ in a finite set $\Psi_0 = \{\psi_0, \dots, \psi_r\}$, some fixed m,a and each pair of increasing n-tuples $\bar{a}, \bar{a}'$, we have $|\psi(\bar{a}, b) - \psi(\bar{a}', b)| \leq \frac{1}{m}$. Assume for induction that S is an infinite subsequence such that this holds for all ψ_i with $i < k$. (For $k = 0$, we set $S = (a_{2i} : i \in \omega)$.) Then we color all finite subsequences $x_1 < \cdots < x_n$ of S with m colors, assigning a tuple color c_j when $\frac{j}{m} \leq \psi_k(x_1, \dots, x_n, b) < \frac{j+1}{m}$. By Ramsey's Theorem, there must be an infinite monochromatic subsequence, which satisfies the induction step.

Once we have this infinite subsequence S, we can select $\bar{a}$ to be an arbitrary increasing subsequence of S. Then we interpret x_d with some a_{2i+1} greater than all of $\bar{a}$, and let $\bar{a}'$ be in S and greater than a_{2i+1}. □

We will now show some useful lemmas for showing that indiscernible sequences are distal.

Lemma 3.5.3 (Generalizes [Sim13, Lemma 2.7])**.** *Assume T is NIP. If I is a dense indiscernible sequence without endpoints, then I is distal if and only if for every partition $I = I_1 + I_2 + I_3$ where I_1, I_2, I_3 have no endpoints, then for all b_1, b_2 such that $I_1 + b_1 + I_2 + I_3$ and $I_1 + I_2 + b_2 + I_3$ are indiscernible, then $I_1 + b_1 + I_2 + b_2 + I_3$ is also.*

Proof. Clearly distality implies this condition, so it suffices to check that such a sequence is distal.

First we observe that this alternative characterization of distality (at least for dense sequences) only depends on the EM-type of I. There exist b_1, b_2 such that $I_1 + b_1 + I_2 + I_3$ and $I_1 + I_2 + b_2 + I_3$ are indiscernible, but $I_1 + b_1 + I_2 + b_2 + I_3$ is not, if and only if there exists some formula $\phi(y_1, x_1, y_2, x_2, y_3)$, an $\varepsilon > 0$, such that when $(y_1, x_1, y_2, x_2, y_3)$ is an increasing tuple of variables, $\phi(y_1, x_1, y_2, x_2, y_3) = 0$ is in the EM-type of I, but $\phi(y_1, x_1, y_2, x_2, y_3) = \varepsilon$ is consistent with (y_1, x_1, y_2, y_3) and (y_1, y_2, x_2, y_3) satisfying the EM-type of I.

Now we will show another property that follows from this condition: for all natural numbers n, if $I = I_0 + I_1 + \dots I_n$ is a partition into dense endpointless pieces, and $b_0, \dots, b_{n-1}$ are such that for each i, $I_0 + \cdots + I_i + b_i + I_{i+1} + \cdots + I_n$ is indiscernible, then $I_0 + b_0 + I_1 +$

$b_1 + \cdots + b_{n-1} + I_n$ is also. We proceed by induction on n, with cases $n = 0, 1$ trivial, and case $n = 2$ assumed. Assuming this works for n for all such sequences, partition or sequence as $I_0 + I_1 + \cdots + I_{n+1}$, and find suitable $b_0, \ldots, b_n$. Then as $I' = I_0 + b_0 + I_1 + I_2 + \cdots + I_n$ is indiscernible, it has the same EM-type as I, so it also has this property. Thus the sequence obtained by inserting b_i into I' is indiscernible for all $i > 0$, so by our induction hypothesis, inserting all n extra elements gives an indiscernible sequence, as desired.

If our sequence I is not distal, then there exists a set B, a tuple d, and sequences $I_1 + I_2$ indiscernible over B, with the same EM-type as I, where I_1 and I_2 are dense and without endpoints, and $I_1 + d + I_2$ is indiscernible but not indiscernible over B.

Thus there is some formula $\phi(x_1, x, x_2)$ with parameters in B, and finite tuples $i_1 \subseteq I_1$ and $i_2 \subseteq I_2$ such that for any $i \in I_1 + I_2$ between i_1 and i_2, $\phi(i_1, i, i_2) = 0$, but $\phi(i_1, d, i_2) = \varepsilon > 0$. By avoiding i_1 and i_2, we can find a final segment $I_1' \subseteq I_1$ and an initial segment $I_2' \subseteq I_2$ such that $I_1' + I_2'$ is indiscernible over Bi_1i_2. By Bi_1i_2-indiscernibility, we see that for any partition of $I_1' + I_2'$ into endpointless pieces, there is some element d' that could be inserted, maintaining indiscernibility, but with $\phi(i_1, d', i_2) = \varepsilon$.

Now partition $I_1' + I_2'$ into a countable infinite sequence $J_0 + J_1 + J_2 + \ldots$ of endpointless parts. For each $n \in \mathbb{N}$, there is d_n such that inserting d_n between J_n and J_{n+1} maintains indiscernibility, but $\phi(i_1, d_n, i_2) = \varepsilon$. Inserting all of these either violates indiscernibility or NIP, as $\phi(i_1, d_n, i_2)$ alternates infinitely often between 0 and ε. We have shown that for each n, inserting all of $d_0, \ldots, d_n$ maintains indiscernibility, so inserting each d_n at once maintains indiscernibility. Thus NIP fails, contradicting our hypothesis. $\square$

This lemma is the metric version of a special case of [Sim13, Lemma 2.8], on strong base change. It is particularly useful in conjunction with Lemma 3.5.3.

Lemma 3.5.4. *Let $I = I_0 + I_1 + I_2$ be an indiscernible sequence, with $A \supset I$ a set of parameters, such that I_0, I_1, I_2 are dense without endpoints. Let a and b be such that $I_0 + a + I_1 + I_2$ and $I_0 + I_1 + b + I_2$ are indiscernible. Then there are a' and b' such that*

$\mathrm{tp}(a'b'/I) = \mathrm{tp}(ab/I)$, $\mathrm{tp}(a'/A) = \lim(I_0/A)$ *and* $\mathrm{tp}(b'/A) = \lim(I_1/A)$.

Proof. Assume that the conclusion is false. Then by compactness, there are closed conditions $\phi(x,y) = 0 \in \mathrm{tp}(ab/I)$, $\psi_0(x) = 0 \in \lim(I_0/A)$ and $\psi_1(y) = 0 \in \lim(I_1/A)$ such that $\{\phi(x,y) = 0, \psi_0(x) = 0, \psi_1(y) = 0\}$ is inconsistent. There is some minimum value ε taken by $\max(\psi_0(x), \psi_1(y))$ on the set of all types in $S_{xy}(A)$ satisfying $\phi(x,y) = 0$, and we see that $\varepsilon > 0$. Let $I_\phi \subset I$ be a finite tuple containing all parameters of ϕ.

Because $\psi_0(x) = 0 \in \lim(I_0/A)$, we can find a final segment $J_{0-} \subseteq I_0$ such that $\psi_0(x) \leq \frac{\varepsilon}{2}$ on all of J_{0-}, and an initial segment J_{0+} of I_1 such that $J_{0-} + J_{0+}$ lies in the space between elements of I_ϕ. We can also find $J_{1-} \subseteq I_1, J_{1+} \subseteq I_2$ satisfying the same properties for ψ_1. As $J_{0-} + J_{0+}$ and $J_{1-} + J_{1+}$ lie between elements of I_ϕ, these sequences are mutually indiscernible over I_ϕ. As a and b also lie in those intervals, we find that for any $a' \in J_{0-} + J_{0+}$ and $b' \in J_{1-} + J_{1+}$, $\phi(a', b') = 0$. This means that there exist e_0, e_1 such that $I_0 + e_0 + I_1 + I_2$ and $I_0 + I_1 + e_1 + I_2$ are indiscernible and $\phi(e_0, e_1) = 0$. Thus for $i = 0$ or $i = 1$, $\psi_i(e_i) \geq \varepsilon$. We now add that value of e_i into the sequence, maintaining indiscernibility, and repartition.

For the sake of simplicity, assume that e_1 is the added value. Then we repartition $J_{1-} + e_1 + J_{1+}$ as $J'_{1-} + J'_{1+}$, where J'_{1-} is a strict initial segment of J_{1-}. We repeat the earlier process, finding e'_0, e'_1 such that $J_{0-} + e_0 + J_{0+}$ and $J'_{1-} + J'_{1+}$ remain mutually indiscernible over I_0, as do $J_{0-} + J_{0+}$ and $J'_{1-} + e'_1 + J'_{1+}$, while maintaining $\phi(e'_0, e'_1) = 0$. Thus we add either e'_0 or e'_1, and repeat infinitely many times.

In conclusion, we have added infinitely many points to either J_{0-} or J_{1-}. Assume without loss of generality it was J_{1-}. Then we have an indiscernible sequence consisting of J_{1-} and the added points where the value of ψ_1 alternates infinitely many times between being $\psi_1(y) \leq \frac{\varepsilon}{2}$, as on all original values of J_{1-}, and $\psi_1(y) \geq \varepsilon$, as on all the new added points. This contradicts NIP. $\square$

In the rest of this section, we will generalize several other definitions of distality, in terms of types and formulas, to continuous logic. We will check that these are the correct generalizations by showing that these definitions are all equivalent to our first definition in terms of indiscernible sequences.

Theorem 3.5.5. *If a metric theory T is NIP, then the following are equivalent:*

1. *T is distal.*

2. *Every global type is distal.*

3. *Every formula admits strong honest definitions.*

4. *Every formula admits an ε-distal cell decomposition for each $\varepsilon > 0$.*

We will prove this over the following subsections by showing that $1 \implies 2, 2 \implies 3$, $3 \implies 4$, and $4 \implies 1$, introducing the definitions of distal types (Definition 3.5.6), strong honest definitions (Definition 3.5.8), and distal cell decompositions (Definition 3.5.17) as we go.

3.5.1 Distal Types

We now restate the definition of distal types in an NIP theory, which also works as-is in the continuous context.

Definition 3.5.6 (Distal types, [Sim15, Def. 9.3])**.** Assume T is NIP. Let p be a global A-invariant type. Then p is *distal* over A when for any tuple b, if $I \vDash p^{(\omega)}|_{Ab}$, then $p|_{AI}$ and $\mathrm{tp}(b/AI)$ are weakly orthogonal. (That means that there is a unique complete type over A extending $p(x) \cup q(y)$.)

If p is distal over all A such that p is invariant over A, then we just say that p is distal, without specifying A.

Theorem 3.5.7. *In a distal theory, all invariant types are distal.*

Proof. Let p be a global A-invariant type, let b be a tuple, and let $I \vDash p^{(\omega)}|_{Ab}$. We wish to show that $p|_{AI}$ is weakly orthogonal to $q(y) = \mathrm{tp}(b/AI)$. One such type is $\mathrm{tp}(a_p b/AI)$ for any

$a_p \models p|_{AIb}$, so for contradiction, assume there is some $a \models p|_{AI}$ such that $a \not\models p_{AIb}$. We then construct another Morley sequence. Let $J \models p^{(\omega)}|_{MIa}$. Then $I+a+J \models p^{(\omega+\omega)}|_A$, and is thus indiscernible over A, while $I+J \models p^{(\omega+\omega)}|_M$, and is thus indiscernible over $Ab \subseteq M$. For any $j \in J$, $j \models p|_{AIb}$, but $a \not\models p|_{AIb}$, so $I+a+J$ is not indiscernible over Ab, contradicting distality. $\square$

3.5.2 Strong Honest Definitions

We will now prove a series of versions of strong honest definitions. As with honest definitions, we start by assuming distality to show a version expressed in terms of pairs, derive a finitary version expressible without pairs, uniformize that finitary version using the (p,q)-theorem, and then prove distality from strong honest definitions, showing that all of these statements are equivalent.

There will be two basic ways to express strong honest definitions. The first is the continuous version of the version from [CS15, Prop. 19].

Definition 3.5.8. Let A be a closed subset of M^y where $M \preceq \mathcal{U}$ and $(M,A) \preceq (M',A')$. Let $\phi(x;y)$ be a definable predicate, let $a \in M$, and let $\theta(x;d)$ be an A'-predicate. We say that $\theta(x;d)$ is a *strong honest definition* for $\phi(a;y)$ over A when

- $M' \vDash \theta(a;d) = 0$
- For all $a' \in M'^x$, $b \in A$, $|\phi(a';b) - \phi(a;b)| \leq \theta(a';d)$.

For either of these definitions, if the same predicate $\theta(x;z)$ works for any choice of M, A, b, then we call $\theta(x;z)$ a strong honest definition for $\phi(x;y)$.

Essentially, $\theta(x;d)$ controls how much the type $\mathrm{tp}_\phi(x/A)$ differs from $\mathrm{tp}_\phi(a/A)$. In classical logic, when ϕ and θ only take values 0 and 1 corresponding to true and false, this definition is equivalent to $M' \vDash \theta(a;d)$ and $\theta(x;d) \vdash \mathrm{tp}_\phi(a/A)$. This is precisely the presentation of strong honest definitions in [CS15, Proposition 19]. We see that as in classical logic, strong honest definitions always exist in distal theories.

Theorem 3.5.9. *Assume T is distal. Let $M \models T$, $A \subseteq M$ closed, $\phi(x; y)$ a definable predicate, and $a \in M^x$. There is some elementary extension $(M, A) \preceq (M', A')$ such that $\phi(a; x)$ admits a strong honest definition $\theta(x; d)$ with $d \in A'^z$.*

Proof. As before, let $S_A \subseteq S_y(\mathcal{U})$ be the set of global types approximately realized in A. We will show that $\mathrm{tp}(a/A') \times S_A|_{A'} \vDash \phi(x; y) = \phi(a; y)$, and then extract the strong honest definition from there.

To do this, let $p(y) \in S_A$ be a global type. We claim that there is $b \in A'$ realizing p over MB for any small $B \subseteq A'$. By the saturation of $\mathcal{M}'$, it suffices to show that the type $p(y)|_{MB} \cup \{P(y) = 0\}$ is consistent. For this, it is enough to show that for every condition $\pi(y) = 0 \in p(y)|_{MB}$, and every $\varepsilon > 0$, $\pi(y) \leq \varepsilon$ is consistent with $P(y) = 0$. As $[\pi(y) < \varepsilon]$ is an open set containing $p(y)$, it must also intersect the set of realizations of A, and thus intersects $[P(y) = 0]$, so $\pi(y) \leq \varepsilon$ is consistent with $P(y) = 0$.

This allows us to construct a Morley sequence I for p over M in A', by recursively defining a_n to be an element of A' realizing $p|_{Ma_0 \ldots a_{n-1}}$. By Theorem 3.5.7, for any $p(y) \in S_A$, $p|_{AI}$ is weakly orthogonal to $\mathrm{tp}(a/AI)$, so $\mathrm{tp}(a/AI) \times p|_{AI} \vDash \phi(x; y) = \phi(a; y)$, and expanding the parameter sets, we see that $\mathrm{tp}(a/A') \times p|_{A'} \vDash \phi(x; y) = \phi(a; y)$. As this holds for all $p \in S_A$, the condition $\phi(x; y) = \phi(a; y)$ holds everywhere on $\mathrm{tp}(a/A') \times S_A|_{A'}$, so the predicate $|\phi(x; y) - \phi(a; y)|$ is zero.

We now apply Lemma 3.3.3 to the partial A' types $\mathrm{tp}(a/A')$ and $S_A|_{A'}(y)$ on (x, y) and the predicate $|\phi(x; y) - \phi(a; y)|$, and find a definable predicate $\theta(x; d)$ with $d \in A'$ such that $\mathrm{tp}(a/A')$ implies $\theta(x; d) = 0$ and for all b satisfying a type in $S_A|_{A'}(y)$, $|\phi(x; b) - \phi(a; b)| \leq \theta(x; d)$. In particular, for all $b \in A$, $|\phi(x; b) - \phi(a; b)| \leq \theta(x; d)$. □

There is another form of strong honest definitions, which is literally an honest definition in the sense of Definition 3.4.4. We call these "strong* honest definitions," as their existence is related to existence of strong honest definitions for the dual predicate.

Definition 3.5.10. Let A be a closed subset of M^x where $M \preceq \mathcal{U}$ and $(M, A) \preceq (M', A')$. Let $\phi(x; b)$ be an M-predicate, and let $\psi(x; d)$ be an A'-predicate. We say that $\psi(x; d)$ is a *strong* honest definition* for $\phi(x; b)$ over A when

- for all $a \in A$, $\phi(a; b) = \psi(a; d)$
- for all $a \in M'^x$, $\phi(a; b) \leq \psi(a; d)$.

If the same predicate $\psi(x; z)$ works for any choice of M, A, b, then we call $\psi(x; z)$ a strong* honest definition for $\phi(x; y)$.

For all $\phi(x; y)$ and $\psi(y; z)$, we also define a predicate

$$\mathrm{SHD}_{\phi,\psi,P}(y; z) = \max\left(\sup_{x:P(x)} |\phi(x; y) - \psi(x; z)|, \sup_{x} \phi(x; y) \dot{-} \psi(x; z)\right).$$

Then for $d \in A'^z$, $(M', A) \vDash \mathrm{SHD}_{\phi,\psi,P}(b; d)$ if and only if $\psi(x; d)$ is a strong* honest definition for $\phi(x; b)$. We will abuse notation later to write $\mathrm{SHD}_{\phi,\psi,A}(b; d)$ for the value of $\mathrm{SHD}_{\phi,\psi,P}(b; d)$ in (M', A).

We see that strong honest definitions imply the existence of strong* honest definitions for the dual predicate.

Lemma 3.5.11. *Let A be a closed subset of M^x where $M \preceq \mathcal{U}$ and $(M, A) \preceq (M', A')$. Let $\phi(x; b)$ be an M-predicate. If $\phi^*(b; x)$ admits a strong honest definition $\theta(y; d)$ over A, then $\phi(x; b)$ admits a strong* honest definition $\psi(x; d)$ over A, with the same parameters, and $\psi(x; z)$ depending only on $\theta(y; z)$.*

Proof. Let $\psi(x; d) = \sup_y(\phi(x; y) \dot{-} \theta(y; d))$. Thus for each $a \in M'^x$, we have $\psi(a; d) = \sup_y(\phi(a; y) \dot{-} \theta(y; d))$. By plugging in $y = b$, we see that $\psi(a; d) \geq \phi(a; b) \dot{-} \theta(b; d) = \phi(a; b)$.

Now let $a \in A$. For all $b' \in M'^y$, we have $|\phi(a;b') - \phi(a;b)| \leq \theta(b';d)$, so $\phi(a;b') \dot{-} \theta(b';d) \leq \phi(a;b)$, and thus $\psi(a;d) \leq \phi(a;b)$, so $\phi(a;b) = \psi(a;d)$. $\square$

We can recover strong honest definitions from strong* honest definitions for both the original predicate and its complement.

Lemma 3.5.12. *If $\phi(x;b)$ and $1 - \phi(x;b)$ admit strong* honest definitions over A then $\phi^*(b;x)$ admits a strong honest definition over A.*

Proof. Assume that $\phi(x;b)$ admits a strong* honest definition $\psi^+(x;d)$ over A, and $1-\phi(x;b)$ admits a strong* honest definition $\psi'(x;d)$ over A. Then by setting $\psi^-(x;d) = 1 - \psi'(x;d)$, we find that

- for all $a \in A$, $\phi(a;b) = \psi^-(a;d) = \psi^+(a;d)$
- for all $a \in M'^x$, $\psi^-(a;d) \leq \phi(a;b) \leq \psi^+(a;d)$.

Then we let $\theta(y;z) = \sup_x \max(\psi^-(x;z) \dot{-} \phi(x;y), \phi(x;y) \dot{-} \psi^+(x;z))$. For every $a \in M'^x$, we have that

$$\psi^-(a;d) \dot{-} \phi(a;b) = \phi(a;b) \dot{-} \psi^+(a;d) = 0,$$

so

$$\theta(b;d) = \sup_x \max(\psi^-(x;d) \dot{-} \phi(x;b), \phi(x;b) \dot{-} \psi^+(x;d)) = 0.$$

Now let $a \in A, b' \in M'^y$. We have that

$$\begin{aligned} |\phi(a;b) - \phi(a;b')| &= \max(\phi(a;b) \dot{-} \phi(a;b'), \phi(a;b') \dot{-} \phi(a;b)) \\ &= \max(\psi^-(a;d) \dot{-} \phi(a;b'), \phi(a;b') \dot{-} \psi^+(a;d)) \\ &\leq \theta(b';d). \end{aligned}$$

$\square$

As with honest definitions, we can take forced limits of approximate strong* honest definitions to get strong* honest definitions, and the proof is essentially the same.

Lemma 3.5.13. *Let A be a closed subset of M^y where $M \preceq \mathcal{U}$, and let $\phi(x;b)$ be an M-predicate. Let $(M,A) \preceq (M',A')$, and let $d \in A'^z$. Let $(\psi_n(x;z) : n \in \mathbb{N})$ be a sequence of definable predicates with $\mathrm{SHD}_{\phi,\psi_n,A}(b;d) \leq 2^{-n}$ for each n. Then $\mathcal{F}\lim \psi_n(x;d)$ is a strong* honest definition for $\phi(x;b)$ over A.*

If instead we have a sequence $(\psi_n(x;z_n) : n \in \mathbb{N})$ with different $d_n \in A'^{z_n}$ for each n such that $\mathrm{SHD}_{\phi,\psi_n,A}(b;d_n) \leq 2^{-n}$, then $\mathcal{F}\lim \psi_n(x;d)$ is a strong* honest definition for $\phi(x;b)$ over A, where d is a concatenation of all the tuples d_n.

We now deduce a finitary version of strong* honest definitions, without having to introduce an elementary extension. The proof is analogous to the proof of 3.4.7.

Lemma 3.5.14. *Let A be a closed subset of M^y where $M \preceq \mathcal{U}$, and let $\phi(x;b)$ be an M-predicate. Fix $(M,A) \preceq (M',A')$ to be $|M|^+$-saturated, $\varepsilon > 0$, and a definable predicate $\psi(x;z)$.*

If there exists $d \in A'^z$ such that $\mathrm{SHD}_{\phi,\psi,A}(b;d) < \varepsilon$, then for all finite $A_0 \subseteq A$, we have there is a tuple $d_{A_0} \in A^z$ such that $\mathrm{SHD}_{\phi,\psi,A_0}(b;d_{A_0}) < \varepsilon$.

Conversely, if for all finite $A_0 \subseteq A$, there is a tuple $d_{A_0} \in A^z$ with $\mathrm{SHD}_{\phi,\psi,A_0}(b;d_{A_0}) \leq \varepsilon$, then there exists $d \in A'^z$ such that $\mathrm{SHD}_{\phi,\psi,A}(b;d) \leq \varepsilon$.

We can now uniformize strong honest definitions using Lemma 3.5.14. The same argument used to prove 3.4.8 and then 3.4.9 applies again:

Theorem 3.5.15. *Assume T is distal. Every definable predicate $\phi(x;y)$ admits a strong* honest definition $\psi(x;z)$. That is, Then there is a definable predicate $\psi(x;z)$ such that for any $\mathcal{M} \models T$, $A \subseteq M$ closed with $|A| \geq 2$, and $b \in M^y$ with $|A| \geq 2$, there is some d such that $\psi(x;d)$ is a strong* honest definition for $\phi(x;y)$ over A.*

Finally, by 3.5.12, we can translate this back into a uniformized version of strong honest definitions.

Theorem 3.5.16. *Assume T is distal. Every definable predicate $\phi(x; y)$ admits a strong honest definition $\theta(y; z)$.*

3.5.3 Distal Cell Decompositions

While our finitary approximation to a strong* honest definition matches our notions for honest definitions, the finitary approximation to strong honest definitions will more closely resemble our approach to UDTFS. As we will use these for more combinatorial applications, we will use the conventions of distal cell decompositions from [CGS20].

Definition 3.5.17. Let $\phi(x; y)$ be a definable predicate, let Ψ be a finite set of definable predicates of the form $\psi(x; y_1, \ldots, y_k)$, where k is finite.

We say that Ψ *weakly defines* a *ε-distal cell decomposition* over M for $\phi(x; y)$ when for every finite $B \subseteq M^y$ with $|B| \geq 2$, there are sets $B_\psi \subseteq B$ for each $\psi \in \Psi$ such that the predicate $\sum_{\psi \in \Psi} \sum_{\bar{b} \in B_\psi} \psi(x; \bar{b})$ is always nonzero, and for each $\psi \in \Psi, \bar{b} \in B_\psi$ and $b \in B$, we have the bound

$$\sup_{x, x'} \min(\psi(x; \bar{b}), \psi(x'; \bar{b}), |\phi(x; b) - \phi(x'; b)| \dot{-} \varepsilon) = 0,$$

indicating that for all a, a' in the support of $\psi(x; \bar{b})$, $|\phi(a; b) - \phi(a'; b)| \leq \varepsilon$.

Let $\Theta = \{\theta_\psi : \psi \in \Psi\}$ where for each $\psi(x; y_1, \ldots, y_k) \in \Psi$, θ_ψ is a definable predicate of the form $\theta(y; y_1, \ldots, y_k)$.

We say that Ψ and Θ *define* a *ε-distal cell decomposition* over M for $\phi(x; y)$ when for every finite $B \subseteq M^y$ with $|B| \geq 2$, we may let $B_\psi = \{(b_1, \ldots, b_k) \in B^k : \forall b \in B, \theta_\psi(b; b_1, \ldots, b_k) = 0\}$ in the above definition.

To recover the classical logic definition from [CGS20], we may choose any $0 < \varepsilon < 1$ and let $\phi = 0$ or $\psi = 0$ denote truth, while $\theta_\psi = 0$ corresponds to falsity.

For most purposes, it suffices to find a weak definition for a distal cell decomposition, as then we can let

$$\theta_\psi(y;\bar{y}) = \sup_{x,x'} \min(\psi(x;\bar{y}), \psi(x';\bar{y}), |\phi(x;y) - \phi(x';y)| \dot{-} \varepsilon),$$

and $\Theta = \{\theta_\psi : \psi \in \Psi\}$ will finish defining the distal cell decomposition.

We justify this definition of distal cell decompositions by showing that their existence is equivalent to distality. First we show that distal cell decompositions follow from strong honest definitions, and then we will show that they imply distality, completing the cycle of equivalences.

Lemma 3.5.18. *Let $\phi(x;y)$ be a definable predicate such that $\phi(x;y)$ admits a strong honest definition. Then $\phi(x;y)$ admits a distal cell decomposition for all $\varepsilon > 0$.*

Proof. Let $\theta(x;z)$ be a strong honest definition for $\phi(x;y)$. Then by the density of formulas in definable predicates, let $\psi(x;z)$ be a formula which is always within $\frac{\varepsilon}{6}$ of $\frac{\varepsilon}{3} \dot{-} \theta(x;z)$.

Fix $B \subseteq M^y$. Then for each $a \in M^x$, there is a tuple d_a in B^z such that $\theta(a;d_a) = 0$, and for all $a' \in M^x$ and $b \in B$, $|\phi(a;b) - \phi(a';b)| \leq \theta(a';d_a)$. Thus $|\psi(x;z) - \frac{\varepsilon}{3}| \leq \frac{\varepsilon}{6}$, so $\psi(a;d_a) \geq \frac{\varepsilon}{6} > 0$. If $a' \in M^x$ is such that $\psi(a';d_a) > 0$, then $\theta(a';d_a) < \frac{\varepsilon}{2}$, so for all $b \in B$, $|\phi(a;b) - \phi(a';b)| \leq \frac{\varepsilon}{2}$, and thus for all $a_1, a_2 \in M^x$ such that $\psi(a_1;d_a) > 0$ and $\psi(a_2;d_a) > 0$, we have $|\phi(a_1;d_a) - \phi(a_2;d_a)| \leq \varepsilon$.

As $\psi(x;z)$ is a formula, it depends on only finitely many variables, so we may select $y_1, \ldots, y_k$ to be copies of y within z including all variables on which ψ depends. Then letting $\Psi = \{\psi(x;y_1,\ldots,y_k)\}$, we check that Ψ weakly defines a ε-distal cell decomposition. If B_ψ is the set of all $\bar{b}$ such that some d_a restricts to $\bar{b}$, we find that for all a, there is some $\bar{b} \in B_\psi$ such that $\psi(a;\bar{b}) > 0$, and for each $\bar{b} \in B_\psi$, and for all $a_1, a_2 \in M^x$ such that $\psi(a_1;\bar{b}) > 0$ and $\psi(a_2;\bar{b}) > 0$, we have $|\phi(a_1;\bar{b}) - \phi(a_2;\bar{b})| \leq \varepsilon$. $\square$

Theorem 3.5.19. *If a metric theory T is such that all formulas admit ε-distal cell decompositions for all $\varepsilon > 0$, then it is distal.*

Proof. Fix $I + d + J$ indiscernible with indiscernible over B and I, J infinite. We will show that $I + d + J$ is indiscernible over B. To do this, let a be a finite tuple from A.

Let ϕ be a formula, and without loss of generality, assume that $\phi(a; b_0, \dots, b_{2n}) = 0$ when $b_0 < \cdots < b_{2n}$ is an increasing sequence in $I + J$. Fix $\varepsilon > 0$. We will show that for any $b_0 < \cdots < b_{n-1} \in I, b_{n+1} < \cdots < b_{2n} \in J$, $\phi(a; b_0, \dots, b_{n-1}, d, b_{n+1}, \dots, b_{2n}) \leq \varepsilon$, implying that $I + d + J$ is A-indiscernible.

Let Ψ weakly define a ε-distal cell decomposition for $\phi(x; y_0, \dots, y_{2n})$. Fix a finite set $I_0 \subseteq I$ with $|I_0| \geq |z| + 2(2n + 1)$. Then there is some $\psi(x; y_1, \dots, y_k) \in \Psi$ and some tuple $\bar{b} \in I^k$ such that $\psi(a, \bar{b}) > 0$ and for all a' with $\psi(a'; \bar{b}) > 0$, for all $\bar{b}' \in I^{2n+1}$, $\phi(a; \bar{b}') \leq \varepsilon$. Thus

$$\sup_x \max(\psi(x; \bar{b}), \varepsilon \dot{-} \phi(x; \bar{b}')) = 0.$$

Because I_0 is large, there is an increasing sequence $b_0 < \cdots < b_{2n}$ in I_0 disjoint from $\bar{b}$, and in particular, all of the elements in the sequence are either less than or greater than the entire tuple $\bar{b}$.

Now let $b'_0, \dots, b'_{n-1} \in I$, $b'_{n+1}, \dots, b'_{2n} \in J$, and we will show that

$$\phi(a; b'_0, \dots, b'_{n-1}, d, b'_{n+1}, \dots, b'_{2n}) \leq \varepsilon.$$

There is some tuple $\bar{b}' \in I + J$ such that the order type of $b'_0, \dots, b'_{n-1}, d, b'_{n+1}, \dots, b'_{2n}, \bar{b}'$ is the same as $b_0, \dots, b_{2n}, \bar{b}$. By the indiscernibility of $I + d + J$, we find that

$$\sup_x \max(\psi(x; \bar{b}'), \varepsilon \dot{-} \phi(x; b'_0, \dots, b'_{n-1}, d, b'_{n+1}, \dots, b'_{2n})) =$$
$$\sup_x \max(\psi(x; \bar{b}), \varepsilon \dot{-} \phi(x; b_0, \dots, b_{2n})) = 0,$$

and by the indiscernibility of $I + J$ over A, we have $\psi(a; \bar{b}') > 0$, so

$$\phi(a; b'_0, \dots, b'_{n-1}, d, b'_{n+1}, \dots, b'_{2n}) \leq \varepsilon,$$

as desired. $\square$

3.5.4 Reductions

Having seen that all of these properties are equivalent to distality, we now provide some more ways to check whether a theory is distal.

We will show that the property of admitting strong honest definitions is closed under continuous combinations, which means that given quantifier elimination, it suffices to check that atomic formulas admit strong honest definitions.

Lemma 3.5.20. *Let $\phi_1(x; y), \dots, \phi_n(x; y)$ be formulas that admit strong honest definitions. Let $u : [0,1]^n \to [0,1]$ be continuous. Then $\phi(x; y) = u(\phi_1(x; y), \dots, \phi_n(x; y))$ admits a strong honest definition.*

Proof. Define $F, G : ([0,1]^n \times [0,1]^n) \to [0,1]$ as follows, using the ℓ_∞-norm on $[0,1]^n$. Let $F(a, a') = |a - a'|_{\ell_\infty}$ and $G(a, a') = |u(a) - u(a')|$. As u is continuous between two compact metric spaces, it is uniformly continuous, so for each $\varepsilon > 0$, there is some $\delta > 0$ such that $|a - a'|_{\ell_\infty} \leq \delta$ implies $|u(a) - u(a')| \leq \varepsilon$. Thus by [BBH08, Proposition 2.10], there is some increasing continuous $\alpha : [0,1] \to [0,1]$ such that $\alpha(0) = 0$ and $\forall a, a', G(a, a') \leq \alpha(F(a, a'))$.

Now for $1 \leq i \leq n$, let $\theta_i(x; z)$ be a strong honest definition for $\phi_i(x; y)$. Then let $\theta(x; z_1, \dots, z_n) = \alpha\left(\min_{1 \leq i \leq n} \theta_i(x; z_i)\right)$. We check that θ is a strong honest definition for ϕ.

Let $M \preceq \mathcal{U}$, A closed in M^y, $a \in M^x$, $(M, A) \preceq (M', A')$ be sufficiently saturated. Let $d_1, \dots, d_n \in A'$ be such that for each i, $\theta_i(x; d_i)$ is a strong honest definition for $\phi_i(a; y)$ over A. Then we see that

$$\theta(a; d) = \alpha\left(\max_{1 \leq i \leq n} \theta_i(a; z_i)\right) = 0.$$

Now let $a' \in M'^x$, $b \in A$. We see that

$$\begin{aligned}
|\phi(a';b) - \phi(a;b)| &= G(\phi_1(a;b), \dots, \phi_n(a;b), \phi_1(a';b), \dots, \phi_n(a';b)) \\
&\leq \alpha(F(\phi_1(a;b), \dots, \phi_n(a;b), \phi_1(a';b), \dots, \phi_n(a';b))) \\
&= \alpha\left(\max_{1 \leq i \leq n} |\phi_i(a';b) - \phi_i(a;b)|\right) \\
&\leq \alpha\left(\max_{1 \leq i \leq n} \theta_i(a';d_i)\right) \\
&= \theta(a';d).
\end{aligned}$$

□

Corollary 3.5.21. *As a corollary, we see that if T eliminates quantifiers and all atomic formulas admit strong honest definitions, then T is distal.*

We can also reduce to one variable.

Theorem 3.5.22. *Let T be an NIP theory. Then T is distal if and only if any of the following equivalent conditions hold:*

- *Any indiscernible $I + d + J$ with $I + J$ indiscernible over a singleton b is indiscernible over b*
- *For any $A \subset \mathcal{M}$, global A-invariant type p and singleton b, if $I \models p^{(\omega)}|_{Ab}$, then $p|_{AI}$ and $\mathrm{tp}(b/AI)$ are weakly orthogonal*
- *Any predicate $\phi(x;y)$ with $|x| = 1$ admits a strong honest definition*
- *Any predicate $\phi(x;y)$ with $|x| = 1$ admits an ε-distal cell decomposition for every $\varepsilon > 0$.*

Proof. Clearly distality implies all of these conditions.

These conditions are all equivalent by following the proofs of the implications in 3.5.5 and keeping track of the length of tuples. We will prove that the indiscernible condition implies distality and the strong honest definition condition implies distality. The first proof is more straightforward, but we will also construct explicit strong honest definitions for predicates with more variables, generalizing the constructions in [And23a, Theorem 3.1] and [ACG22, Proposition 1.9].

First we show that if distality fails, the first condition fails. Let $I + d + J$ be indiscernible, and let b be a tuple such that $I + J$ is indiscernible over b, but $I + d + J$ is not indiscernible over b. Then $I + d + J$ it is not indiscernible over some finite subtuple of b, and we may assume b is finite. Let n be minimal such that there exists $b = (b_1, \ldots, b_n)$ satisfying these properties.

For a sequence S and a tuple b', let $S^\frown b'$ be the tuple obtained by concatenating b' to each term of S. Then S is indiscernible over b' if and only if $S^\frown b'$ is indiscernible.

We know that $I + d + J$ is indiscernible over $(b_1, \ldots, b_{n-1})$, so $(I + d + J)^\frown(b_1, \ldots, b_{n-1})$ is indiscernible, and $(I + J)^\frown(b_1, \ldots, b_{n-1})$ is indiscernible over b_n, but $(I + d + J)^\frown(b_1, \ldots, b_{n-1})$ is not indiscernible over b_n, so this sequence fails the first criterion over the singleton b_n.

Now we provide an explicit construction of strong honest definitions. Let T be a theory in which every definable predicate $\phi(x; y)$ with $|x| = 1$ admits a strong honest definition. To show that every definable predicate $\phi(x; y)$ admits a strong honest definition, it suffices to show it for all predicates with $|x|$ finite, as every predicate is a uniform limit of such predicates, and by Lemma 3.5.13, uniform limits of predicates with strong honest definitions have strong honest definitions.

Assume for induction that this holds for every definable predicate with $|x| \leq n$, and let $\phi(x_0, x; y)$ be a definable predicate with $|x| = n$. We will now repartition the variables of ϕ several ways, and find strong(*) honest definitions for each repartition. Then by assumption, there exists a strong honest definition $\theta_0(x_0; z_0)$ for $\phi(x_0; x, y)$. As z_0 is a (possibly countable) tuple of copies of (x, y), and we will be interested in considering θ_0 as a strong honest definition over sets of the form $\{a\} \times A$ for $A \subseteq M^y$, we will assume that each copy of x is equal, and write the predicate as $\theta_0(x_0; x, z_0)$, where z_0 is a tuple of copies of y. Then we let $\psi^+(x; y, z_0) = \sup_{x_0} \phi(x_0, x; y) \dot{-} \theta_0(x_0; x, z_0)$, and $\psi^-(x; y, z_0) = 1 - \sup_{x_0}(1 - \phi(x_0, x; y)) \dot{-} \theta_0(x_0; x, z_0)$. As $|x| = n$, there are also strong honest definitions $\theta^+(x; z_+), \theta^-(x; z_-)$ for $\psi^+(x; y, z_0), \psi^-(x; y, z_0)$ respectively.

We claim that $\theta(x_0, x; z_0, z_+, z_-) = \theta_0(x_0; x, z_0) + \theta^+(x; z_+) + \theta^-(x; z_-)$ is a strong honest definition for $\phi(x_0, x; y)$. Now fix $A \subseteq M^y$, $a_0 \in M, a \in M^x$. Let d_0 be such that $\theta_0(x_0; a, d_0)$ is a strong honest definition for $\phi(a_0; x, y)$ over $\{a\} \times A$, and let $d_\pm$ be such that $\theta^\pm(x; d_\pm)$ is a strong honest definition for $\psi^\pm(a; y, z_0)$ over $A \times \{d_0\}$. By definition, we will have $\theta(a_0, a; d_0, d'_+, d'_-) = 0 + 0 + 0$. For any $a'_0 \in M$, as $\theta_0(a'_0; a, d_0) \geq |\phi(a'_0, a; b) - \phi(a_0, a; b)|$ and thus $\phi(a'_0, a; b) \leq \phi(a_0, a; b) + \theta_0(a'_0; a, d_0)$, we have $\psi_+(a; b, d_0) = \sup_{x_0} \phi(a'_0, a; b) \dot{-} \theta_0(a'_0; a, d_0) \leq \phi(a_0, a; b)$. A similar calculation shows that $\psi_-(a; b, d_0) \geq \phi(a_0, a; b)$.

Now let $a'_0 \in M$, $a' \in M^n$, and $b \in A$, and we will show that $|\phi(a_0, a; b) - \phi(a'_0, a'; b)| \leq \theta(a'_0, a'; d)$. First we will show that $\phi(a'_0, a'; b) \leq \phi(a_0, a; b) + \theta(a'_0, a'; d)$.

We see that $|\psi^+(a'; b, d_0) - \psi^+(a; b, d_0)| \leq \theta^+(a'; d_+)$, so

$$\phi(a'_0, a'; b) \dot{-} \theta_0(a'_0; a', d_0) \leq \psi^+(a'; b, d_0) \leq \psi^+(a; b, d_0) + \theta^+(a'; d_+) \leq \phi(a_0, a; b) + \theta^+(a'; d_+)$$

and thus

$$\phi(a'_0, a'; b) \leq \phi(a_0, a; b) + \theta_0(a'_0; a', d_0) + \theta^+(a'; d_+) \leq \phi(a_0, a; b) + \theta(a'_0; a', d_0, d_+, d_-).$$

By similar logic,

$$\phi(a'_0, a'; b) \geq \phi(a_0, a; b) - \theta_0(a'_0; a', d_0) - \theta^-(a'; d_-) \geq \phi(a_0, a; b) - \theta(a'_0; a', d_0, d_+, d_-).$$

$\square$

CHAPTER 4

Generically Stable Measures and Distal Regularity in Continuous Logic

In this chapter, we develop a theory of generically stable and smooth Keisler measures in NIP metric theories, generalizing the case of classical logic. Using smooth extensions, we verify that fundamental properties of (Borel)-definable measures and the Morley product hold in the NIP metric setting. With these results, we prove that as in discrete logic, generic stability can be defined equivalently through definability properties, statistical properties, or behavior under the Morley product. We also examine weakly orthogonal Keisler measures, characterizing weak orthogonality in terms of various analytic regularity properties.

We then examine Keisler measures in distal metric theories, proving that as in discrete logic, distality is characterized by all generically stable measures being smooth, or by all pairs of generically stable measures being weakly orthogonal. We then use this, together with our results on weak orthogonality and a cutting lemma, to find analytic versions of distal regularity and the strong Erdős-Hajnal property.

4.1 Introduction

This chapter continues the study of distal theories in continuous logic begun in Chapter 3. In that chapter, we characterized distal metric structures in terms of the behavior of their indiscernible sequences and a continuous version of strong honest definitions, generalizing [Sim13] and [CS15]. It is just as fundamental to define distal structures as those structures where all generically stable Keisler measures are smooth.

Keisler measures, as a real-valued generalization of types, lend themselves naturally to continuous logic. Despite this, while many properties of types such as definability, finite satisfiability, and generic stability have been generalized both to Keisler measures[HPS13] and to types in continuous logic[Ben10b][CGH23b][Kha22], the literature is comparatively lacking in simultaneous generalizations to measures in continuous logic. Thus before we can examine distal metric structures from a Keisler measure perspective, we must generalize these properties, extending the theory of Keisler measures over metric structures from papers such as [BK09], [Ben09], and [CCP24].

Once we understand generically stable Keisler measures in continuous logic, and prove that the Keisler measure definition of distality is equivalent to all other definitions for metric structures, we may use these measures for combinatorial applications of distality. We develop continuous logic versions of the distal regularity lemma and (definable) strong Erdős-Hajnal property of [CS18]. A forthcoming paper with Ben Yaacov will provide several examples of metric structures to which these results apply[AB24].

This contributes to a growing subject of "tame regularity" in the analytic setting. Analytic regularity lemmas replace the graphs of Szemerédi's original regularity lemma with real-valued functions, which are decomposed into structured, pseudorandom, and error parts[LS07]. Under a tameness assumption, such as the function being definable in an NIP[LS10], n-dependent[CT20], or stable[CCP24] metric structure, this decomposition can be simplified.

The distal analytic regularity lemma, Theorem 4.5.5, implies that for every $\varepsilon > 0$, any definable predicate $\phi(x_1, \ldots, x_n)$ in a distal structure can be expressed as a the sum of a structured part of bounded complexity and a particularly well-behaved error part, which is bounded in magnitude by ε everywhere except on a structured set of small measure.

This in turn implies an analytic version of the strong Erdős-Hajnal property: We say that a predicate $\phi(x_1, \dots, x_n)$ has the *strong Erdős-Hajnal property* in some structure M when for every $\varepsilon > 0$, there exists $\delta > 0$ such that for any finite sets $A_i \subseteq M^{x_i}$, there are subsets $B_i \subseteq A_i$ such that $|B_i| \geq \delta|A_i|$ and for all $b, b' \in B_1 \times \cdots \times B_n$, $|\phi(b) - \phi(b')| \leq \varepsilon$. Just as [CS18, Theorem 3.1] proves in the discrete case, we show that in continuous logic, distality is equivalent to every definable predicate having a definable version of the strong Erdős-Hajnal property, where the counting measures on the sets A_i can be replaced with generically stable Keisler measures, and the sets B_i can be defined uniformly.

Overview and Results

Section 4.2 lays out the basic theory of Keisler measures in continuous logic. These can be understood either as regular Borel measures on the space of types, or equivalently, as certain linear functionals on the space of definable predicates[CCP24]. Most importantly for studying distality, we study weak orthogonality and smooth measures, following the approach of [Sim16]. We characterize weakly orthogonal measures as those where the following equivalent conditions hold:

Corollary 4.1.1 (Corollary 4.2.21)**.** *Let $x_1, \dots, x_n$ be variable tuples, and let $\mu_i \in \mathfrak{M}_{x_i}(M)$ be Keisler measures on x_i for each i. The measures μ_i are weakly orthogonal, meaning that there is a unique measure $\omega \in \mathfrak{M}_{x_1,\dots,x_n}(M)$ on $(x_1, \dots, x_n)$ extending the product measure of $\mu_1, \dots, \mu_n$, if and only if for every M-definable predicate $\phi(x_1, \dots, x_n)$ and every $\varepsilon > 0$, there exist M-definable predicates $\psi^-(x_1, \dots, x_n), \psi^+(x_1, \dots, x_n)$, where $\psi^\pm(x_1, \dots, x_n)$ are each of the form $\sum_{j=1}^m \prod_{i=1}^n \theta_{ij}^\pm(x_i)$, such that*

- *For all $(x_1, \dots, x_n)$, $\psi^-(x_1, \dots, x_n) \leq \phi(x_1, \dots, x_n) \leq \psi^+(x_1, \dots, x_n)$.*
- *For any product measure ω of $\mu_1, \dots, \mu_n$, $\int_{S_{x_1 \dots x_n}(M)} (\psi^+ - \psi^-)\, d\omega \leq \varepsilon$.*

From this perspective, we consider smooth measures - all measures such that there is a small model M such that $\mu|_M$ has a unique global extension. We characterize them also as the measures that are weakly orthogonal to all measures, or equivalently all types. We also examine invariant and (Borel)-definable measures, extending the careful work in [CGH23a] on Morley products in NIP to continuous logic. This approach revolves around the fact that any measure in an NIP theory admits a smooth extension, which we verify for continuous logic in Lemma 4.2.30. We are then able to use smooth extensions of measures as we would use realizations of types.

In Section 4.3, we turn to generically stable measures, finding many equivalent continuous logic characterizations of these versatile measures, culminating with a generalization of [HPS13, Theorem 3.2] to continuous logic:

Theorem 4.1.2 (Thm 4.3.1)**.** *Assume T is an NIP metric theory. For any small model $M \subseteq \mathcal{U}$, if μ is a global M-invariant measure, the following are equivalent:*

1. *μ is a frequency interpretation measure (fim) over M (see Definition 4.2.2)*
2. *μ is a finitely approximated measure (fam) over M (see Definition 4.2.2)*
3. *μ is definable over and approximately realized in M (see Definition 4.2.2)*
4. *$\mu(x) \otimes \mu(y) = \mu(y) \otimes \mu(x)$ (see Definition 4.2.8)*
5. *$\mu^{(\omega)}(x_0, x_1, \dots)|_M$ is totally indiscernible (see Definition 4.2.8).*

This connects the topological properties of generically stable measures (definability and approximate realizability) and the behavior of the Morley product to the property of being a frequency interpretation measure (fim). Classically, these are measures against which formulas obey a version of the VC-theorem. We show that in continuous logic, definable predicates and generically stable measures satisfy various properties that were shown in [And23b, Section 3.2] for definable predicates and finitely-supported measures. This includes a Glivenko-Cantelli property analogous to the VC-Theorem (our definition of fim) as well as bounds on the sizes of ε-approximations (Corollary 4.3.7) and ε-nets (Theorem 4.3.11).

Before approaching distal regularity directly, we connect weak orthogonality of measures to regularity properties in Section 4.4. In [Sim16], the distal regularity lemma are proven using weak orthogonality. Before assuming distality, we develop the nomenclature for expressing this regularity lemma and the (definable) strong Erdős-Hajnal property in continuous logic, generalizing [CS18, Theorems 3.1 and 5.8] in the discrete case, and we are able to prove non-uniform versions of these regularity lemmas for any weakly orthogonal measures:

Theorem 4.1.3 (Theorem 4.4.10)**.** *Let $\mu_1 \in \mathfrak{M}_{x_1}(M), \dots, \mu_n \in \mathfrak{M}_{x_n}(M)$. The following are equivalent:*

- *The measures $\mu_1, \dots, \mu_n$ are weakly orthogonal.*
- *For each M-definable predicate $\phi(x_1, \dots, x_n)$ and each $\varepsilon, \delta > 0$, there is some C such that ϕ admits a definable (ε, δ)-distal regularity partition (see Definitions 4.4.1 and 4.4.9)*
- *For each M-definable predicate $\phi(x_1, \dots, x_n)$ and each $\varepsilon, \delta > 0$, there is some C such that ϕ admits a constructible (ε, δ)-distal regularity partition (see Definitions 4.4.1 and 4.4.9)*
- *For each M-definable predicate $\phi(x_1, \dots, x_n)$ and each $\varepsilon > \gamma \geq 0$, there is some $\delta > 0$ such that for any product measure ω of continuous localizations of $\mu_1, \dots, \mu_n$, if $\int_{S_{x_1 \dots x_n}(M)} \phi \, d\omega \geq \varepsilon$, then there are M-definable predicates $\psi_i(x_i)$ such that $\psi_i(a_i) > 0$ for each i, $\phi(a_1, \dots, a_n) \geq \gamma$, and for each i, $\int_{S_{x_i}(M)} \psi_i(x_i) \, d\mu_i \geq \delta$.*
- *For each M-definable predicate $\phi(x_1, \dots, x_n)$ and each $\varepsilon > 0$, ϕ has the definable ε-SEH with respect to any continuous localizations of $\mu_1, \dots, \mu_n$ (see Definition 4.4.8).*

Furthermore, if these hold, then the (ε, δ)-distal regularity partitions can be chosen to be grid partitions of size $O(\delta^{-C})$ for some constant C depending on $\phi, \varepsilon, \mu_1, \dots, \mu_n$.

Having explored Keisler measures in NIP metric structures, we turn to distality in Section 4.5. First we characterize distal metric theories in terms of Keisler measures:

Theorem 4.1.4 (Theorem 4.5.1)**.** *The following are equivalent:*

- *The theory T is distal*
- *Every generically stable measure is smooth*
- *All pairs of generically stable measures are weakly orthogonal.*

We then apply distality to the regularity results of Section 4.4, showing that the regularity lemmas hold uniformly, getting a continuous logic version of the distal regularity lemma from [CS18]:

Theorem 4.1.5 (Theorem 4.5.5)**.** *Assume T is distal.*

For every definable predicate $\phi(x_1, \ldots, x_n; y)$ and $\varepsilon > 0$, there exist predicates $\psi_i(x_i; z_i)$, which can be chosen to be either definable or constructible, and a constant C such that if $\mu_1 \in \mathfrak{M}_{x_1}(M), \ldots, \mu_n \in \mathfrak{M}_{x_n}(M)$ are such that for $i < n$, μ_i is generically stable, $b \in M^y$, and $\delta > 0$, the following all hold: The predicate $\prod_{i=1}^n \psi_i(x_i; z_i)$ defines a (ε, δ)-distal regularity grid partition for $\phi(x_1, \ldots, x_n; b)$ of size $O(\delta^{-C})$.

4.2 Keisler Measures in Metric Theories

In this section, we translate some of the theory of Keisler measures to continuous logic, building on the definitions in [BK09], [Ben09], and [CCP24]. Throughout, T will be a complete metric theory with a monster model $\mathcal{U}$, and for any $A \subseteq \mathcal{U}$, $S_x(A)$ will be the space of types in variables x with parameters in A. For background on metric structures, see Chapter 3.

Definition 4.2.1. A *Keisler measure* on $S_x(A)$ is a regular Borel probability measure on $S_x(A)$. We denote the space of such measures $\mathfrak{M}_x(A)$.

It is noted in [CCP24] that these are in bijection with *Keisler functionals*, that is, positive linear functionals on $\mathcal{C}(S_x(A), \mathbb{R})$ with $\|f\| = 1$. We give the space $\mathfrak{M}_x(A)$ of Keisler measures the weak* topology as positive linear functionals, the coarsest topology such that every definable predicate $\phi(x)$ with parameters in A, $\mu \mapsto \int_{S_x(A)} \phi(x)\, d\mu$ is continuous. This generalizes the *compact Hausdorff topology* used for Keisler measures in classical logic in [Gan20]. We also see that for every definable predicate $\phi(x)$ with parameters in A, as $\phi(x)$ is the uniform limit of a sequence of formulas, $\mu \mapsto \int_{S_x(A)} \phi(x)\, d\mu$ is the uniform limit of a sequence of integrals of formulas, each of which is a continuous function, and is thus continuous.

We now present continuous analogs for several key properties that global Keisler measures (measures in $\mathfrak{M}_x(\mathcal{U})$) can have.

Definition 4.2.2. Let μ be a global Keisler measure, and let $A \subseteq \mathcal{U}$ be a small set, and $M \preceq \mathcal{U}$ a small model.

- We say μ is A-*invariant* when for any tuples $a \equiv_A b$ in $\mathcal{U}^y$, and any formula $\phi(x; y) \in \mathcal{L}(A)$, $\int \phi(x; a)\, d\mu = \int \phi(x; b)\, d\mu$. Equivalently, any automorphism of $\mathcal{U}$ fixing A preserves μ.

- If μ is A-invariant, define the map $F^{\phi}_{\mu,A} : S_y(A) \to [0, 1]$ by $F^{\phi}_{\mu,A}(p) = \int \phi(x; b)\, d\mu$ for $b \models p$.

- We say μ is A-*Borel definable* when it is A-*invariant* and for all $\phi(x; y) \in \mathcal{L}(A)$, the map $F^{\phi}_{\mu,A}$ is Borel.

- We say μ is A-*definable* when it is A-*invariant* and for all $\phi(x; y) \in \mathcal{L}(A)$, the map $F^{\phi}_{\mu,A}$ is continuous (and thus a definable predicate).

- We say μ is *approximately realized in* A when μ is in the topological closure of the convex hull of the Dirac measures at types of points in A. This corresponds to finite satisfiability.

- Keeping discrete notation, we call a definable, approximately realized measure *dfs* (for definable, finitely satisfiable).

- We say μ is *finitely approximated in* M when for every $\varphi(x;y) \in \mathcal{L}(M)$ and every $\varepsilon > 0$, there exists a tuple $(a_1, \dots, a_n) \in (M^x)^n$ which is a ε-approximation for the family $\{\varphi(x;b) : b \in \mathcal{U}^y\}$ with respect to μ. We abbreviate this property as *fam*.

- We say μ is a *frequency interpretation measure over* M when for every $\varphi(x;y) \in \mathcal{L}(M)$, there is a family of formulas $(\theta_n(x_1, \dots, x_n) : n \in \omega)$ with parameters in M such that $\lim_{n \to \infty} \mu^{(n)}(\theta_n(x_1, \dots, x_n)) = 1$, and for every $\varepsilon > 0$, for large enough n, any $\bar{a} \in (\mathcal{U}^x)^n$ satisfying $\theta_n(\bar{a})$ is a ε-approximation to $\varphi(x;y)$ with respect to μ. We abbreviate this property as *fim*.

- We say μ is *smooth* over M when for every N with $M \preceq N$, there exists a unique extension $\mu' \in \mathfrak{M}_x(N)$ of $\mu|_M$.

Note that if μ is A-invariant, then $F^{\phi}_{\mu,A}$ can also be defined for $\phi(x;y)$ a definable predicate. Any definable predicate is a uniform limit of formulas, so not only will $F^{\phi}_{\mu,A}$ be well-defined, but it will be the uniform limit of functions of the form $F^{\psi}_{\mu,A}$ where ψ is a formula. Thus if μ is A-Borel definable, the function $F^{\phi}_{\mu,A}$ will be Borel for ϕ a definable predicate, and if μ is A-definable, $F^{\phi}_{\mu,A}$ will be continuous. While we will often prove results for Borel definable measures for full generality, we will eventually show that in the NIP context, these are the same as invariant measures (see Lemma 4.2.7).

We will need to be able to consider sequences which are indiscernible with respect to μ in a certain sense, for which we will need the following definitions.

Definition 4.2.3. Let $\mathcal{L}_{\mathbb{E}}$ be an extension of the language $\mathcal{L}$ to add a relation symbol $\mathbb{E}_{\psi}(y)$ for each restricted formula $\psi(x;y)$, with $\mathbb{E}_{\psi}(y)$ having the same Lipschitz constant as ψ.

If M is a model and $\mu \in \mathfrak{M}_x(M)$, let $(M;\mu)$ be the $\mathcal{L}_{\mathbb{E}}$-structure so that for all $b \in M^y$, $\mathbb{E}_{\psi}(b) = \int_{S_x(M)} \psi(x;b)\, d\mu$.

The metric structure $(M;\mu)$ is valid because the integral of a C-Lipschitz function is also C-Lipschitz. Then by density and the fact that uniform limits commute with integrals, for any $\mathcal{L}$-definable predicate $\psi(x;y)$, we can define a $\mathcal{L}_{\mathbb{E}}$-definable predicate $\mathbb{E}_\psi(y)$ interpreted as $\int_{S_x(M)} \psi(x;y)\, d\mu$.

Lemma 4.2.4 (Generalizes [Sim15, Prop 7.5])**.** *Let $\mu \in \mathfrak{M}_x(\mathcal{U})$ be a global measure, $(b_i : i < \omega)$ an indiscernible sequence. Let $\phi(x;y)$ be a formula, and let $0 < s < r$ be such that*

$$\int_{S_x(\mathcal{U})} (\phi(x;b_i))\, d\mu \geq r$$

for all $i < \omega$. Then the partial type $\{\phi(x;b_i) \geq s : i < \omega\}$ is consistent.

Proof. We can use Ramsey and compactness to extract an $\mathcal{L}_{\mathbb{E}}$-indiscernible in an elementary extension of $(M;\mu)$ satisfying the EM-type of $(b_i : i < \omega)$. In particular, for every formula $\psi(x;y_1,\ldots,y_n)$ (not just the restricted ones), $\int_{S_x(M')} \psi(x;b_{i_1},\ldots,b_{in})\, d\mu'$ takes the same value for all $i_1 < \cdots < i_n \in \mathbb{N}$. Thus we can assume that the sequence $(b_i : i < \omega)$ was already indiscernible in this extended language.

Assume for contradiction that $\{\phi(x;b_i) \geq s : i < \omega\}$ is inconsistent. Thus for some N,

$$\min_{i=0}^{N} \left(\phi(x;b_i) \dot{-} s\right) = 0$$

indentically, and in particular,

$$\int_{S_x(M)} \min_{i=0}^{N} \left(\phi(x;b_i) \dot{-} s\right)\, d\mu = 0.$$

Let N be the minimal such value, and let $t = \int_{S_x(M)} \min_{i=0}^{N-1} \left(\phi(x;b_i) \dot{-} s\right)\, d\mu'$. Note that $t > 0$, as

$$t = \int_{S_x(M)} \min_{i=0}^{N-1} \left(\phi(x;b_i) \dot{-} s\right)\, d\mu' \geq \int_{S_x(M)} \phi(x;b_0)\, d\mu' - s = r - s > 0.$$

Then define

$$\psi_k(x) = \min_{i=0}^{N-1} \left(\phi(x;b_i) \dot{-} s\right),$$

and observe that

$$\min(\psi_0(x), \psi_1(x)) = \min_{i=0}^{2N-1} \left(\phi(x;b_i) \dot{-} s\right) \leq \min_{i=0}^{N} \left(\phi(x;b_i) \dot{-} s\right) = 0,$$

so by indiscernibility, for all $i < j$, $\int_{S_x(M)} \min(\psi_i(x), \psi_j(x))\, d\mu = 0$.

Thus for any indices $i_1 < \cdots < i_m$,

$$\int_{S_x(M)} \max_{1 \leq j \leq m} \psi_{i_j}(x)\, d\mu = \int_{S_x(M)} \sum_{j=1}^{m} \psi_{i_j}(x)\, d\mu = mt$$

and for $m > \frac{1}{t}$, this gives $\int_{S_x(M)} \max_{1 \leq j \leq m} \psi_{i_j}(x)\, d\mu' > 1$, a contradiction because we can bound $\max_{1 \leq j \leq m} \psi_{i_j}(x) \leq 1$. □

Definition 4.2.5. For any measure $\mu \in \mathfrak{M}_x(A)$, define $S_\mu(x)$ to be the partial type consisting of all closed A-conditions with μ-measure 1. We also define $S(\mu) \subseteq S_x(A)$ to be the set of all types satisfying $S_\mu(x)$, we call this *the support* of μ.

Clearly the intersection of finitely many closed conditions in $S_\mu(x)$ has μ-measure 1, so any finite subtype of $S_\mu(x)$ is satisfiable.

Lemma 4.2.6. *Assume that T is NIP. Let $A \subset \mathcal{U}$ be such that $\mathcal{U}$ is $|A|^+$-saturated, let $\mathfrak{M}_x(\mathcal{U})$ be an A-invariant measure, and $p(x) \in S(\mu)$. Then p is A-invariant, meaning that for any formula $\phi(x;y)$, and any $b, b' \in \mathcal{U}$ with $b \equiv_A b'$, $(\phi(x;b) = \phi(x;b')) \in p(x)$.*

Proof. Let $p(x) \in S(\mu)$ and let $\phi(x;y)$ be an A-formula. Then for any $b \equiv_A b'$, by A-invariance, $\int_{S_x(\mathcal{U})} \phi(x;b)\, d\mu = \int_{S_x(\mathcal{U})} \phi(x;b')\, d\mu$. Assume $p(x)$ is not A-invariant. Then there exist $b \equiv_A b'$ with $(\phi(x;b) = \phi(x;b')) \notin p(x)$. Then without loss of generality, there is some

$\varepsilon > 0$ with $\phi(x; b) = \phi(x; b') + \varepsilon \in p(x)$. Meanwhile, $\int_{S_x(\mathcal{U})} \phi(x; b)\, d\mu = \int_{S_x(\mathcal{U})} \phi(x; b')\, d\mu$. We will show that $\int_{S_x(\mathcal{U})} |\phi(x; b) - \phi(x; b')|\, d\mu = 0$.

Assume for contradiction that $\int_{S_x(\mathcal{U})} |\phi(x; b) - \phi(x; b')|\, d\mu = \varepsilon > 0$. Then because $b \equiv_A b'$, we may find an A-indiscernible sequence $(b_i : i < \omega)$ with $b_0 = b$, $b_1 = b'$. For all i, we have $b_{2i}b_{2i+1} \equiv_A bb'$, and by invariance of μ, $\int_{S_x(\mathcal{U})} |\phi(x; b_{2i}) - \phi(x; b_{2i+1})|\, d\mu = \int_{S_x(\mathcal{U})} |\phi(x; b) - \phi(x; b')|\, d\mu = \varepsilon$. Thus by Lemma 4.2.4, the partial type $\left\{|\phi(x; b_{2i}) - \phi(x; b_{2i+1})| \geq \frac{\varepsilon}{2} : i < \omega\right\}$ is consistent. This contradicts NIP. $\square$

4.2.1 (Borel) Definable Measures and the Morley Product

Lemma 4.2.7. *Let $\mu \in \mathfrak{M}_x(\mathcal{U})$ be A-(Borel) definable, and let $A \subseteq B \subseteq \mathcal{U}$. Then μ is A-(Borel) definable if and only if μ is B-(Borel) definable. In particular, if either holds, μ is C-(Borel) definable whenever μ is C-invariant.*

Proof. The proof is essentially the same as the version for discrete logic (see [Gan20, Proposition 2.22] [CGH23a, Corollary 2.2]).

The map $\pi_{B,A} : S_x(B) \to S_x(A)$ given by $\pi_{B,A}(p) = p|_A$ is continuous, surjective, and closed[BBH08, Prop. 8.11], and $F^\phi_{\mu,B} = F^\phi_{\mu,A} \circ \pi_{B,A}$. These properties of $\pi_{B,A}$ imply that $F^\phi_{\mu,A}$ is continuous/Borel if and only if $F^\phi_{\mu,A} \circ \pi_{B,A}$ is. Most of these implications are straightforward, but it is nontrivial that $F^\phi_{\mu,B}$ being Borel implies $F^\phi_{\mu,A}$ is as well. $\square$

Borel definable measures are important largely because they are the measures for which we can define the Morley product of Keisler measures.

Definition 4.2.8. Given an A-Borel definable measure μ and a global measure ν, let $f_{\mu\otimes\nu}$ be the Keisler functional defined by

$$f_{\mu\otimes\nu}(\phi(x; y)) = \int_{S_y(A')} F^\phi_{\mu,A'}(y)\, d\nu|_{A'}$$

where $\phi(x; y)$ is a formula, and A' contains A and the parameters of ϕ. Let $\mu \otimes \nu$ be the corresponding Keisler measure, so that

$$\int_{S_{xy}(\mathcal{U})} \phi(x; y) \, d(\mu \otimes \nu) = \int_{S_y(A')} F^{\phi}_{\mu,A'} \, d\nu|_{A'}$$

for all formulas $\phi(x; y)$ with parameters in $A' \supset A$.

First we check that this definition does not depend on the choice of A'. It is enough to see that if A' is enlarged to $B \supset A'$, that the value will not change. In this case, if $\pi : S_y(B) \to S_y(A')$ is the projection map, then it is easy to see that $\nu|_{A'}$ is equal to the pushforward measure $\pi_* \nu|_B$. Also, $F^{\phi}_{\mu,A'} = F^{\phi}_{\mu,A'} \circ \pi$. Thus

$$\int_{S_y(A')} F^{\phi}_{\mu,A'}(y) \, d\nu|_{A'} = \int_{S_y(B)} F^{\phi}_{\mu,A'}(y) \circ \pi \, d\nu|_B = \int_{S_y(B)} F^{\phi}_{\mu,B}(y), d\nu|_B.$$

This indeed defines a valid Keisler functional, as it is clearly linear and

$$f_{\mu\otimes\nu}(1) = \int_{S_y(A)} F^{1}_{\mu,A} \, d\nu|_A = 1.$$

It is also easy to see that if μ is A-Borel definable and ν is A-invariant, then $\mu \otimes \nu$ is also A-invariant. Also, we see that for any A such that μ is A-Borel definable and the parameters of $\phi(x; y)$ are contained in A, the value of $\int_{S_y(\mathcal{U})} \phi(x; y) \, d(\mu \otimes \nu)$ depends only on $\nu|_A$.

Lemma 4.2.9 (Generalizing [CG20, Prop. 2.6])**.** *If $\mu \in \mathfrak{M}_x(\mathcal{U}), \nu \in \mathfrak{M}_y(\mathcal{U}), \lambda \in \mathfrak{M}_z(\mathcal{U})$ are M-definable measures, then $\mu \otimes \nu$ is M-definable, and $(\mu \otimes \nu) \otimes \lambda = \mu \otimes (\nu \otimes \lambda)$.*

Proof. First we will show that $\mu \otimes \nu$ is definable by showing that for all formulas $\phi(x, y; z) \in \mathcal{L}(M)$, the function $F^{\phi(x,y;z)}_{\mu\otimes\nu} : S_z(M) \to [0, 1]$ is continuous. We can see that

$$F^{\phi(x,y;z)}_{\mu\otimes\nu} = \int_{S_{xy}(M)} \phi(x, y; z) \, d(\mu \otimes \nu) = \int_{S_y(M)} \left(\int_{S_x(M)} \phi(x, y; z) \, d\mu \right) d\nu.$$

As μ is definable, the function $F^{\phi(x;y,z)}_{\mu,M} = \int_{S_x(M)} \phi(x;y,z)\,d\mu$ is continuous, and is thus a definable predicate on (y,z). Thus as ν is definable, $\int_{S_y(M)} \left(\int_{S_x(M)} \phi(x,y;z)\,d\mu\right) d\nu$ is continuous as desired.

Now to verify associativity, it is enough to show that for all formulas $\phi(x,y,z) \in \mathcal{L}(M)$, $\int \phi(x,y,z)\,d((\mu\otimes\nu)\otimes\lambda) = \int \phi(x,y,z)\,d(\mu\otimes(\nu\otimes\lambda))$. We can see this in the simple-looking calculation

$$\begin{aligned}
\int_{S_{xyz}(M)} \phi(x,y,z)\,d((\mu\otimes\nu)\otimes\lambda) &= \int_{S_z(M)} \left(\int_{S_{xy}(M)} \phi(x,y,z)\,d(\mu\otimes\nu)\right) d\lambda \\
&= \int_{S_z(M)} \left(\int_{S_y(M)} \left(\int_{S_x(M)} \phi(x,y,z)\,d\mu\right) d\nu\right) d\lambda \\
&= \int_{S_{yz}(M)} \left(\int_{S_x(M)} \phi(x,y,z)\,d\mu\right) d(\nu\otimes\lambda) \\
&= \int \phi(x,y,z)\,d(\mu\otimes(\nu\otimes\lambda)).
\end{aligned}$$

These equations are justified by the definition of the Morley product, together with the fact that all the functions being integrated are continuous, and thus are definable predicates. This continuity follows from the definability of $\mu,\nu,\mu\otimes\nu$. □

It will also be useful to generalize some of the behavior of continuous functions with respect to Morley products of definable measures to characteristic functions of open sets.

Lemma 4.2.10 (Generalizing [CGH23a, Prop. 2.17]). *If $\mu \in \mathfrak{M}_x(\mathcal{U})$ is A-definable, then for any open set $U \subseteq S_{xy}(A)$, the function $\int_{S_x(A)} \chi_U(x,y)\,d\mu$ is itself Borel, and for any $\nu \in \mathfrak{M}_y(\mathcal{U})$,*

$$\int_{S_y(A)} \int_{S_x(A)} \chi_U(x,y)\,d\mu\,d\nu = \int_{S_{xy}(A)} \chi_U(x,y)\,d\mu\otimes\nu = (\mu\otimes\nu)(U).$$

Proof. Fix μ and U. Let $\mathcal{F}$ be the set of all continuous functions $f : S_{xy}(A) \to [0,1]$ such that $f \leq \chi_U$ pointwise (in other words, the support of f is contained in U). The set $\mathcal{F}$ is a directed partial order (with pointwise $\leq$), and thus the function $f \mapsto \int_{S_x(A)} f(x,y)\,d\mu$ with domain $\mathcal{F}$ is an increasing net of continuous functions. We can show that the pointwise limit of this net is $\int_{S_x(A)} \chi_U(x,y)\,d\mu$. As for all functions $f \in \mathcal{F}$, $\int_{S_x(A)} f(x,y)\,d\mu \leq \int_{S_x(A)} \chi_U(x,y)\,d\mu$, and the net is increasing, it suffices to show that for each $q \in S_y(A)$ with $b \vDash q$ and each $\varepsilon > 0$, there is some $f \in \mathcal{F}$ with $f(q) \geq \int_{S_x(A)} \chi_U(x,b)\,d\mu - \varepsilon$. Let $C \subseteq S_x(A)$ be a closed subset of the open fiber $U_b = \{p \in S_x(A) : (a,b) \in U \text{ for } a \vDash p\}$ with $\mu(C) \geq \mu(U_b) - \varepsilon$. Let $C' \subseteq S_{xy}(A)$ be the closed set $\{\mathrm{tp}(a,b/A) : \mathrm{tp}(a/A) \in C\}$. By Urysohn's lemma, there is a continuous function f with support contained in U with value 1 on all of C'. Thus $f \leq \chi_U(x,y)$, and

$$\int_{S_x(A)} f(x,b)\,d\mu \geq \mu(C) \geq \mu(U_b) - \varepsilon = \int_{S_x(A)} \chi_U(x,y)\,d\mu - \varepsilon,$$

so f is the function we desired.

By the monotone convergence theorem for nets ([RS80, Theorem IV.15]), the pointwise limit of an increasing net of uniformly bounded continuous functions is Borel, and its integral relative to a regular Borel measure such as ν is the limit of the integrals of the functions in the net. Thus $\int_{S_x(A)} \chi_U(x,y)\,d\mu$ is Borel, and

$$\int_{S_y(A)} \int_{S_x(A)} \chi_U(x,y)\,d\mu\,d\nu = \lim_{f\in\mathcal{F}} \int_{S_y(A)} \int_{S_x(A)} f(x,y)\,d\mu\,d\nu.$$

By the definition of the Morley product,

$$\lim_{f\in\mathcal{F}} \int_{S_y(A)} \int_{S_x(A)} f(x,y)\,d\mu\,d\nu = \lim_{f\in\mathcal{F}} \int_{S_{xy}(A)} f(x,y)\,d\mu \otimes \nu,$$

and once again using Urysohn's lemma, it is straightforward to find $f \in \mathcal{F}$ with

$$\int_{S_{xy}(A)} f(x,y)\,d\mu \otimes \nu \geq (\mu \otimes \nu)(U) - \varepsilon$$

for each ε, so this limit is $(\mu \otimes \nu)(U)$. $\square$

4.2.2 Approximately Realizable Measures

We provide another characterization of approximately realized measures, which justifies the name:

Lemma 4.2.11. *Let $\mu \in \mathfrak{M}_x(\mathcal{U})$ be a global measure.*

Then μ is approximately realized in A if and only if the following holds:

For every predicate $\phi(x)$ with parameters in $\mathcal{U}$, if $\phi(a) = 0$ for all $a \in A^x$, then $\int_{S_x(\mathcal{U})} \phi(x)\, d\mu = 0$.

Proof. First, we show that this holds for all approximately realized measures.

Assume $\phi(x)$ is a predicate such that for all $a \in A^x$, $\phi(a) = 0$.

Let ν be a convex combination of Dirac measures at types of points in A - specifically, let $\nu = \sum_{i=1}^n \lambda_i \delta_{a_i}$, where δ_{a_i} is the Dirac measure at the type realized by $a_i \in A$, and $\lambda_i \geq 0$, $\sum_{i=1}^n \lambda n = 1$. Then $\int_{S_x(\mathcal{U})} \phi(x)\, d\nu = \sum_{i=1}^n \lambda_i \phi(a_i) = 0$.

We then recall that $\nu \mapsto \int_{S_x(\mathcal{U})} \phi(x)\, d\nu$ is continuous, so as this continuous function takes the value 0 everywhere in a set, it must take the value 0 everywhere in its closure - the set of approximately realized measures.

Now we will show that any measure with this property is approximately realized in A. Assume μ is not approximately realized in A, and we will find some predicate $\phi(x)$ such that $\phi(a) = 0$ for all $a \in A^x$, but $0 < \int_{S_x(\mathcal{U})} \phi(x)\, d\mu$.

Because μ is not approximately realized in A, μ is contained in an open set that does not contain any convex combinations of Dirac measures of types realized in A. We may assume that the open set is basic - a finite intersection of sets of the form $\{\nu : r < \int_{S_x(\mathcal{U})} \phi(x)\, d\nu < s\}$ where $r < s$ and $\phi(x)$ is a formula with parameters. By potentially replacing $\phi(x)$ with $1 - \phi(x)$, we may assume that this set is an intersection of sets of the form $\{\nu : r <$

$\int_{S_x(\mathcal{U})} \phi(x)\, d\nu\}$, and by replacing $\phi(x)$ with $\phi(x) \dot{-} r$, we can replace these with sets of the form $\{\nu : 0 < \int_{S_x(\mathcal{U})} \phi(x)\, d\nu\}$. Thus assume there are formulas $\phi_1, \dots, \phi_n$ such that for each i, $0 < \int_{S_x(\mathcal{U})} \phi_i(x)\, d\mu$, but for each convex combination ν of Dirac measures at types realized in A, $\int_{S_x(\mathcal{U})} \phi_i(x)\, d\nu = 0$ for some i. We wish to show that for some i, $\phi_i(a) = 0$ for all $a \in A^x$. If not, then for each i, let $a_i \in A^x$ be such that $\phi_i(a_i) > 0$. Then let $\nu = \frac{1}{n} \sum_{i=1}^n \delta_{a_i}$, and note that $\int_{S_x(\mathcal{U})} \phi_i(x)\, d\nu = \frac{1}{n} \sum_{i=1}^n \phi_i(a_i) > 0$, a contradiction.

□

We also note that all approximately realized measures are invariant:

Lemma 4.2.12. *The set of A-invariant measures is closed in $\mathfrak{M}_x(\mathcal{U})$, and all measures approximately realized in A are A-invariant.*

Proof. The set of A-invariant measures is

$$\bigcap_{\phi(x;y),\, a \equiv_A b} \left\{\mu : \int_{S_x(\mathcal{U})} \phi(x;a)\, d\mu = \int_{S_x(\mathcal{U})} \phi(x;a)\, d\mu\right\}$$

. As for each predicate $\phi(x;y)$ and each $a \in \mathcal{U}$, the function $\mu \mapsto \int_{S_x(\mathcal{U})} \phi(x;a)\, d\mu$ is continuous, each set in this intersection is closed, so the intersection itself is.

The set of approximately realized measures is the topological closure of the convex hull of the Dirac measures at types of points in A. It is clear that the type of a point in A is A-invariant, and that a convex combination of A-invariant measures is A-invariant. This is thus the closure of a set of A-invariant measures, which must then be contained in the closed set of A-invariant measures. □

The choice of model does not matter for defining approximately realized measures, as long as they are invariant.

Lemma 4.2.13. *Let μ be a measure approximately realized/dfs in A, and invariant over a small model M. Then μ is approximately realized/dfs in M.*

Proof. The result for dfs will follow from the result for approximate realization, as it holds for definability by Lemma 4.2.7. Now assume μ is approximately realized in A and M-invariant. Let N be a measure extending $A \cup M$. Then clearly μ is approximately realized in N. $\square$

Approximately realized measures are also closed under Morley products:

Lemma 4.2.14. *Let $\mu \in \mathfrak{M}_x(\mathcal{U})$ and $\nu \in \mathfrak{M}_y(\mathcal{U})$ be approximately realized in A. Then $\mu \otimes \nu$ is as well.*

Proof. Let us use the characterization from Lemma 4.2.11. It suffices to show that for every predicate $\phi(x;y)$ with parameters from $\mathcal{U}$, if $\int_{S_{xy}(\mathcal{U})} \phi(x;y)\, d\mu \otimes \nu > 0$, then $\phi(a;b) > 0$ for some $ab \in A^{xy}$.

If $\int_{S_{xy}(\mathcal{U})} \phi(x;y)\, d\mu \otimes \nu = \int_{S_y(\mathcal{U})} F_\mu^\phi(y)\, d\nu > 0$, then as ν is approximately realized, there is some $b \in A^y$ such that $F_\mu^\phi(b) > 0$. Thus $\int_{S_x(\mathcal{U})} \phi(x;b)\, d\mu > 0$, so as μ is approximately realized, there is some $a \in A^x$ with $\phi(a;b) > 0$. $\square$

4.2.3 Extensions and Orthogonality

In order to understand extensions of Keisler measures to larger sets of parameters, let us focus on the positive linear functional perspective, and apply a specialized version of Hahn-Banach.

First, we observe that for any model M, the space $\mathcal{C}(S_x(M), \mathbb{R})$ is an ordered vector space, with positive cone C consisting of all $f \in \mathcal{C}(S_x(M), \mathbb{R})$ such that $f(x) \geq 0$ always. We can view it as an ordered topological vector space by giving it the ℓ_∞-norm. Note that this means the interior points of the positive cone C are exactly the functions f such that $\inf_{p \in S_x(M)} f(p) > 0$. As the functions in this space have a compact domain $S_x(M)$, these are all the strictly positive functions.

By the following fact, it is clear that every positive linear functional on $\mathcal{C}(S_x(M), \mathbb{R})$, and thus every Keisler functional, is continuous with respect to the ℓ_∞-norm.

Fact 4.2.15 ([SW99, Theorem 5.5]). *Let E be an ordered topological vector space with positive cone C, such that C has nonempty interior. Then every positive linear form on E is continuous.*

We can combine that conclusion with this fact, guaranteeing continuous positive extensions of continuous positive linear functionals defined on subspaces.

Fact 4.2.16 ([SW99, Corollary 2 of Theorem 5.4]). *Let E be an ordered topological vector space with positive cone C, and suppose that V is a vector subspace of E such that $C \cap V$ contains an interior point of C. Then every continuous, positive linear form on V can be extended to E, preserving continuity and positivity.*

Corollary 4.2.17. *Let V be a vector subspace of $\mathcal{C}(S_x(M), \mathbb{R})$, containing the constant functions, and $f : V \to \mathbb{R}$ a positive linear functional with $f(1) = 1$. Then f can be extended to a Keisler functional in $\mathfrak{M}_x(M)$.*

Proof. We see that V contains an interior point of C, namely the constant function 1. Thus V, as a subspace of $\mathcal{C}(S_x(M), \mathbb{R})$, is an ordered topological vector space whose positive cone has nonempty interior, so by Fact 4.2.15, f is continuous. Thus also by Fact 4.2.16, f has an extension to all of $\mathcal{C}(S_x(M), \mathbb{R})$, which is positive, and is thus a Keisler functional, as $f(1) = 1$. □

Lemma 4.2.18. *Let V be a vector subspace of $\mathcal{C}(S_x(M), \mathbb{R})$, containing the constant functions, and $f : V \to \mathbb{R}$ a positive linear functional with $f(1) = 1$. Let $\phi(x) \in \mathcal{C}(S_x(M), \mathbb{R})$. Then the set of possible values $\hat{f}(\phi)$ where $\hat{f}$ is a Keisler functional extending f is exactly the interval*

$$\left[\sup_{\psi \in V : \psi \leq \phi} f(\psi), \inf_{\psi \in V : \psi \geq \phi} f(\psi)\right].$$

Proof. First, we note that if $\hat{f}$ is a Keisler functional extension of f, then $\hat{f}(\phi)$ must be in that interval, because for any $\psi \in V : \psi \leq \phi$, we have $f(\psi) = \hat{f}(\psi) \leq \hat{f}(\phi)$, and similarly for any $\psi \in V : \psi \geq \phi$, we have $f(\psi) = \hat{f}(\psi) \geq \hat{f}(\phi)$.

Now fix $r \in \left[\sup_{\psi \in V:\psi \leq \phi} f(\psi), \inf_{\psi \in V:\psi \geq \phi} f(\psi)\right]$. Then we define f_ϕ on the vector subspace $V + \mathbb{R}\phi$ by $f_\phi(\theta + a\phi) = f(\theta) + ar$ for all $\theta \in V$ and $a \in \mathbb{R}$. This is clearly an extension of f if $\phi \notin V$, and if $\phi \in V$, then our conditions already guarantee $r = f(\phi)$, so $f_\phi = f$. It suffices to show that f_ϕ is positive, as Corollary 4.2.17 will then guarantee that $\hat{f}$ extends to a Keisler functional, which has $\hat{f}(\phi) = f_\phi(\phi) = r$.

To show that f_ϕ is positive, consider $\theta \in V$ and $a \in \mathbb{R}$ such that $\theta + a\phi$ is always nonnegative. Then we must show that $f_\phi(\theta + a\phi) = f(\theta) + ar$ is always nonnegative. If $a = 0$, this is guaranteed by the positivity of f. If a is positive, we need to check that $r \geq -a^{-1}f(\theta)$. This is true because for each x, $\theta(x) + a\phi(x) \geq 0$, so $-a^{-1}\theta(x) \leq \phi(x)$. Thus $f(-a^{-1}\theta) \leq \sup_{\psi \in V:\psi \leq \phi} f(\psi) \leq r$. Similarly, if a is negative, then $f(-a^{-1}\theta) \geq \inf_{\psi \in V:\psi \geq \phi} f(\psi) \geq r$, so $f(\theta) + ar \geq 0$. □

Our first application of Lemma 4.2.18 is extending Keisler measures to larger parameter sets.

Corollary 4.2.19. *Let $M \subseteq N$ be models, let $\mu \in \mathfrak{M}_x(M)$, and let $\phi(x)$ be an N-definable predicate. Then for $r \in [0,1]$, there is a Keisler measure $\nu \in \mathfrak{M}_x(N)$ extending μ such that $\phi(x) = r$ if and only if*

$$\sup_{\psi:\psi \leq \phi} f(\psi) \leq r \leq \inf_{\psi:\psi \geq \phi} f(\psi)$$

where the sup *and* inf *are over M-definable predicates $\psi(x)$.*

Proof. This follows from applying Lemma 4.2.18 to the image of $\mathcal{C}(S_x(M), \mathbb{R})$ in $\mathcal{C}(S_x(N), \mathbb{R})$. □

Our second will be an application to product measures, which will require generalizing a few more basic definitions to continuous logic.

Definition 4.2.20. Let $x_1, \dots, x_n$ be variable tuples, with $x = (x_1, \dots, x_n)$. If $\mu_i \in \mathfrak{M}_{x_i}(M)$ is a family of Keisler measures, then we use the notation $\mu_1 \times \dots \times \mu_n$ to denote the partial

Keisler "measure" (actually a functional) defined by

$$\int_{S_x(M)} \prod_{i=1}^{n} \psi_i(x_i; b)\, d\mu_1 \times \cdots \times \mu_n = \prod_{i=1}^{n} \int_{S_{x_i}(M)} \psi_i(x_i; b)\, d\mu_i$$

whenever $\psi_i(x_i; y)$ are formulae (or definable predicates) and $b \in M^y$.

A measure $\mu \in \mathfrak{M}_x(M)$ is a *product measure* of $\mu_i \in \mathfrak{M}_{x_i}(M)$ when it extends $\mu_1 \times \cdots \times \mu_n$. The measures $\mu_i \in \mathfrak{M}_{x_i}(M)$ are *weakly orthogonal* when they have a unique product measure. If $M = \mathcal{U}$, we say that they are *orthogonal*.

Note that unlike with types, if $\mu \in \mathfrak{M}_{xy}(M)$ is a measure, and $\mu|_x, \mu|_y$ are the restrictions to the appropriate variables, then μ need not be a product measure of $\mu|_x$ and $\mu|_y$.

Corollary 4.2.21. *Let $x_1, \ldots, x_n$ be variable tuples, and let $\mu_i \in \mathfrak{M}_{x_i}(M)$ for each i. The measures μ_i are weakly orthogonal if and only if for every M-definable predicate $\phi(x_1, \ldots, x_n)$ and every $\varepsilon > 0$, there exist M-definable predicates $\psi^-(x_1, \ldots, x_n), \psi^+(x_1, \ldots, x_n)$, where $\psi^\pm(x_1, \ldots, x_n)$ are each of the form $\sum_{j=1}^{m} \prod_{i=1}^{n} \theta_{ij}^{\pm}(x_i)$, such that*

- *For all $(x_1, \ldots, x_n)$, $\psi^-(x_1, \ldots, x_n) \leq \phi(x_1, \ldots, x_n) \leq \psi^+(x_1, \ldots, x_n)$.*
- *For any product measure ω of $\mu_1, \ldots, \mu_n$, $\int_{S_{x_1 \ldots x_n}(M)} (\psi^+ - \psi^-)\, d\omega \leq \varepsilon$.*

Proof. First, we assume the measures are weakly orthogonal. Let $x = x_1 \ldots x_n$. Then consider the vector subspace V of $\mathcal{C}(S_x(M), \mathbb{R})$ spanned by products $\prod_{i=1}^{n} \psi_i(x_i)$ where each $\psi_i(x_i) \in \mathcal{C}(S_{x_i}(M), \mathbb{R})$. Define a positive linear functional f on V so that

$$f\left(\prod_{i=1}^{n} \psi_i(x_i)\right) = \prod_{i=1}^{n} \int_{S_{x_i}(M)} \psi_i(x_i)\, d\mu_i.$$

Any positive linear extension of this to $\mathcal{C}(S_x(M), \mathbb{R})$ gives rise to a product measure in $\mathfrak{M}_x(M)$, but we know that this is unique.

Fix $\phi(x)$ and $\varepsilon > 0$. By Lemma 4.2.18 and the uniqueness of the extension of f, we know that $\left[\sup_{\psi \in V : \psi \leq \phi} f(\psi) = \inf_{\psi \in V : \psi \geq \phi} f(\psi)\right]$, so choose $\psi^-(x), \psi^+(x) \in V$ such that $\psi^-(x) \leq \phi(x) \leq \psi^+(x)$ and $f(\psi^+ - \psi^-) \leq \varepsilon$. Then $\int_{S_x(M)} (\psi^+ - \psi^-)\, d\omega = f(\psi^+ - \psi^-) \leq \varepsilon$.

Now assume that the result holds, and we will show the measures are weakly orthogonal. If $\omega_1, \omega_2 \in \mathfrak{M}_x(M)$ are measures extending $\mu_1 \times \cdots \times \mu_n$, and $\phi(x_1, \ldots, x_n)$ is a definable predicate, then for every $\varepsilon > 0$, we may find $\psi^\pm$ as above. We see that $\int_{S_{x_1 \ldots x_n}(M)} \psi^\pm\, d\omega_1 = \int_{S_{x_1 \ldots x_n}(M)} \psi^\pm\, d\omega_2$, so the integrals $\int_{S_{x_1 \ldots x_n}(M)} \phi\, d\omega_j$ for $j = 1, 2$ must lie in the interval

$$\left[\int_{S_{x_1 \ldots x_n}(M)} \psi^-\, d\omega_1, \int_{S_{x_1 \ldots x_n}(M)} \psi^+\, d\omega_1\right]$$

of width at most ε. Thus $\int_{S_{x_1 \ldots x_n}(M)} \phi\, d\omega_1 = \int_{S_{x_1 \ldots x_n}(M)} \phi\, d\omega_2$, and the measures are weakly orthogonal. $\square$

We will show that weak orthogonality is preserved under localization to a positive-measure Borel set or positive-integral function. Let $\mu \in \mathfrak{M}_x(M)$ be a Keisler measure, and let $\phi : S_x(M) \to [0, 1]$ be Borel with $\int_{S_x(M)} \phi(x)\, d\mu > 0$. Then the localization $\mu_\phi \in \mathfrak{M}_x(M)$ is the measure given by

$$\int_{S_x(M)} \psi(x)\, d\mu_\phi = \frac{\int_{S_x(M)} \phi(x)\psi(x)\, d\mu}{\int_{S_x(M)} \phi(x)\, d\mu}.$$

If ϕ is the characteristic function of a Borel set $X \subseteq S_x(M)$, we may also call the localization μ_X.

Lemma 4.2.22. *Let $\mu_i \in \mathfrak{M}_{x_i}(M)$ for $1 \leq i \leq n$ be weakly orthogonal, and $\theta_i : S_{x_i}(M) \to [0, 1]$ be Borel with $\int_{S_{x_i}(M)} \theta_i(x_i)\, d\mu_i > 0$. Then the measures $(\mu_i)_{\theta_i}$ are weakly orthogonal also.*

Proof. For readability, let $A_i = \int_{S_{x_i}(M)} \theta_i(x_i)\, d\mu_i$ for $1 \leq i \leq n$, and let $A = \prod_{i=1}^n A_i$. Let ω be the unique extension of $\mu_1 \times \cdots \times \mu_n$. If $\theta(x_1, \ldots, x_n) = \prod_{i=1}^n \theta_i(x_i)$, then ω_θ extends

$(\mu_1)_{\theta_1} \times \cdots \times (\mu_n)_{\theta_n}$. Suppose that ν also extends $(\mu_1)_{\theta_1} \times \cdots \times (\mu_n)_{\theta_n}$. Then we define a measure ν' on any M-definable predicate $\phi(x_1, \ldots, x_n)$ by

$$\begin{aligned}&\int_{S_{x_1 \ldots x_n}(M)} \phi(x_1, \ldots, x_n)\, d\nu' \\ =&A \int_{S_{x_1 \ldots x_n}(M)} \phi(x_1, \ldots, x_n)\, d\nu + \int_{S_{x_1 \ldots x_n}(M)} \phi(x_1, \ldots, x_n)(1 - \theta(x_1, \ldots, x_n))\, d\omega.\end{aligned}$$

If for $1 \leq i \leq n$, $\psi_i(x_i)$ is an M-definable predicate, then

$$\begin{aligned}&\int_{S_{x_1 \ldots x_n}(M)} \prod_{i=1}^{n} \psi_i(x_i)\, d\nu' \\ =&A \int_{S_{x_1 \ldots x_n}(M)} \prod_{i=1}^{n} \psi_i(x_i)\, d\nu' + \int_{S_{x_1 \ldots x_n}(M)} \left(\prod_{i=1}^{n} \psi_i(x_i) - \prod_{i=1}^{n} \psi_i(x_i)\theta_i(x_i) \right) d\omega \\ =&A \prod_{i=1}^{n} \frac{1}{A_i} \int_{S_{x_i}(M)} \psi_i(x_i)\theta_i(x_i)\, d\nu + \prod_{i=1}^{n} \int_{S_{x_i}(M)} \psi_i(x_i)\, d\mu_i - \prod_{i=1}^{n} \int_{S_{x_i}(M)} \psi_i(x_i)\theta_i(x_i)\, d\mu_i \\ =&\prod_{i=1}^{n} \psi_i(x_i)\, d\mu_i,\end{aligned}$$

so ν' extends $\mu_1 \times \cdots \times \mu_n$, and thus equals ω.

Thus for any ϕ,

$$A \int_{S_{x_1 \ldots x_n}(M)} \phi(x_1, \ldots, x_n)\, d\nu = \int_{S_{x_1 \ldots x_n}(M)} \phi(x_1, \ldots, x_n)\theta(x_1, \ldots, x_n)\, d\omega,$$

so $\nu = \omega_\theta$, showing the uniqueness of extensions of $(\mu_1)_{\theta_1} \times \cdots \times (\mu_n)_{\theta_n}$ and weak orthogonality of the localizations. $\square$

4.2.4 Smooth Measures

In this subsection, we will update to the continuous setting several results about smooth measures that do not require NIP, and then the important result that in an NIP theory, all measures over models have smooth extensions.

The most important characterization of smooth measures is the following lemma, analogous to [Sim15, Lemma 7.8]. This result was already known by James Hanson, but we provide a proof here for completeness.

Lemma 4.2.23. *Let $\mu \in \mathfrak{M}_x(\mathcal{U})$ be a global measure. Then μ is smooth over M if and only if for every definable predicate $\phi(x; y)$ with parameters in M, and $\varepsilon > 0$, there are open conditions $U_i(y)$ and definable predicates $\psi_i^+(x), \psi_i^-(x)$ with parameters in M for $i = 1, \ldots, n$ such that*

- *$U_1(y), \ldots, U_n(y)$ cover $S_y(M)$*
- *For all $1 \leq i \leq n$, if $\vDash U_i(b)$, then $\forall x, \psi_i^-(x) \leq \phi(x; b) \leq \psi_i^+(x)$.*
- *For all $1 \leq i \leq n$, $\int_{S_x(\mathcal{U})} \psi_i^+(x) - \psi_i^-(x)\, d\mu < \varepsilon$.*

Proof. Assume that μ satisfies these requirements, and fix an M-definable predicate $\phi(x; y)$. We will show that for all $b \in \mathcal{U}^y$, the value of $\int_{S_x(\mathcal{U})} \phi(x; b)\, d\nu = \int_{S_x(\mathcal{U})} \phi(x; b)\, d\mu$ is determined for all global measures ν extending $\mu|_M$. Specifically, for every $\varepsilon > 0$, we show that $|\int_{S_x(\mathcal{U})} \phi(x; b)\, d\nu - \int_{S_x(\mathcal{U})} \phi(x; b)\, d\mu| \leq \varepsilon$. There must be some i such that $\vDash U_i(b)$. Thus $\psi_i^-(x) \leq \phi(x; b) \leq \psi_i^+(x)$, so

$$\int_{S_x(M)} \psi_i^-(x)\, d\mu|_M \leq \int_{S_x(\mathcal{U})} \phi(x; b)\, d\nu \leq \int_{S_x(M)} \psi_i^+(x)\, d\mu|_M,$$

and $\int_{S_x(\mathcal{U})} \phi(x; b)\, d\mu$ lies in that same interval of width at most ε. Thus their difference is at most ε.

Now assume that μ is smooth over M. For every $b \in \mathcal{U}^y$, we will find ψ_b^-, ψ_b^+ such that $\forall x, \psi_b^-(x) \leq \phi(x; b) \leq \psi_b^+(x)$, and $\int_{S_x(\mathcal{U})} \psi_b^+(x) - \psi_b^-\, d\mu < \varepsilon$. We will then apply compactness.

By smoothness, for any global extension ν of $\mu|_M$, the integral $\int_{S_x(\mathcal{U})} \phi(x; b)\, d\nu$ has the same value, namely $\int_{S_x(\mathcal{U})} \phi(x; b)\, d\nu = \int_{S_x(\mathcal{U})} \phi(x; b)\, d\mu$. Thus by Corollary 4.2.19, we must have

$$\sup_{\psi:\psi\leq\phi}\int_{S_x(M)}\psi(x)\,d\mu|_M=\int_{S_x(\mathcal{U})}\phi(x;b)\,d\mu=\inf_{\psi:\psi\geq\phi}\int_{S_x(M)}\psi(x)\,d\mu|_M,$$

where as above, $\psi(x)$ ranges over M-definable predicates. Thus there exist $\psi_b^-(x)$ and $\psi_b^+(x)$ with $\psi_b^-(x) \leq \phi(x;b) \leq \psi_b^+(x)$ and $\int_{S_x(\mathcal{U})}\phi(x;b)\,d\mu - \int_{S_x(M)}\psi_b^-(x)\,d\mu < \frac{\varepsilon}{2}$ and $\int_{S_x(M)}\psi_b^+(x)\,d\mu - \int_{S_x(\mathcal{U})}\phi(x;b)\,d\mu < \frac{\varepsilon}{2}$. Thus $\int_{S_x(M)}\psi_b^+(x)\,d\mu - \int_{S_x(M)}\psi_b^-(x)\,d\mu < \varepsilon$ as desired.

For the compactness argument, it will be simpler to assume that $\phi(x;y)$ only takes values in the open interval $(0,1)$. We will prove the result for those ϕ first, and then make a correction for all ϕ. For any $b\in\mathcal{U}^y$, by adding a very small amount to $\psi_b^+(x)$ and $\psi_b^-(x)$, we can guarantee that at every point $\psi_b^-(x) < \phi(x;b) < \psi_b^+(x)$, while still ensuring that $\int_{S_x(M)}\psi_b^+(x)\,d\mu - \int_{S_x(M)}\psi_b^-(x)\,d\mu < \varepsilon$. By the compactness of $S_x(M)$, there is some $\delta > 0$ such that $\inf_{x\in S_x(M)}\psi_b^+(x)-\phi(x;b) > \delta$ and $\inf_{x\in S_x(M)}\phi(x;b)-\psi_b^-(x) > \delta$. Thus there is an open subset $U_b(y)$ of $S_y(M)$, containing the type of b, defined by $\inf_{x\in S_x(M)}\psi_b^+(x)-\phi(x;y) > \delta$ and $\inf_{x\in S_x(M)}\phi(x;y)-\psi_b^-(x) > \delta$, such that $\psi_b^-(x) < \phi(x;y) < \psi_b^+(x)$ at every point satisfying $U_b(y)$. Let $U_1(y),\dots,U_n(y)$ be a finite subcover of these, with $U_i(y) = U_{b_i}(y)$. Then taking $\psi_i^\pm(x) = \psi_{b_i}^\pm(x)$, we have the desired result.

Now for the correction. If $\phi(x;y)$ is any M-definable predicate, possibly taking the values 0 or 1, we apply the result to the predicate $\phi(x;y)' = \frac{1}{2}\phi(x;y)+\frac{1}{4}$, whose range is bounded to $\left[\frac{1}{4},\frac{3}{4}\right]$, finding open conditions $U_i(y)$ and definable predicates $\psi_i^{+'}(x),\psi_i^{-'}(x)$ with parameters in M for $i=1,\dots,n$ such that

- $U_1(y),\dots,U_n(y)$ cover $S_y(M)$
- For all $1\leq i\leq n$, if $\vDash U_i(b)$, then $\forall x, \psi_i^{-'}(x)\leq\phi'(x;b)\leq\psi_i^{+'}(x)$.
- For all $1\leq i\leq n$, $\int_{S_x(\mathcal{U})}\psi_i^{+'}(x)-\psi_i^{-'}\,d\mu < \frac{\varepsilon}{2}$.

By taking $\psi_i^\pm(x) = \min(\max(2\psi_i^{\pm'}(x)-\frac{1}{2},0),1)$, we find that for all $1\leq i\leq n$, if $\vDash U_i(b)$, then $\forall x, \psi_i^-(x)\leq\phi(x;b)\leq\psi_i^+(x)$, and $\int_{S_x(\mathcal{U})}\psi_i^+(x)-\psi_i^-\,d\mu<\varepsilon$.

$\square$

In the analogous characterization from discrete logic, the conditions $U_1(y), \dots, U_n(y)$ can be chosen to be disjoint. By taking Boolean combinations, we can enforce disjointness and end up with a Borel partition. In other applications, it will be more convenient to replace the open cover with a partition of unity using [And23b, Fact 3.3.5]. Using that fact, we can choose the open conditions $U_i(y)$ to each be the support of some $u_i(y)$, with $\forall y, u_1(y) + \cdots + u_n(y) = 1$.

This allows us to characterize smooth measures in terms of weak orthogonality.

Lemma 4.2.24. *Let $\mu \in \mathfrak{M}_x(\mathcal{U})$ be a global measure, and let $M \subset \mathcal{U}$ be small. The following are equivalent:*

- *μ is smooth over M*
- *$\mu|_M$ is weakly orthogonal to all types $p(y) \in S_y(M)$*
- *$\mu|_M$ is weakly orthogonal to all measures $\nu(y) \in \mathfrak{M}_y(M)$.*

Proof. Clearly if $\mu|_M$ is weakly orthogonal to all measures over M, then it is weakly orthogonal to all types over M.

Assume that $\mu|_M$ is weakly orthogonal to all types $p(y) \in S_y(M)$. Then fix a formula $\phi(x; y)$, and $b \in \mathcal{U}^y$. As $\mu|_M$ is weakly orthogonal to $\mathrm{tp}(b/M)$, there is a unique measure $\omega \in \mathfrak{M}_{xy}(M)$ extending $\mu|_M \times \mathrm{tp}(b/M)$, so for any μ' extending $\mu|_M$ and $c \vDash \mathrm{tp}(b/M)$, the measure $\lambda \in \mathfrak{M}_x(M)$ given by $\int_{S_{xy}(M)} \phi(x; y)\, d\lambda = \int_{S_x(\mathcal{U})} \phi(x; b)\, d\mu'$ equals ω, so $\int_{S_x(\mathcal{U})} \phi(x; b)\, d\mu'$ is uniquely determined, from which we can conclude that μ is smooth over M.

Now assume that μ is smooth over M, and let $\nu \in \mathfrak{M}_y(M)$ be another measure. We will show that $\mu|_M$ is weakly orthogonal to ν. Let $\lambda \in \mathfrak{M}_{xy}(M)$ extend $\mu|_M \times \nu$ and fix an M-definable predicate $\phi(x; y)$ and $\varepsilon > 0$. Let $u_i(y), \psi_i^+(x), \psi_i^-(x)$ for $1 \leq i \leq n$ be definable predicates as given by Lemma 4.2.23, with the u_is forming a partition of unity as in the remarks following that lemma. Then we see that $\forall x, \forall y, \sum_{i=1}^n \psi_i^-(x)u_i(y) \leq \phi(x;y) \leq \sum_{i=1}^n \psi_i^+(x)u_i(y)$. These bounds together with the separated amalgam property tell us that

$$\int_{S_{xy}(\mathcal{U})} \phi(x;y)\,d\lambda \leq \sum_{i=1}^n \int_{S_{xy}(\mathcal{U})} \psi_i^+(x)u_i(y)\,d\lambda = \sum_{i=1}^n \int_{S_x(\mathcal{U})} \psi_i^+(x)\,d\mu|_M \int_{S_y(\mathcal{U})} u_i(y)\,d\nu,$$

and that

$$\sum_{i=1}^n \int_{S_x(\mathcal{U})} \psi_i^-(x)\,d\mu|_M \int_{S_y(\mathcal{U})} u_i(y)\,d\nu \leq \int_{S_{xy}(\mathcal{U})} \phi(x;y)\,d\lambda.$$

This places $\int_{S_{xy}(\mathcal{U})} \phi(x;y)\,d\lambda$ in an interval not depending on λ, of width

$$\sum_{i=1}^n \int_{S_x(\mathcal{U})} \psi_i^+(x)\,d\mu|_M \int_{S_y(\mathcal{U})} u_i(y)\,d\nu - \sum_{i=1}^n \int_{S_x(\mathcal{U})} \psi_i^-(x)\,d\mu|_M \int_{S_y(\mathcal{U})} u_i(y)\,d\nu \leq \varepsilon,$$

which determines $\int_{S_{xy}(\mathcal{U})} \phi(x;y)\,d\lambda$ uniquely as ε was arbitrary. □

Corollary 4.2.25 (Generalizes [HPS13, Corollary 2.5])**.** *Let $\mu \in \mathfrak{M}_x(\mathcal{U})$ be smooth over a model M, and let $\nu \in \mathfrak{M}_y(\mathcal{U})$ be Borel-definable over M. Then $\mu \otimes \nu = \nu \otimes \mu$.*

Proof. Both $\mu \otimes \nu$ and $\nu \otimes \mu$ are separated amalgams of μ and ν, thus by Lemma 4.2.24, they are equal. □

Smooth measures are also preserved under localization to a positive-measure Borel set or positive-integral function.

Corollary 4.2.26. *Let $\mu \in \mathfrak{M}_x(\mathcal{U})$ be a smooth measure and let $\theta : S_x(M) \to [0,1]$ be Borel with $\int_{S_x(M)} \theta(x)\,d\mu > 0$. Then the measures μ_θ is weakly orthogonal also.*

Proof. By Lemma 4.2.24, μ is smooth if and only if it is weakly orthogonal to all global types, and by Lemma 4.2.22, μ_θ is also weakly orthogonal to all global types. □

Lemma 4.2.27 (Generalizes [Sim15, Lemma 7.17(i)])**.** *Let $\mu \in \mathfrak{M}_x(\mathcal{U})$ be a global measure smooth over M. Then μ is dfs over M.*

Proof. First we show that μ is approximately realized in M. By Lemma 4.2.11, it suffices to show that for any $\phi(x;b)$, if $\int_{S_x(\mathcal{U})} \phi(x;b)\, d\mu > 0$, then for some $a \in M^x$, $\phi(a;b) > 0$. Assume $\int_{S_x(\mathcal{U})} \phi(x;b)\, d\mu > \varepsilon > 0$. Then by Lemma 4.2.23, there are $\psi^-(x), \psi^+(x)$ such that $\forall x, \psi^-(x) \leq \phi(x;b) \leq \psi^+(x)$, and $\int_{S_x(\mathcal{U})} \psi_i^+(x) - \psi_i^-\, d\mu < \varepsilon$. Thus

$$\varepsilon < \int_{S_x(\mathcal{U})} \phi(x;b)\, d\mu \leq \int_{S_x(\mathcal{U})} \psi_i^+\, d\mu < \int_{S_x(\mathcal{U})} \psi_i^-\, d\mu + \varepsilon,$$

so $\psi^-(x)$ must take a positive value at some $a \in \mathcal{U}^x$. By elementary equivalence, it must also take a positive value at some $a' \in \mathcal{U}^x$, where we have $0 < \psi^-(a') \leq \phi(a';b)$.

Now we show that μ is definable over M. Specifically, we fix $\phi(x;y)$, and wish to show that $F^{\phi}_{\mu,A}$ is continuous, by showing that if $r < s$, the set $\{p \in S_y(A) : r < F^{\phi}_{\mu,A}(p) < s\}$ is open. Let $p \in S_y(A)$ be such that $r < F^{\phi}(\mu, A)(p) < s$, and let $b \vDash p$. Fix $0 < \varepsilon < \min(s - F^{\phi}(\mu, A)(p), F^{\phi}(\mu, A)(p) - r)$. By Lemma 4.2.23, there is an open condition $U(y)$ such that $\vDash U(b)$, and $\psi^-(x), \psi^+(x)$ such that $\int_{S_x(A)} \psi^+(x) - \psi^-(x)\, d\mu < \varepsilon$ and for all b' with $\vDash U(b')$, $\forall x, \psi^i(x) \leq \phi(x;b') \leq \psi^+(x)$. We will show that for all $q \in S_y(A)$ in the open neighborhood defined by $U(y)$, $r < F^{\phi}(\mu, A)(q) < s$.

We have $F^{\phi}(\mu, A)(p) = \int_{S_x(\mathcal{U})} \phi(x;b)\, d\mu$, so

$$\begin{aligned} r &< F^{\phi}(\mu, A)(p) - \varepsilon \\ &\leq \int_{S_x(\mathcal{U})} \psi^-(x)\, d\mu \\ &\leq F^{\phi}(\mu, A)(p) \\ &\leq \int_{S_x(\mathcal{U})} \phi^+(x)\, d\mu \\ &\leq F^{\phi}(\mu, A)(p) + \varepsilon < s. \end{aligned}$$

Now let $q \in S_y(A)$ in the open neighborhood defined by $U(y)$, and let $b' \vDash q$. Then $\vDash U(b')$, and thus

$$r < \int_{S_x(\mathcal{U})} \psi^-(x)\, d\mu \leq \int_{S_x(\mathcal{U})} \phi(x;b')\, d\mu \leq \int_{S_x(\mathcal{U})} \phi^+(x)\, d\mu < s,$$

so $F^\phi(\mu, A)(q) = \int_{S_x(\mathcal{U})} \phi(x;b')\, d\mu$ is in (r,s). $\square$

Now we note that the choice of small model is not critical when defining smoothness.

Lemma 4.2.28. *[Generalizing [Sim15, Lemma 7.17]] Let $\mu \in \mathfrak{M}_x(\mathcal{U})$ be smooth over M, and invariant over another small model N. Then μ is smooth over N also.*

Proof. We may assume that $N \preceq M$, as otherwise we may replace M with an elementary extension of N containing $M \cup N$.

Fix $\varepsilon > 0$, an N-definable predicate $\phi(x;y)$, and $p \in S_y(M)$. Then by smoothness, $\mu|_M$ and p are weakly orthogonal, so there exist definable predicates $\psi^-(x,y,z), \psi^+(x,y,z)$, where $\psi^\pm(x,y,z)$ are each of the form $\sum_{j=1}^m \theta^\pm_{1j}(x)\theta^\pm_{2j}(y)\theta^\pm_{3j}(z)$, and some $c \in M^z$, such that for all $a, b \in M^z$, $\psi^-(a,b,c) \leq \phi(a,b) \leq \psi^+(a,b,c)$ and for all b satisfying p, $\int_{S_x(M)}(\psi^+(x;b) - \psi^-(x;b))\, d\mu < \varepsilon$.

Then c satisfies $\inf_{x,y}(\psi^-(x,y,z) \dot{-} \phi(x,y)) < \varepsilon$ and $\inf_{x,y}(\phi(x,y) \dot{-} \psi^+(x,y,z)) < \varepsilon$, both open conditions in z which have parameters only in N.

As μ is definable over M and invariant over N, it is also definable over N, so it is also an open condition with parameters in N that $F^{\psi^+}_{\mu,N}(p|_N, z) - F^{\psi^-}_{\mu,N}(p|_N, z) < \varepsilon$, and this open condition also applies to c.

The conjunction of all of these open conditions is an open condition, so as such a c realizes it in M, there is some c' realizing it in N. Thus letting $\chi^-(x,y,c') = \psi^-(x,y,c') - \varepsilon$ and $\chi^+(x,y,c') = \psi^+(x,y,c') + \varepsilon$, we see that for all $a \in N^z$, $\chi^-(a,b,c') \leq \phi(a,b) \leq \chi^+(a,b,c')$, while $F^{\chi^+}_{\mu,N}(p|_N, c') - F^{\chi^-}_{\mu,N}(p|_N, c') < 3\varepsilon$, showing that $\mu|_N$ and $p|_N$ are weakly orthogonal, so by the generality of p, μ is smooth over N. $\square$

Lemma 4.2.29 (Generalizes [CG21, Corollary 1.3])**.** *Let $\mu \in \mathfrak{M}_x(\mathcal{U})$ and $\nu \in \mathfrak{M}_y(\mathcal{U})$ be smooth over a model M, Then $\mu \otimes \nu$ is smooth over M.*

Proof. It suffices to show that if $\lambda|_M = (\mu\otimes\nu)|_M$, then $\lambda = \mu\otimes\nu$. As $(\mu\otimes\nu)|_M$ is a separated amalgam of μ, ν and μ and ν are smooth, by Lemma 4.2.24 it suffices to show that λ is as well.

Let $\phi(x), \psi(y)$ be formulas (with parameters), and fix $\varepsilon > 0$. We will show that

$$\left|\int_{S_{xy}(\mathcal{U})} \phi(x)\psi(y)\, d\lambda - \int_{S_x(\mathcal{U})} \phi(x)\, d\mu \int_{S_y(\mathcal{U})} \psi(y)\, d\nu\right|$$
$$< \left(\int_{S_x(\mathcal{U})} \phi(x)\, d\mu + \int_{S_y(\mathcal{U})} \psi(y)\, d\nu\right)\varepsilon + \varepsilon^2,$$

which will show that λ is a separated amalgam, as ε is arbitrary. By Lemma 4.2.23, there are formulas $\theta^-(x), \theta^+(x), \chi^-(y), \chi^+(y)$ with parameters from M such that

- $\forall x, \theta^-(x) \leq \phi(x) \leq \theta^+(x)$
- $\forall y, \chi^-(y) \leq \psi(y) \leq \chi^+(y)$
- $\int_{S_x(\mathcal{U})} \theta^+(x) - \theta^-(x)\, d\mu < \varepsilon$
- $\int_{S_y(\mathcal{U})} \chi^+(y) - \chi^-(y)\, d\nu < \varepsilon.$

We will explicitly prove the upper bound on $\int_{S_{xy}(\mathcal{U})} \phi(x)\psi(y)\, d\lambda$, the lower bound will follow by the same logic. Using the fact that $\lambda|_M = (\mu \otimes \nu)|_M$, we see that

$$\begin{aligned}\int_{S_{xy}(\mathcal{U})} \phi(x)\psi(y)\, d\lambda &\leq \int_{S_{xy}(\mathcal{U})} \theta^+(x)\chi^+(y)\, d\lambda\\ &= \int_{S_{xy}(\mathcal{U})} \theta^+(x)\chi^+(y)\, d(\mu \otimes \nu)\\ &= \int_{S_x(\mathcal{U})} \theta^+(x)\, d\mu \int_{S_y(\mathcal{U})} \chi^+(y)\, d\nu.\end{aligned}$$

We now note that $\int_{S_x(\mathcal{U})} \theta^+(x)\,d\mu < \int_{S_x(\mathcal{U})} \phi(x)\,d\mu + \varepsilon$ and $\int_{S_y(\mathcal{U})} \chi^+(y)\,d\nu < \int_{S_y(\mathcal{U})} \psi(y)\,d\nu + \varepsilon$, and these inequalities give us the desired upper bound. $\square$

We now assume NIP, and see that every Keisler measure over a model admits a smooth extension.

Lemma 4.2.30 (Generalizes [Kei87, 3.26])**.** *Every Keisler measure $\mu \in \mathfrak{M}_x(M)$ over a small model M admits a smooth extension over some $M \preceq N$.*

Proof. Assume for contradiction that μ has no smooth extensions. Then we inductively build a chain of extensions of measures indexed by the ordinal $|\mathcal{L}|^+$.

That is, we will construct $((M_\alpha, \mu_\alpha) : \alpha < |\mathcal{L}|^+)$, with $(M_0, \mu_0) = (M, \mu)$ and for each $\alpha < \beta$, $M_\alpha \subseteq M_\beta$ and $\mu_\alpha = \mu_\beta|_{M_\alpha}$. At limit stages, we can take a union of the models and the measures, so we can just define the successor steps. Let (M_α, μ_α) be defined. As μ_α extends μ, it is not smooth, so let μ^+, μ^- be two distinct global extensions of μ_α. As they are distinct, there is some formula $\phi_\alpha(x; b_\alpha)$ with $\phi_\alpha \in \mathcal{F}_{xy}$ such that $\int_{S_x(\mathcal{U})} \phi_\alpha(x; b_\alpha)\,d\mu^+ \geq \int_{S_x(\mathcal{U})} \phi_\alpha(x; b_\alpha)\,d\mu^- + \varepsilon_\alpha$ for some $\varepsilon_\alpha > 0$. We let $M_{\alpha+1}$ be a model containing M_α and b_α, and let $\mu_{\alpha+1} = \left(\frac{1}{2}\left(\mu^+ + \mu^-\right)\right)|_{M_{\alpha+1}}$. We then see that for any $\theta(x)$ with parameters in M_α, as $\int_{S_x(\mathcal{U})} \theta(x)\,d\mu^+ = \int_{S_x(\mathcal{U})} \theta(x)\,d\mu^-$,

$$\begin{aligned}
&\int_{S_x(\mathcal{U})} |\theta(x) - \phi_\alpha(x; b_\alpha)|\; d\mu^+ + \int_{S_x(\mathcal{U})} |\theta(x) - \phi_\alpha(x; b_\alpha)|\; d\mu^- \\
\geq &\left|\int_{S_x(\mathcal{U})} \theta(x) - \phi_\alpha(x; b_\alpha)\,d\mu^+\right| + \left|\int_{S_x(\mathcal{U})} \theta(x) - \phi_\alpha(x; b_\alpha)\,d\mu^-\right| \\
\geq &\left|\int_{S_x(\mathcal{U})} \phi_\alpha(x; b_\alpha)\,d\mu^+ - \int_{S_x(\mathcal{U})} \phi_\alpha(x; b_\alpha)\,d\mu^-\right| \\
\geq &\varepsilon_\alpha.
\end{aligned}$$

Thus either $\int_{S_x(\mathcal{U})} |\theta(x) - \phi_\alpha(x; b_\alpha)|\; d\mu^+ \geq \frac{\varepsilon_\alpha}{2}$ or $\int_{S_x(\mathcal{U})} |\theta(x) - \phi_\alpha(x; b_\alpha)|\; d\mu^- \geq \frac{\varepsilon_\alpha}{2}$, so

$$\int_{S_x(M_{\alpha+1})} |\theta(x) - \phi_\alpha(x; b_\alpha)|\; d\mu_{\alpha+1} \geq \frac{\varepsilon_\alpha}{4}.$$

We may assume that each ε_α is rational. Then by an infinite pigeonhole principle, and the fact that there are at most $|\mathcal{L}|$ choices of $(\phi_\alpha, \varepsilon_\alpha)$, we can restrict to a subsequence of the same length such that ϕ_α and ε_α are constant. We call these constant values simply ϕ and ε. Thus if we let M' be the union of all the models and μ' be the union of all the measures in our new sequence, have an infinite sequence $(b_\alpha : \alpha < |\mathcal{L}|^+)$ such that for all $\alpha < \beta$,

$$\int_{S_x(M')} |\phi(x; b_\alpha) - \phi(x; b_\beta)|\, d\mu' \geq \varepsilon.$$

Now using the same Ramsey and compactness argument as in the proof of Lemma 4.2.4, we extract an $\mathcal{L}_{\mathbb{E}}$-indiscernible sequence with the same EM-type as $(b_\alpha : \alpha < |\mathcal{L}|^+)$. Call this indiscernible $(b'_i : i < \omega)$, and let $(M^*; \mu^*)$ be an elementary extension of $(M'; \mu')$ containing it.

Then in particular, for all $i < \omega$,

$$\int_{S_x(M^*)} |\phi(x; b'_{2i}) - \phi(x; b'_{2i+1})|\, d\mu^* \geq \varepsilon.$$

so by Lemma 4.2.4, the partial type $\{|\phi(x; b'_{2i}) - \phi(x; b'_{2i+1})| \geq \frac{\varepsilon}{2} : i < \omega\}$ is consistent. As $(b'_i : i < \omega)$ is indiscernible, this contradicts NIP. $\square$

We can now use smooth extensions to prove that the Morley product is associative in an NIP context.

Lemma 4.2.31. *Assume T is NIP. Let $\mu \in \mathfrak{M}_x(\mathcal{U}), \nu \in \mathfrak{M}_y(\mathcal{U}), \lambda \in \mathfrak{M}_z(\mathcal{U})$ be M-Borel definable, with $\mu \otimes \nu$ also M-Borel definable. Then $(\mu \otimes \nu) \otimes \lambda = \mu \otimes (\nu \otimes \lambda)$.*

Proof. The proof in [CG21, Section 3.1] suffices, as we have proven all of the ingredients of that proof still hold in the case of metric structures. Specifically, it only requires the following tools,

- associativity of the Morley product for smooth (or just definable) measures (Lemma 4.2.9)

- the existence of smooth extensions (Lemma 4.2.30)

- the fact that the Morley product of smooth measures is smooth (Lemma 4.2.29)

- the fact that if μ is A-invariant, $(\mu \otimes \nu)|_A$ depends only on $\nu|_A$

all of which we have established in continuous logic. $\square$

4.3 Generically Stable Measures

In this section, we will obtain a continuous version of [HPS13, Theorem 3.2], which characterize generically stable measures in NIP theories. We will prove the following properties are equivalent, and we will call any measure satisfying them *generically stable*. We then show how to find generically stable measures in metric structures using ultraproducts or averaging indiscernible segments.

Theorem 4.3.1. *Assume T is NIP. For any small model $M \subseteq \mathcal{U}$, if μ is a global M-invariant measure, the following are equivalent:*

1. *μ is fim over M*

2. *μ is fam over M*

3. *μ is dfs over M*

4. *$\mu(x) \otimes \mu(y) = \mu(y) \otimes \mu(x)$*

5. *$\mu^{(\omega)}(x_0, x_1, \dots)|_M$ is totally indiscernible.*

Once we have this equivalence, we can see that by Lemma 4.2.27, smooth measures are generically stable in NIP.

Several of these implications follow without the NIP assumption:

Lemma 4.3.2. *For any small model $M \subseteq \mathcal{U}$, if μ is a global M-invariant measure, then each property implies the next:*

1. *μ is fim over M*
2. *μ is fam over M*
3. *μ is dfs over M*

Proof. $1 \implies 2$ follows by definition.

$2 \implies 3$: assume μ is fam over M. First we check that μ is approximately realized in M, which will imply that μ is M-invariant. Fix $\varepsilon > 0$ and $\phi(x; b)$ such that $\int \phi(x)\, d\mu < \varepsilon$. As μ is fam, there exists some $\frac{1}{2}(\varepsilon - \int \phi(x; b)\, d\mu)$-approximation $(a_1, \ldots, a_n) \in (M^x)^n$ for $\phi(x; b)$ with respect to μ. Thus $|\mathrm{Av}(a_1, \ldots, a_n; \phi(x; b)) - \int \phi(x; b)\, d\mu)| \leq \frac{1}{2}(\varepsilon - \int \phi(x)\, d\mu)$, from which we conclude that $\mathrm{Av}(a_1, \ldots, a_n; \phi(x; b)) < \varepsilon$. This means that for at least one a_i, $\vDash \phi(a_i; b) < \varepsilon$.

Now we check definability. Fix $\varphi(x; y) \in \mathcal{L}$. Then for each $\varepsilon > 0$, there is a tuple $\bar{a}$ such that $|F^{\phi}_{\mu,M}(y) - F^{\phi}_{\mathrm{Av}(\bar{a}),M}(y)| < \varepsilon$ for all y. Thus if $(\bar{a}_n : n \in \omega)$ is a sequence of tuples with $|F^{\phi}_{\mu,M}(y) - F^{\phi}_{\mathrm{Av}(\bar{a}_n),M}(y)| < 2^{-n}$, then $\lim_{n\to\infty} F^{\phi}_{\mathrm{Av}(\bar{a}_n),M}(y) = F^{\phi}_{\mu,M}(y)$ is a uniform limit of continuous functions, which is thus continuous. $\square$

For the rest of this subsection, we will assume T is NIP.

The following lemma shows that if μ is dfs over M, then μ commutes with itself.

Lemma 4.3.3 (Generalizing [Sim15, Prop. 2.26])**.** *Let $\mu \in \mathfrak{M}_x(\mathcal{U}), \nu \in \mathfrak{M}_y(\mathcal{U})$, with μ M-definable and ν approximately realized in M. Then $\mu \otimes \nu = \nu \otimes \mu$.*

Proof. We take the general approach to this from [CGH23a].

First, we assume that ν is the Dirac measure of a type realized by some tuple b in M. Then we see that for any $\phi(x;y)$, $F_{\nu}^{\phi^*}(x) = \phi(x;b)$, so

$$\begin{aligned}\int_{S_{xy}(\mathcal{U})} \phi(x;y)\,d\mu\otimes\nu =& \int_{S_y(\mathcal{U})} F_{\mu}^{\phi}(y)\,d\nu \\ =& F_{\mu}^{\phi}(b) \\ =& \int_{S_x(\mathcal{U})} \phi(x;b)\,d\mu \\ =& \int_{S_{xy}(\mathcal{U})} \phi(x;y)\,d\nu\otimes\mu\end{aligned}$$

Then we see that if we take a convex combination of measures that commute with μ, they will also commute with ν, as convex combinations commute with integration. Thus ν is the limit of a net of measures that commute with μ, call these $(\nu_i : i \in I)$ for a directed set I.

Now let $\phi(x;y)$ be a formula with parameters. By enlarging M if necessary, we may assume that M contains all the parameters of ϕ, as μ will still be M-definable and ν still approximately realized in M. By Lemma 4.2.30, let $\hat{\mu}$ be a global smooth extension of $\mu|_M$. We now see that

$$\begin{aligned}\int_{S_{xy}(\mathcal{U})} \phi(x;y)\,d\mu\otimes\nu =& \lim_{i\in I}\int_{S_{xy}(\mathcal{U})} \phi(x;y)\,d\mu\otimes\nu_i \\ =& \lim_{i\in I}\int_{S_{xy}(\mathcal{U})} \phi(x;y)\,d\nu_i\otimes\mu \\ =& \lim_{i\in I}\int_{S_{xy}(\mathcal{U})} \phi(x;y)\,d\nu_i\otimes\hat{\mu} \\ =& \lim_{i\in I}\int_{S_{xy}(\mathcal{U})} \phi(x;y)\,d\hat{\mu}\otimes\nu_i \\ =& \int_{S_{xy}(\mathcal{U})} \phi(x;y)\,d\hat{\mu}\otimes\nu \\ =& \int_{S_{xy}(\mathcal{U})} \phi(x;y)\,d\nu\otimes\hat{\mu} \\ =& \int_{S_{xy}(\mathcal{U})} \phi(x;y)\,d\nu\otimes\mu.\end{aligned}$$

The first and fifth equations are justified by the fact that for any definable μ', the map $\nu' \mapsto \int_{S_{xy}(\mathcal{U})} \phi(x;y)\, d\mu' \otimes \nu'$ is continuous. This is true as $\int_{S_{xy}(\mathcal{U})} \phi(x;y)\, d\mu' \otimes \nu' = \int_{S_y(\mathcal{U})} F^{\phi}_{\mu'}\, d\nu'$, and by the definition of the topology of $\mathfrak{M}_x(\mathcal{U})$, the integral of a continuous predicate is continuous as a function on the space of measures.

The third and seventh equations are justified by the fact that for any A-invariant μ, ν', $(\nu' \otimes \mu')|_A$ only depends on $\mu'|_A$. The second follows from our observation that the ν_is commute with μ, and the fourth and sixth follow from Corollary 4.2.25. $\square$

The total indiscernibility of $\mu^{(\omega)}(x_0, x_1, x_2, \dots)$ follows from the associativity of the Morley product in NIP combined with Lemma 4.3.3,

We now work towards showing that the indiscernibility of $\mu^{(\omega)}(x_0, x_1, x_2, \dots)$ implies fim, following [HPS13, Theorem 3.2]. We will assume that μ is a measure such that $\mu^{(\omega)}(x_0, x_1, x_2, \dots)$ is well-defined. That is, we can recursively define $\mu^{(n)}(x_0, x_1, \dots, x_{n-1})$, and each time it will be M-Borel definable. At the moment, we know that this is true if μ is M-definable.

For the following lemmas, we will need some more notation, following [Sim15, Section 7.5]. If $\phi(x;y)$ is a formula and $n \in \mathbb{N}$, define the formula

$$f_n^{\phi}(\bar{x}, \bar{x}') = \sup_y |\mathrm{Av}(x_1, \dots, x_n; \phi(x;y)) - \mathrm{Av}(x'_1, \dots, x'_n; \phi(x;y))|.$$

Lemma 4.3.4 (Generalizes [Sim15, Lemma 7.24])**.** *Let $\phi(x;y)$ be a formula. Then for any n, any Keisler measure $\mu \in \mathfrak{M}_x(M)$ with $\mu^{(2n)}$ totally indiscernible, and any $\varepsilon > 0$,*

$$\mu^{(2n)}\left(f_n^{\phi}(\bar{x}, \bar{x}') > \varepsilon\right) \leq 4\mathcal{N}_{\phi(x;y),\varepsilon/4}(n) \exp\left(-\frac{n\varepsilon^2}{32}\right).$$

Proof. Let $R = \{-1, 1\}^n$. We claim that

$$\mu^{(2n)}(f_n^{\phi}(\bar{x}, \bar{x}') > \varepsilon) \leq \frac{1}{2^{n-1}} \sum_{\sigma \in R} \mu^{(n)}\left(\left\{\bar{x} : \sup_y \frac{1}{n}\left|\sum_{i=1}^{n} \sigma_i \phi(x_i; y)\right| > \frac{\varepsilon}{2}\right\}\right).$$

By definition, we have

$$\mu^{(2n)}(f_n^{\phi}(\bar{x},\bar{x}') > \varepsilon) = \mu\left(\sup_y \frac{1}{n}\left|\sum_{i=1}^{n}(\phi(x_i;y) - \phi(x_i';y))\right| > \varepsilon\right),$$

and by symmetry, this equals

$$\begin{aligned}
&\frac{1}{2^n}\sum_{\sigma\in R}\mu_{2n}\left(\sup_y \frac{1}{n}\left|\sum_{i=1}^{n}\sigma_i(\phi(x_i;y) - \phi(x_i';y))\right| > \varepsilon\right)\\
\leq&\frac{1}{2^n}\sum_{\sigma\in R}\mu_{2n}\left(\sup_y \frac{1}{n}\left|\sum_{i=1}^{n}\sigma_i\phi(x_i;y)\right| > \frac{\varepsilon}{2} \text{ or } \sup_y \frac{1}{n}\left|\sum_{i=1}^{n}\sigma_i\phi(x_i';y)\right| > \frac{\varepsilon}{2}\right)\\
\leq&\frac{1}{2^{n-1}}\sum_{\sigma\in R}\mu_{n}\left(\sup_y \frac{1}{n}\left|\sum_{i=1}^{n}\sigma_i\phi(x_i;y)\right| > \frac{\varepsilon}{2}\right)
\end{aligned}$$

where in the last inequality we use symmetry and a union bound, proving the claim.

For any vector $\bar{c} \in [0,1]^n$, and $\delta > 0$, let $R(\bar{c},\delta)$ be the set of vectors $\sigma \in R$ such that $\frac{1}{n}|\sigma \cdot \bar{c}| > \delta$. For each $\bar{a} \in (M^x)^n$, let $\phi(\bar{a};y) = (\phi(a_1;y),\ldots,\phi(a_n;y))$. We will bound $|\bigcup_y R(\phi(\bar{a};y),\frac{\varepsilon}{2})|$. If we show that this is at most $2^{n-1}\cdot 4\mathcal{N}_{\phi(x;y),\varepsilon/4}(n)\exp\left(-\frac{n\varepsilon^2}{32}\right)$, then the lemma follows.

We first observe that for δ, δ' and vectors $\bar{c},\bar{c}' \in [0,1]^n$, if $|\bar{c}-\bar{c}'|_\infty < \delta'$, then $R(\bar{c},\delta) \subseteq R(\bar{c}',\delta-\delta')$. To see this, let $\sigma \in R(\bar{c},\delta)$, that is, $\frac{1}{n}|\sigma\cdot\bar{c}| > \delta$, so

$$\frac{1}{n}|\sigma\cdot\bar{c}'| > \frac{1}{n}(|\sigma\cdot\bar{c}| - |\sigma\cdot(\bar{c}-\bar{c}')|) \geq \delta - |\bar{c}-\bar{c}'|_\infty \geq \delta-\delta'.$$

For any given $\bar{a} \in X^n$, there exists a set $C \subset [0,1]^n$ of size $\mathcal{N}_{\phi(x;y),\varepsilon/4}(n)$ such that for every $b \in \mathcal{U}^y$, there is $\bar{c} \in C^n$ such that $|\phi(\bar{a};b) - \bar{c}|_\infty \leq \frac{\varepsilon}{4}$, and thus $R(\phi(\bar{a};b),\frac{\varepsilon}{2}) \subseteq R(\bar{c},\frac{\varepsilon}{4})$. Thus $\bigcup_{b\in\mathcal{U}^y} R(\phi(\bar{a};b),\frac{\varepsilon}{2}) \subseteq \bigcup_{\bar{c}\in C} R(\bar{c},\frac{\varepsilon}{4})$, and

$$\left|\bigcup_{b\in\mathcal{U}^y} R(\phi(\bar{a};y),\frac{\varepsilon}{2})\right| \leq |C|\max_{\bar{c}\in C}\left|R\left(\bar{c},\frac{\varepsilon}{4}\right)\right| \leq \mathcal{N}_{\phi(x;y),\varepsilon/4}(n)\max_{\bar{c}\in C}\left|R\left(\bar{c},\frac{\varepsilon}{4}\right)\right|.$$

For each individual vector $\bar{c}$, we can think of $\frac{1}{2^n}\left|R\left(\bar{c}, \frac{\varepsilon}{4}\right)\right|$ probabilistically. If σ is chosen uniformly at random from R, then this is $\mathbb{P}_\sigma[\frac{1}{n}|\sigma \cdot \bar{c}| > \frac{\varepsilon}{4}]$. As $\sigma \cdot \bar{c} = \sum_i \sigma_i c_i$ is the sum of n independent random variables of mean 0 supported on $[-\frac{1}{n}, \frac{1}{n}]$, we can apply Hoeffding's inequality to find that $\mathbb{P}_\sigma[\frac{1}{n}|\sigma \cdot \bar{c}| > \frac{\varepsilon}{4}] \leq 2\exp\left(-\frac{n\varepsilon^2}{32}\right)$. Thus $|\bigcup_{b \in \mathcal{U}^y} R(\phi(\bar{a}; b), \frac{\varepsilon}{2})| \leq 2^n \mathcal{N}_{\phi(x;y),\varepsilon/4}(n)\left(2\exp\left(-\frac{n\varepsilon^2}{32}\right)\right)$, as desired. □

Lemma 4.3.5 (See [Sim15, Proposition 7.26])**.** *Let $\phi(x; y)$ be a formula. Let any n, $\varepsilon > 0$ be such that $n \geq \frac{9}{2\varepsilon^2}$ and $\mathcal{N}_{\phi(x;y),\varepsilon/12}(n)\mathrm{exp}\left(-\frac{n\varepsilon^2}{96}\right) < \frac{1}{8}$.*

Then for any Keisler measure $\mu \in \mathfrak{M}_x(M)$ with $\mu^{(2n)}|_M$ totally indiscernible, there is a formula $\theta_{n,\varepsilon}(x_1, \ldots, x_n)$ with parameters in M such that

- *Any $\bar{a} \in (\mathcal{U}^x)^n$ satisfying $\theta_{n,\varepsilon}(\bar{a}) = 0$ is a ε-approximation to $\phi(x; y)$ with respect to μ.*
- *$\mu^{(n)}(\theta_{n,\varepsilon}(x_1, \ldots, x_n)) \geq 1 - 8\mathcal{N}_{\phi(x;y),\varepsilon/12}(n)\exp\left(-\frac{n\varepsilon^2}{96}\right)$*

Proof. Let $\theta'_{n,\varepsilon}(x_1, \ldots, x_n, x'_1, \ldots, x'_n) = f^\phi_n(x_1, \ldots, x_n, x'_1, \ldots, x'_n) \dot{-} \frac{\varepsilon}{3}$.

Then by Lemma 4.3.4, $\mu^{(2n)}(\theta'_{n,\varepsilon}) \geq 1 - 4\mathcal{N}_{\phi(x;y),\varepsilon/12}(n)\mathrm{exp}\left(-\frac{n\varepsilon^2}{96}\right)$. Thus there exists $\bar{a}' \in (\mathcal{U}^x)^n$ such that $\mu^{(n)}(\theta'_{n,\varepsilon}(x_1, \ldots, x_n; \bar{a}')) \geq 1 - 4\mathcal{N}_{\phi(x;y),\varepsilon/12}(n)\mathrm{exp}\left(-\frac{n\varepsilon^2}{96}\right) > \frac{1}{2}$. Then let $\theta_{n,\varepsilon} = \theta'_{n,\varepsilon}(\bar{x}; \bar{a}')$.

It now suffices to show that for all $\bar{a}$ satisfying $\theta'_{n,\varepsilon}(\bar{a}; \bar{a}') = 0$, and any $b \in \mathcal{U}^y$, $\bar{a}$ is a ε-approximation to $\phi(x; b)$ with respect to μ. Fix b, and let $\zeta_n(x_1, \ldots, x_n) = |\mathrm{Av}(x_1, \ldots, x_n) - \mathbb{E}_\mu[\phi(x; b)]| \dot{-} \frac{\varepsilon}{3}$. By the weak law of large numbers, as the functions $\phi(x_i; b)$ are $[0, 1]$-valued i.i.d. random variables with respect to μ, $\mu^{(n)}(\zeta_n = 0) \geq 1 - \frac{1}{4n(\varepsilon/3)^2} = 1 - \frac{9}{4n\varepsilon^2}$. Thus as $n \geq \frac{9}{2\varepsilon^2}$, we have $\mu^{(n)}(\zeta_n) \geq \frac{1}{2}$. As $\mu^{(n)}(\theta_{n,\varepsilon} = 0) > \frac{1}{2}$ also, $\mu^{(n)}(\theta_{n,\varepsilon} \wedge \zeta_n = 0) > 0$, so let $\bar{a}^*$ be such that $\vDash \theta_{n,\varepsilon}(\bar{a}^*) = 0$ and $\zeta_n(\bar{a}^*) = 0$. Thus $|\mathrm{Av}(\bar{a}') - \mathbb{E}_\mu[\phi(x; b)]| \leq |\mathrm{Av}(\bar{a}') - \mathrm{Av}(\bar{a}^*)| + |\mathrm{Av}(\bar{a}^*) - \mathbb{E}_\mu[\phi(x; b)]| \leq \frac{2\varepsilon}{3}$. Thus if $\bar{a}$ is such that $\vDash \theta'_{n,\varepsilon}(\bar{a}; \bar{a}') = 0$, we also have $|\mathrm{Av}(\bar{a}) - \mathbb{E}_\mu[\phi(x; b)]| \leq \varepsilon$. □

Theorem 4.3.6 (Generalizes [Sim15, Proposition 7.26], [ABC97, Lemma 3.3])**.** *Any Keisler measure $\mu \in \mathfrak{M}_x(\mathcal{U})$ with $\mu^{(\omega)}|_M$ totally indiscernible is fim over M.*

Proof. Let $\phi(x;y)$ be a formula. By [And23b, Lemma 3.4.12], for all $\varepsilon > 0$, there exist C, k such that $\mathcal{N}_{\phi(x;y),\varepsilon/12}(n) \leq Cn^k$ for all $n \geq 2$. Thus we have

$$\lim_{n\to\infty} \mathcal{N}_{\phi(x;y),\varepsilon/12}(n)\exp\left(-\frac{n\varepsilon^2}{96}\right) \leq \lim_{n\to\infty} n^k\exp\left(-\frac{n\varepsilon^2}{96}\right) = 0.$$

Thus for each $m \in \mathbb{N}$, we can let $n_m \in \mathbb{N}$ be such that $n_m \geq \frac{9}{2}m^2$ and $n^k\exp\left(-\frac{n}{96m^2}\right) < \frac{1}{8m}$.

Then for each $n \in \mathbb{N}$, let $\theta_n(x_1, \ldots, x_n) = \theta_{n,1/m_n}(x_1, \ldots, x_n)$, where $m_n \in \mathbb{N}$ is the greatest natural number such that $n_{m_n} \leq n$. Then by Lemma 4.3.5, any $\bar{a} \in (\mathcal{U}^x)^n$ satisfying $\theta_{n,1/m_n}(\bar{a}) = 0$ is a $\frac{1}{m_n}$-approximation to $\phi(x;y)$ with respect to μ, and

$$\mu^{(n)}(\theta_n(x_1, \ldots, x_n)) \geq 1 - 8\mathcal{N}_{\phi(x;y),1/(12m_n)}(n) \exp\left(-\frac{n}{96m_n^2}\right) \geq 1 - \frac{1}{m_n}.$$

Now it suffices to show that $\lim_{n\to\infty} m_n = \infty$. This is true as for each m, if $n \geq n_m$, then $m_n \geq m$. $\square$

Corollary 4.3.7. *If $\mathrm{vc}_{\varepsilon/12}(\phi(x;y)) \leq d$, and $\mu \in \mathfrak{M}_x(\mathcal{U})$ is such that $\mu^{(\omega)}|_M$ is totally indiscernible, then $\phi(x;y)$ admits a ε-approximation of size at most $O(\frac{d}{\varepsilon^2}\ln\frac{d}{\varepsilon})$ with respect to μ.*

Proof. It suffices to find n such that $n \geq \frac{9}{2\varepsilon^2}$ and $\mathcal{N}_{\phi(x;y),\varepsilon/12}(n)\exp\left(-\frac{n\varepsilon^2}{96}\right) < \frac{1}{8}$. Then by Lemma 4.3.5, there is a positive-measure (with respect to $\mu^{(n)}$) set of ε-approximations of size n to $\phi(x;y)$ with respect to μ.

Let C be the constant depending on d, ε such that $\mathcal{N}_{\phi(x;y),\varepsilon/12}(n) \leq n^{C\ln n} = e^{C\ln^2 n}$, according to [And23b, Fact 3.2.18]. $\square$

Lemma 4.3.8. *If $\mu \in \mathfrak{M}_x(\mathcal{U})$ is dfs over M, then it is fim over M.*

Proof. Let μ be dfs. We know by Lemma 4.3.3 that $\mu(x) \otimes \mu(y) = \mu(y) \otimes \mu(x)$. By definability of μ and Lemma 4.2.9, we know that $\mu^{(\omega)}(x_0, \ldots, x_{n-1})|_M$ is well-defined, and by commutativity, it is totally indiscernible. Thus by Theorem 4.3.6, μ is fim over M. $\square$

In particular, smooth measures are fim, so any measure admits a fim extension, from which we can show that every measure is locally approximated by types in its support.

Lemma 4.3.9. *Let $\mu \in \mathfrak{M}_x(M)$ be a Keisler measure, $\phi(x; y)$ a definable predicate, $\varepsilon > 0$.*

There are types $p_1, \dots, p_n \in S(\mu)$ such that for every $b \in M^y$,

$$\left| \int_{S_x(M)} \phi(x; b)\, d\mu - \mathrm{Av}(p_1, \dots, p_n; \phi(x; b)) \right| \leq \varepsilon.$$

Proof. Let $\nu \in \mathfrak{M}_x(N)$ be a fim extension of μ, with $M \preceq N$. Then for every $\phi(x; y)$, there is some closed N-condition $\theta(x_1, \dots, x_n)$ with $\nu^{(n)}(\theta(x_1, \dots, x_n)) > \frac{1}{2}$, such that any $(a_1, \dots, a_n)$ with $\vDash \theta(a_1, \dots, a_n)$ is a ε-approximation to $\phi(x; y)$ with respect to μ.

We claim that there are some $a_1, \dots, a_n$ such that for each i, $\mathrm{tp}(a_i/M) \in S(\mu)$ and $\vDash \theta(a_1, \dots, a_n)$. Then we can let $p_1, \dots, p_n$ be the types of $a_1, \dots, a_n$ over M. To do that, we just have to show that any finite set of closed conditions in $\{\theta_n(x_1, \dots, x_n)\} \cup \bigcup_{i=1}^n S_\mu(x_i)$ is satisfiable, where $S_\mu(x_i)$ is the partial type indicating that $\mathrm{tp}(x_i/M) \in S(\mu)$, consisting of all closed M-conditions with positive μ-measure. As ν is an extension of μ, the $\nu^{(n)}$-measure of the intersection of the finite set is at least $\nu(\theta_n(x_1, \dots, x_n)) > \frac{1}{2}$. Thus this finite partial type is satisfiable. □

Finally we are able to show that assuming NIP, every M-invariant measure is M-Borel definable, simplifying many of our earlier results. This was originally shown for classical logic in [HP11] using a VC-Theorem argument.

Lemma 4.3.10. *Let $\mu \in \mathfrak{M}_x(\mathcal{U})$ be M-invariant. Then μ is M-Borel definable.*

Proof. Let $\phi(x; y)$ be a definable predicate. For every $\varepsilon > 0$, we will find a Borel function $f_\varepsilon : S_y(\mathcal{U}) \to [0, 1]$ such that $|F^\phi_{\mu,M} - f_\varepsilon(y)| \leq \varepsilon$. Then $F^\phi_{\mu,M}$ is a uniform limit of Borel functions, and as Borel functions are closed under even pointwise limits, $F^\phi_{\mu,M}$ is Borel.

By Lemma 4.3.9, there are types $p_1, \dots, p_n \in S(\mu)$ such that for every $b \in M^y$,

$$\left| \int_{S_x(M)} \phi(x;b)\, d\mu - \mathrm{Av}(p_1, \dots, p_n; \phi(x;b)) \right| \leq \varepsilon.$$

Let $f_\varepsilon(y) = \mathrm{Av}(p_1, \dots, p_n; \phi(x;b))$. As each p_i is M-invariant by Lemma 4.2.6, and an average of Borel functions is Borel, it is enough to show that the Dirac measure of an invariant type is Borel-definable.

Let p be an M-invariant type. It suffices to show that for each $r > 0$, the set $\{q : \phi(x;b) < r \in p(x) \text{ for } b \vDash q\}$ is Borel.

Fix $b \in \mathcal{U}$, $\varepsilon > 0$. By NIP, there is some maximal N such that there is $(a_i : i \leq N) \vDash p^{(N)}(x)|_M$ with $|\phi(a_i;b) - \phi(a_{i+1};b)| \geq \varepsilon$ for $0 \leq i < N$. By the maximality of N, we see that $|\phi(p;b) - \phi(a_N;b)| < \varepsilon$.

Now let $A_{N,\varepsilon}(y)$ indicate the set in $S_y(M)$ such that there exist $a_0, \dots, a_N$ satisfying $p^{(N)}|_M$ such that $|\phi(a_i;y) - \phi(a_{i+1};y)| \geq \varepsilon$ for all $i < N$, and $\phi(a_N;y) \leq r - \varepsilon$. This set is closed, as by saturation, this holds if and only if for each closed condition $\chi(x_0, \dots, x_n) = 0 \in p^{(N)}|_M$,

$$\inf_{x_0,\dots,x_n} \max\left(\chi(x_0, \dots, x_n), \phi(a_N;y) \dot{-} (r - \varepsilon), \max_{i<N} \left(\varepsilon \dot{-} |\phi(a_i;y) - \phi(a_{i+1};y)|\right)\right) = 0$$

holds.

Let $B_{N,\varepsilon}(y)$ be the weaker condition that there exist $a_0, \dots, a_N$ satisfying $p^{(N)}|_M$ such that $|\phi(a_i;y) - \phi(a_{i+1};y)| \geq \varepsilon$ for all $i < N$. Then by NIP, for every $\varepsilon > 0$, every $b \in \mathcal{U}^y$, there is some N such that $B_{N,\varepsilon}(b)$ holds, but $B_{N+1,\varepsilon}(b)$ does not. If in addition, $p \vDash \phi(x;b) < r - 2\varepsilon$, then we know that $A_{N,\varepsilon}(b)$ holds, as in any maximal sequence $a_0, \dots, a_N$ witnessing $B_{N,\varepsilon}(b)$, we have $|\phi(p;b) - \phi(a_N;b)| < \varepsilon$, so $\phi(a_N;b) < r - \varepsilon$. Also, if b is such that $A_{N,\varepsilon}(y)$ holds but $B_{N+1,\varepsilon}(b)$ does not, then $p \vDash \phi(x;b) < r$, as in any witness sequence $a_0, \dots, a_N$, we have $|\phi(p;b) - \phi(a_N;b)| < \varepsilon$ and $\phi(a_N;b) < r - \varepsilon$. Thus $\{q : \phi(x;b) < r \in p(x) \text{ for } b \vDash q\} =$

$\bigcup_{N,m\in\mathbb{N}} \left(A_{N,1/m}(y) \setminus B_{N+1,1/m}(y)\right)$, which is a countable union of boolean combinations of closed sets, and is thus Borel.

□

We can also use the indiscernibility of $\mu^{(\omega)}$ to prove a version of [And23b, Theorem 3.2.25] with respect to generically stable measures:

Theorem 4.3.11. *For any $\varepsilon > 0$, $d \in \mathbb{N}$, $0 \leq r < s \leq 1$, there is $N = O_{r,s}(d\varepsilon^{-1}\log\varepsilon^{-1})$ such that if $\phi(x;y)$ is a definable predicate with $\mathrm{vc}_{r',s'}(\phi(x;y)) \leq d$ for some $r < r' < s' < s$, and $\mu \in \mathfrak{M}_x(M)$ is generically stable, then there is an ε-net A for the fuzzy set system $\phi_{r,s}^{M^y}$ with respect to μ with $|A| \leq N$.*

Proof. This proof generalizes the argument by Haussler and Welzl used in [Mat02, Theorem 10.2.4].

Fix $\varepsilon > 0$, d, $\phi(x;y)$ a formula, and $\mu \in \mathfrak{M}_x(M)$ generically stable. Let r', s' be such that that $r < r' < s' < s$ and $\mathrm{vc}_{r',s'}(\phi(x;y)) \leq d$.

Let $N = Cd\varepsilon\log(\varepsilon^{-1})$, with C to be determined later.

We will define an open set $E_0 \subseteq S_N(M)$ such that for every tuple $(a_1,\ldots,a_N)$ that is not a ε-net, $\mathrm{tp}(a_1,\ldots,a_N/M) \in E_0$. Then we will find conditions on N that guarantee $\mu^{(N)}(E_0) < 1$, implying that there exists some $(a_1,\ldots,a_N)$ with $\mathrm{tp}(a_1,\ldots,a_N/M) \notin E_0$, which must therefore be a ε-net.

For $b \in M^y$, let $E_{0,b} \subseteq S_N(M)$ be $\bigcap_{i=1}^N\{\phi(x_i;b) < r'\}$, and let $E_0 = \bigcup_{b\in M,\mu(\phi(x;b)\geq s)\geq\varepsilon} E_{0,b}$. We see that each $E_{0,b}$ is open, and thus E_0 is open. (The purpose of using $< r'$ instead of $\leq r$ is to guarantee measurability of E_0.) If $(a_1,\ldots,a_N)$ is not a ε-net, then there exists some $b \in M^y$ such that $\mu(\phi(x;b) \geq s) \geq \varepsilon$ and for all $1 \leq i \leq N$, $\phi(a_i;b) \leq r < r'$, so $\mathrm{tp}(a_1,\ldots,a_N/M) \in E_{0,b} \subseteq E_0$.

Now define $E_{1,b} \subseteq S_{2N}(M)$ be the (open) event that for all $1 \leq i \leq N$, $\phi(x_i;b) < r'$, and for at least $k = \lceil\frac{N\varepsilon}{2}\rceil$ values of $1 \leq i \leq N$, $\phi(x_{N+i};b) > s'$, and let E_1 be $\bigcup_{b\in M,\mu(\phi(x;b)\geq s)\geq\varepsilon} E_{1,b}$.

We wish to show that $\mu^{(N)}(E_0) \leq 2\mu^{(2N)}(E_1)$ and that $\mu^{(2N)}(E_1) > \frac{1}{2}$.

In order to show that $\mu^{(N)}(E_0) \leq 2\mu^{(2N)}(E_1)$, we will split up the tuple of variables $(x_1, \dots, x_{2N})$ into $\bar{x} = (x_1, \dots, x_N)$ and $\bar{x}' = (x_{N+1}, \dots, x_{2N})$, and look at conditional probability. By Lemma 4.2.10, as E_1 is open, the function defined by $\mu^{(N)}(E_1|p) := \mu^{(N)}_{\bar{x}'}((\bar{a}, \bar{x}') \in E_1)$ where $\bar{a} \vDash p$ is Borel, and $\mu^{(2N)}(E_1) = \int \mu^{(N)}(E_1|p)\, d(\mu^{(N)})$. Thus it suffices to show that for all p, $\chi_{E_0}(p) \leq 2\mu^N(E_1|p)$, where χ_{E_0} is the characteristic function of E_0.

Fix $p \in S_N(M)$ and $\bar{a} \vDash p$ with $\bar{a} = (a_1, \dots, a_N)$. If $p \notin E_0$, then for all $q \in S_{2N}(M)$ extending p, $q \notin E_1$, so $\chi_{E_0}(p) = 0 \leq 0 = 2\mu^N(E_1|p)$.

Now assume $p \in E_0$, and let b be such that $p \in E_{0,b}$. Let I_i for $1 \leq i \leq N$ be the indicator random variables (on $S_{2N}(M)$) for $\phi(x_{N+i}; b) > s'$, and $I = I_1 + \dots + I_N$. We have that $\mu^{(N)}(E_1|p) = \mu^{(N)}(\{I \geq k\}|p)$. The I_is are i.i.d. random variables, equalling 1 with probability $\mu(\{\phi(x; b) > s'\}) \geq \varepsilon$. By a standard Chernoff tail bound for binomial distributions, we have that $\mu^{(N)}(\{I \geq k\}|p) \geq \frac{1}{2} = \frac{1}{2}\chi_{E_0}(p)$. Thus in general, $\mu^{(N)}(E_0) \leq 2\mu^{(2N)}(E_1)$.

To show that $\mu^{(2N)}(E_1) < \frac{1}{2}$, we will instead condition on the multiset $\{x_1, \dots, x_{2N}\}$. Given a tuple $\bar{a} = (a_1, \dots, a_{2N})$ and a permutation σ of $\{1, \dots, 2N\}$, let $\sigma(\bar{a})$ refer to $(a_{\sigma(1)}, \dots, a_{\sigma(2N)})$. To formally condition on the multiset $\{x_1, \dots, x_{2N}\}$, we will show that for every tuple $\bar{a}$, $\mathbb{P}_\sigma[\sigma(\bar{a}) \in E_1] > \frac{1}{2}$, where σ is a permutation on $\{1, \dots, 2N\}$ selected uniformly at random. This probability is calculated as

$$\mathbb{P}_\sigma[\sigma(\bar{a}) \in E_1] = \frac{1}{n!} \sum_\sigma \chi_{E_1}(\sigma(\bar{a})),$$

and we also see that because $\mu^{(\omega)}|_M$ is totally indiscernible, for any σ we have

$$\int_{S_{2N}(M)} \chi_{E_1}(\bar{x})\, \mu^{(2N)} = \int_{S_{2N}(M)} \chi_{E_1}(\sigma(\bar{x}))\, \mu^{(2N)}.$$

Thus we see that

$$
\begin{aligned}
\mu^{(2N)}(E_1) &= \int_{S_{2N}(M)} \chi_{E_1}(\bar{x})\,\mu^{(2N)} \\
&= \frac{1}{n!}\sum_{\sigma}\int_{S_{2N}(M)} \chi_{E_1}(\bar{x})\,\mu^{(2N)} \\
&= \frac{1}{n!}\sum_{\sigma}\int_{S_{2N}(M)} \chi_{E_1}(\sigma(\bar{x}))\,\mu^{(2N)} \\
&= \int_{S_{2N}(M)} \frac{1}{n!}\sum_{\sigma} \chi_{E_1}(\sigma(\bar{x}))\,\mu^{(2N)} \\
&= \int_{S_{2N}(M)} \mathbb{P}_{\sigma}[(\sigma(\bar{x})) \in E_1]\,\mu^{(2N)} \\
&< \frac{1}{2}
\end{aligned}
$$

finishing the reduction.

We now fix $\bar{a}$ and work with finite probability, selecting a random permutation σ of the variables in $\bar{a}$.

Let $\mathcal{F}$ be the fuzzy set system on $\{1, \dots, 2N\}$ consisting of the fuzzy sets S_b for $b \in M^y$ where $S_{b+} = \{i : \phi(a_i; b) > s'\}$ and $S_{b-} = \{i : \phi(a_i; b) < r'\}$. By [And23b, Lemma 3.2.6], there is a strong disambiguation $\mathcal{F}'$ for $\mathcal{F}$ of size $|\mathcal{F}'| = (2N)^{O(d \log(2N))}$, or as we will prefer later, there is C' such that $|\mathcal{F}'| \leq (2N)^{C'(d \log(2N))}$. Recall that this means that for all $b \in M^y$, there is some $S \in \mathcal{F}'$ with $\{i : \phi(a_i; b) > s'\} \subseteq S$ and $\{i : \phi(a_i; b) < r'\} \cap S = \emptyset$. Given a set $S \in \mathcal{F}'$, let E_S be the event that $\{\sigma(1), \dots, \sigma(N)\} \cap S = \emptyset$ and $|\{\sigma(N+1), \dots, \sigma(2N)\} \setminus S| \geq k$. We see that if $\sigma(\bar{a}) \in E_1$, then there is some b with $\sigma(\bar{a}) \in E_{1,b}$. There is also some $S \in \mathcal{F}'$ with $\{i : \phi(a_i; b) > s'\} \subseteq S$ and $\{i : \phi(a_i; b) < r'\} \cap S = \emptyset$, so E_S occurs. Thus $\mathbb{P}_\sigma[\sigma(\bar{a}) \in E_1] \leq \sum_{S \in \mathcal{F}'} \mathbb{P}_\sigma[E_S]$. For each S, if $|S| < k$, then $\mathbb{P}_\sigma[E_S] = 0$, but if $|S| \geq k$, then $\mathbb{P}_\sigma[E_S]$ is the probability that when a permutation σ is selected uniformly at random,

$\{\sigma(1),\ldots,\sigma(N)\}\cap S=\emptyset$. This is at most

$$\frac{\binom{2N-|D\cap S'|}{N}}{\binom{2N}{N}}\leq\frac{\binom{2N-k}{N}}{\binom{2N}{N}}\leq\left(1-\frac{k}{2N}\right)^{N}\leq e^{-(k/2N)N}=\varepsilon^{Cd/4}.$$

Now we bound $\mathbb{P}_\sigma[\sigma(\bar{a})\in E_1]$:

$$\begin{aligned}\mathbb{P}_\sigma[\sigma(\bar{a})\in E_1]&\leq\sum_{S\in\mathcal{F}'}\mathbb{P}_\sigma[E_S]\\&\leq|\mathbb{F}'|\varepsilon^{-Cd/4}\\&\leq(2N)^{C'(d\log(2N))}\varepsilon^{Cd/4}\\&=\left((2Cd\varepsilon^{-1}\log\varepsilon^{-1})^{C'(\log\left(2Cd\varepsilon^{-1}\log\varepsilon^{-1}\right))}\varepsilon^{C/4}\right)^{d}\end{aligned}$$

While this expression is somewhat complicated, it is still clear that an increasing quasipolynomial function of C times a decreasing exponential of C will limit to 0, so for large enough C, we find that $\mathbb{P}_\sigma[\sigma(\bar{a})\in E_1]<\frac{1}{2}$. □

4.3.1 Ultraproducts

In this subsection, we recall the definition of the ultraproduct of a family of Keisler measures from discrete logic. See for instance [Sim15, Page 98]. Let I be an index set, $(M_i : i\in I)$ a family of models, $(\mu_i : i\in I)$ a family of Keisler measures in $\mathfrak{M}_x(M_i)$, and U an ultrafilter on I. By $[a_i : i\in I]$, if $a_i\in M_i^x$ for each i, we denote the equivalence class of $(a_i : i\in I)$, as an element of the sort over the ultraproduct $(\prod_U M_i)^x$.

Definition 4.3.12. Then we define the ultraproduct $\prod_U\mu_i$ to be the Keisler measure in $\mathfrak{M}_x(\prod_U M_i)$ such that for all $\phi(x;y)$, and all $b=[b_i : i\in I]\in(\prod_U M_i)^y$, we have

$$\int\phi(x;b)\,d\prod_U\mu_i=\lim_U\int\phi(x;b_i)\,d\mu_i,$$

where $\lim_U$ is the ultralimit, defined as the values lie in a compact subset of $\mathbb{R}$.

Lemma 4.3.13 (See [Sim16, Corollary 1.3])**.** *Assume NIP. If $(\mu_i : i \in I)$ is a sequence of generically stable measures and U an ultrafilter, then $\prod_U \mu_i$ is generically stable.*

Proof. Let $\phi(x, y, z)$ be a formula, and fix $\varepsilon > 0$ and $c = [c_i : i \in I] \in \prod_U M_i$. We will show that $\prod_U \mu_i$ is fam by finding an ε-approximation to the family $\{\phi(x, b; c) : b \in (\prod_U M_i)^y\}$.

First we observe that by Lemma 4.3.5, as each μ_i is generically stable, there is some n depending only on $\phi(x, y, z)$ and ε such that for each i, there exists an ε-approximation $(a_i^1, \dots, a_i^n)$ to $\{\phi(x, b_i; c_i) : b_i \in M_i^y\}$.

Now for $1 \leq j \leq n$, let $a^j = [a_i^j : i \in I]$. We claim that $(a^1, \dots, a^n)$ is a ε-approximation to $\{\phi(x, b; c) : b \in (\prod_U M_i)^y\}$. Fix $b = [b_i : i \in I]\,(\prod_U M_i)^y$. Then $|\mathrm{Av}(a_i^1, \dots, a_i^n; \phi(x; b_i, c_i)) - \int_{S_x(M_i)} \phi(x; b_i, c_i)\, d\mu_i| \leq \varepsilon$.

By the definition of an ultraproduct, we know that

$$\mathrm{Av}(a^1, \dots, a^n; \phi(x; b, c)) = \lim_U \mathrm{Av}(a_i^1, \dots, a_i^n; \phi(x; b_i, c_i))$$

and

$$\int_{S_x(\prod_U M_i)} \phi(x; b, c)\, d\prod_U \mu_i = \lim_U \int_{S_x(M_i)} \phi(x; b_i, c_i)\, d\mu_i,$$

and these ultralimits differ by at most ε, as the sequences do pointwise. $\square$

4.3.2 Indiscernible Segments

An *indiscernible segment* is an indiscernible sequence indexed by the order $[0, 1] \subseteq \mathbb{R}$. We will prove one characterization of NIP using indiscernible segments, and then assume T is NIP for the rest of this subsection.

Lemma 4.3.14. *A theory T is NIP if and only if the following holds: For every indiscernible segment $I = (a_i : t \in [0, 1])$ with $|a_t| = |x|$ and any definable predicate $\phi(x; b)$ with $b \in \mathcal{U}^y$, the* *function $t \mapsto \phi(a_t; b)$ is regulated (a uniform limit of step functions), and thus measurable.*

Proof. NIP is equivalent to every $\phi(x; b)$ having a limit on every indiscernible sequence of order type ω.

First, assume that T is NIP. Thus on every indiscernible segment, for any increasing or decreasing sequence $(t_n : n \in \mathbb{N}) \in [0,1]^{\mathbb{N}}$, $\lim_n \phi(a_{t_n}; b)$ exists, so at every point $t_0 \in (0,1)$, the limits $\lim_{t\to t_0^+} \phi(a_t; b)$ and $\lim_{t\to t_0^-} \phi(a_t; b)$ both exist. If they did not, we could find an increasing or decreasing sequence limiting to t_0 on which $\phi(x; b)$ has no limit. By [Bou07, Théorème 3, FVR II.5], a function on $[0,1]$ is regulated if and only if its left and right limits all exist.

Now assume that T is not NIP - there must be some indiscernible sequence $(a_n : n \in \mathbb{N})$ and some $\phi(x; b)$ with $\lim_n \phi(a_n; b)$ undefined. Then by restricting to a non-Cauchy subsequence, we find that there is some $\varepsilon > 0$ such that $|\phi(a_n; b) - \phi(a_{n+1}; b)| \geq \varepsilon$. We claim that there also exists an indiscernible segment $(a'_t : t \in [0,1])$ and some b' where $\lim_n \phi(a'_{1/n}; b')$ also does not exist, making $\phi(x_t; b')$ not regulated. We can find this counterexample by realizing the type given by $\psi(x_{t_1}, \ldots, x_{t_n}) = \psi(x_{t'_1}, \ldots, x_{t'_n})$ for all increasing tuples $t_1 < \cdots < t_n$ and $t'_1 < \cdots < t'_n$ and $|\phi(x_{1/n}; y) - \phi(x_{1/(n+1)}; y)| \geq \varepsilon$, every finite subtype of which is realized by any finite subsequence of $(a_n : n \in \mathbb{N})$. □

We now assume T is NIP. This allows us to define the average measure of an indiscernible segment.

Definition 4.3.15. If $I = (a_t : t \in [0,1])$ is an indiscernible segment, define the *average measure* of I, $\mu_I \in \mathfrak{M}_x(\mathcal{U})$, to be the unique global Keisler measure with

$$\int_{S_x(\mathcal{U})} \phi(x; b)\, d\mu_I = \int_0^1 \phi(a_t; b)\, dt.$$

We will show that these measures are generically stable using the following lemma:

Lemma 4.3.16. *If F is a family of functions $[0,1] \to [0,1]$ such that for each $f \in F$, there is no sequence $0 \le t_1 < t_1' < \cdots < t_N < t_N' \le 1$ with $|f(t_i) - f(t_i')| > \frac{\varepsilon}{2}$ for all $1 \le i \le N$, then for any $M > \frac{2N}{\varepsilon}$, the set $A = \left\{ \frac{k}{M} : 0 \le k < M \right\}$ is a ε-approximation to F with respect to the Lebesgue measure on $[0,1]$.*

Proof. For any $f \in F$,

$$\left|\mathrm{Av}_{a\in A} f(a) - \int_0^1 f(t)\,dt\right| = \frac{1}{M}\left|\sum_{k=0}^{M-1}\left(f\left(\frac{k}{M}\right) - M\int_{k/M}^{(k+1)/M} f(t)\,dt\right)\right|$$
$$\le \frac{1}{M}\sum_{k=0}^{M-1}\left|f\left(\frac{k}{M}\right) - M\int_{k/M}^{(k+1)/M} f(t)\,dt\right|,$$

so it suffices to show that there are few values of k such that $\left|f\left(\frac{k}{M}\right) - M\int_{k/M}^{(k+1)/M} f(t)\,dt\right|$ is large.

For any integer $0 \le k < M$, either $|f\left(\frac{k}{M}\right) - M\int_{k/M}^{(k+1)/M} f(t)\,dt| \le \frac{\varepsilon}{2}$, or there is some $c \in [\frac{k}{M}, \frac{k+1}{M}]$ such that $|f\left(\frac{k}{M}\right) - \phi(\sigma_c; b)| > \frac{\varepsilon}{2}$. By the choice of N, there are at most N values of k such that the latter case holds, for which we can bound $\left|\phi(a_{k/M}; b) - M\int_{k/M}^{(k+1)/M} f(t)\,dt\right| \le 1$. Thus $\sum_{k=0}^{M-1}\left|f\left(\frac{k}{M}\right) - M\int_{k/M}^{(k+1)/M} f(t)\,dt\right|$ is bounded by $M\frac{\varepsilon}{2}$ for most intervals plus N for exceptional intervals, yielding

$$\frac{1}{M}\sum_{k=0}^{M-1}\left|f\left(\frac{k}{M}\right) - \int_{k/M}^{(k+1)/M} f(t)\,dt\right| \le \frac{1}{M}\left(\frac{M\varepsilon}{2} + N\right) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

as desired. □

Lemma 4.3.17. *If $\sigma \in \mathfrak{M}_{[0,1]}(\mathcal{U})$ is an indiscernible segment, then μ_σ is generically stable.*

Proof. Specifically, we show that μ_I is fam in I. Fix a definable predicate $\phi(x; y)$ and $\varepsilon > 0$. By Lemma 4.3.14, for all $b \in \mathcal{U}^y$, there is some N such that there is no sequence $0 \le t_1 < t_1' < \cdots < t_N < t_N' \le 1$ with $|\phi(a_{t_i}; b) - \phi(a_{t_i'}; b)| > \frac{\varepsilon}{2}$ for all $1 \le i \le N$. By compactness, we may choose some N that will work simultaneously for all $b \in \mathcal{U}^y$. Now fix an integer $M > \frac{2N}{\varepsilon}$, and let $A = \{a_{k/M} : 0 \leq k < M\}$. We can apply Lemma 4.3.16 to the family $F = \{f_b(t) : b \in \mathcal{U}^y\}$ defined by $f_b(t) = \phi(a_t; b)$, and find that $\left\{\frac{k}{M} : 0 \leq k < M\right\}$ is a ε-approximation to F with respect to Lebesgue measure. Then as $\int_{S_x(\mathcal{U})} \phi(a_t; b)\, d\mu_I = \int_0^1 f_b(t)\, dt$, we see that A is a ε-approximation to $\{\phi(a_t; b) : b \in \mathcal{U}^y\}$ with respect to μ_I, as desired. □

4.4 Weak Orthogonality and Regularity

In this section, we will characterize weak orthogonality of measures with several Szemerédi-style regularity properties. In the next section, we will use uniform versions of these properties to characterize distality. In order to best explain our techniques and choices, we will first prove some versions of NIP regularity. NIP regularity theorems already exist for classical logic, for instance in [CS21], and indeed for continuous logic, in the form of [CT20, Theorem 6.6], which proves a regularity lemma for real-valued definable predicates in a generalization of NIP structures. However, we find that NIP regularity is the correct setting to first develop the definitions we will need for distal regularity, namely definable and constructible regularity partitions.

4.4.1 NIP Regularity

NIP regularity is a consequence of the ability to approximate definable predicates relative to generically stable measures.

In order to understand it, we must first understand how the classic partitioning into ϕ-types over finite sets works in continuous logic. As even for finite parameter sets B, the set $S_\phi(B)$ of consistent ϕ-types $\mathrm{tp}_\phi(x; y)(a/B)$ is usually infinite in continuous logic, we have to look at partitions where the type $\mathrm{tp}_\phi(x; y)(a/B)$ varies by at most ε on the support of each piece of the partition. Even considering partitions may require us to sacrifice definability of the pieces in continuous logic, leaving us with two options. We can either look at partitions of unity on $S_x(M)$ into definable predicates, get an actual partition, but settle for Borel sets of uniform low complexity.

Definition 4.4.1. Inspired by terminology from algebraic geometry, call a subset of $S_x(A)$ *constructible* when it is a finite boolean combination of closed sets. We will also refer to the indicator functions of constructible sets in type spaces, or their restrictions to models, *constructible predicates.*

If P is a finite partition of unity on M^x, we say that a function $\psi(x;z) : S_{xz}(M) \to [0,1]$ *defines* P when for each piece $\pi \in P$, there is $d \in M^z$ such that π is the support of $\psi(x;d)$. If $\psi(x;z)$ is a definable predicate (that is, continuous), then we call P *definable*, and if $\psi(x;z)$ is a constructible predicate, we call P *constructible*, and as its pieces are all $\{0,1\}$-valued, we may identify P with the partition into the supports of its pieces. When different partitions of unity are definable by the same ψ, we say that they are *uniformly* definable/constructible.

If P is a partition of unity on $M^{x_1 \dots x_n}$ defined by a definable or constructible predicate of the form $\prod_{i=1}^n \psi_i(x_i; z_i)$, then we call P *rectangular.*

If $P_1, \dots, P_n$ are partitions of unity on M^{x_i}, then let $\otimes_{i=1}^n P_i$ denote the partition on $M^{x_1 \dots x_n}$ given by $(\prod_{i=1}^n \pi_i(x_i) : \pi_1 \in P_1, \dots, \pi_n \in P_n)$. We call such a partition a *grid partition.*

If P is a partition of unity on $S_x(M)$ and $B \subseteq M^y$, we call P a (ϕ, ε)*-partition* over B when for each $\pi(x) \in P$, if $a_1, a_2 \in M^x$ both satisfy $\pi(a_i) > 0$, then $|\phi(a_1; b) - \phi(a_2; b)| \leq \varepsilon$ for all $b \in B$.

In the classical case, there is a unique minimal such partition whose size is bounded by the VC-dimension of ϕ. In this lemma, we will show that a suitable partition exists with size given by a covering number bound.

Lemma 4.4.2. *Let $\phi(x;y;w)$ be a definable predicate. Then for any $\varepsilon > 0$ and $n \in \mathbb{N}$, there is a definable predicate $\psi(x;z)$ such that for any $B \subset M^y$ with $|B| \leq n$, and any $c \in M^w$, there is a set $D \subset M^z$ with $|D| \leq \mathcal{N}_{\phi(x;y),0.49\varepsilon}(n)$ such that $(\psi(x;d) : d \in D)$ forms a (ϕ, ε)-partition over B, and ψ is a continuous combination of instances of ϕ over D.*

We may alternately choose ψ to be constructible with the slightly better bound $|D| \leq \mathcal{N}_{\phi(x;y),\varepsilon/2}(n)$, although ψ will no longer be a continuous combination of instances of ϕ.

Proof. Fix c and write $\phi(x;y) = \phi(x;y;w)$, and we will see that the resulting formula $\psi(x;z)$ is constructed from $\phi(x;y;w)$ in a uniform way, with the parameter c reoccuring in the parameters $d \in D$.

For ease of notation, let $m = \mathcal{N}_{\phi(x;y),0.49\varepsilon}(n)$. Define the predicate

$$\theta(x;x',y_1,\ldots,y_n) = \max_{1\leq i\leq n} \frac{\varepsilon}{2} \dot{-} |\phi(x;y_i) - \phi(x';y_i)|,$$

so that for any $a, a' \in M^x, b_1, \ldots, b_m \in M^y$, $\theta(a;a',b_1,\ldots,b_m) > 0$ if and only if $|\phi(a;b_i) - \phi(a';b_i)| < \frac{\varepsilon}{2}$ for all $1 \leq i \leq n$. Similarly, define

$$\psi(x;x_0,\ldots,x_m,y_1,\ldots,y_n) = \frac{\theta(x;x_0,y_1,\ldots,y_n)}{\sum_{i=1}^m \theta(x;x_i,y_1,\ldots,y_n)}.$$

Now if $B \subseteq M^y$ has $|B| \leq n$, express $B = \{b_1,\ldots,b_n\}$, and let $A = \{a_1,\ldots,a_m\}$ be a set such that for each $a \in M^x$, there exists $1 \leq i \leq m$ such that for all $b \in B$, $|\phi(a;b) - \phi(a_i;b)| \leq 0.49\varepsilon$. Then $\{\psi(x;a,a_1,\ldots,a_m,b_1,\ldots,b_n) : a \in A\}$ constitutes a partition of unity, and on each piece $\psi(x;a,a_1,\ldots,a_m,b_1,\ldots,b_n) > 0$, for each $b \in B$, $\phi(x;b)$ varies from $\phi(a;b)$ by at most $\frac{\varepsilon}{2}$, so overall, $\phi(x;b)$ varies by at most ε.

If instead we wish ψ to be constructible, then we instead let $m = \mathcal{N}_{\phi(x;y),\varepsilon/2}(n)$, and we can let $A = \{a_1,\ldots,a_m\}$ be a set such that for each $a \in M^x$, there exists $1 \leq i \leq m$ such that for all $b \in B$, $|\phi(a;b) - \phi(a_i;b)| \leq \frac{\varepsilon}{2}$. Then the sets $X_i = \{|\phi(x;b) - \phi(a_i;b)| \leq \frac{\varepsilon}{2}\}$ cover $S_x(M)$, so we can use standard coding tricks to let $\psi(x;z)$ define any of the sets $X_i \setminus \bigcup_{j<i} X_i$. $\square$

Lemma 4.4.3 (See [CS21, Proposition 2.18])**.** *Let M be an NIP structure, let $\phi(x;y;w)$ be a definable predicate, let $\varepsilon > 0$, and let $\mathrm{vc}_{\varepsilon/12}(\phi(x;y)) \leq d$. Then for any generically stable measure $\mu \in \mathfrak{M}_x(M)$ and any $c \in M^w$, there is a set $A \subseteq M^x$ of size $O(\frac{d}{\varepsilon^2} \ln \frac{d}{\varepsilon})$ such that if*

P is a $(\phi^*(y;x), \varepsilon)$-partition over A, then for $b, b' \in M^y$ in the support of the same piece of P, $\int_{S_x(M)} |\phi(x;b;c) - \phi(x;b';c)|\, d\mu \leq \varepsilon$.

Proof. Fix c and write $\phi(x;y) = \phi(x;y;w)$, and the resulting predicate $\psi(y;z)$ will be constructed from $\phi(x;y;w)$ in a uniform way, based on the predicate defined in Lemma 4.4.2.

Let $\chi(x;y,y') = |\phi(x;y) - \phi(x';y)|$, and using Corollary 4.3.7, there is $n = O(\frac{d}{\varepsilon^2}\ln\frac{d}{\varepsilon})$ such that for each μ, there is a $\frac{\varepsilon}{2}$-approximation with respect to μ of size at most n. Now for every μ, let $A = \{a_1, \ldots, a_n\}$ be a $\frac{\varepsilon}{2}$-approximation to $\chi(x;y,y')$ with respect to μ, and let P be a $(\phi^*(y;x), \varepsilon)$-partition over A. If $\pi(x) \in P$ and $b_1, b_2 \in M^y$ both satisfy $\pi(x) > 0$, then for all $a \in A$, $|\phi(a;b_1) - \phi(a;b_2)| \leq \frac{\varepsilon}{2}$, so $\chi(a;b_1,b_2) \leq \frac{\varepsilon}{2}$, and thus by the $\frac{\varepsilon}{2}$-approximation definition of A,

$$\int_{S_x(M)} |\phi(x;b_1) - \phi(x;b_2)|\, d\mu = \int_{S_x(M)} \chi(x;b_1,b_2)\, d\nu \leq \frac{1}{n}\sum_{i=1}^{n} \chi(a_i;b_1,b_2) + \frac{\varepsilon}{2} \leq \varepsilon.$$

□

Lemma 4.4.4 (See [CS21, Theorem 2.19])**.** *Assume T is NIP and $M \models T$. Fix a definable predicate $\phi(x_1, \ldots, x_n; w)$ and $\varepsilon > 0$. Then there is a definable predicate $\theta(x_1, \ldots, x_n; z)$ of the form $\sum_{j=1}^{m}\prod_{i=1}^{n}\theta_{ij}(x_i;z_i)$ with m depending only on ϕ and ε such that if $\mu_i \in \mathfrak{M}_{x_i}(M)$ are measures with μ_i generically stable for $i < n$ and $c \in M^w$, then there is $d \in M^z$ such that*

$$\int_{S_{x_1\ldots x_n}(M)} |\phi(x_1, \ldots, x_n; c) - \theta(x_1, \ldots, x_n; d)|\; d\mu_1 \otimes \cdots \otimes \mu_n \leq \varepsilon.$$

Proof. As before, fix c and write $\phi(x_1, \ldots, x_n) = \phi(x_1, \ldots, x_n; c)$, and the resulting predicates will be constructed from $\phi(x;y;w)$ in a uniform way, based on the predicate defined in Lemma 4.4.3.

We start by proving the two-dimensional case. Let $\phi(x;y)$ be an M-definable predicate and let $\varepsilon > 0$. Let $\psi(y;z), k$ be as given in Lemma 4.4.3, and let m be the upper bound on the size of the resulting partition of unity. Then we will show that for any measures $\mu \in \mathfrak{M}_x(M), \nu \in \mathfrak{M}_y(M)$ with μ generically stable, there are parameters $b_1, \dots, b_m \in M^y, d_1, \dots, d_m \in M^z$ such that

$$\int_{S_{xy}(M)} \left| \phi(x; y) - \sum_{i=1}^{m} \phi(x; b_i)\psi(y; d_i) \right| d\mu \otimes \nu \leq \varepsilon.$$

Given μ, ν, let $D = \{d_1, \dots, d_m\}$ be such that $(\psi(y; d_i) : 1 \leq i \leq m)$ is the partition of unity given by Lemma 4.4.3, and for each $1 \leq i \leq m$, let $b_i \in M^y$ be such that $\psi(b_i; d_i) > 0$. Then for all $b \in M^y$, we calculate that

$$\begin{aligned}
&\int_{S_x(M)} \left| \phi(x; b) - \sum_{i=1}^{m} \phi(x; b_i)\psi(b; d_i) \right| d\mu \\
=&\int_{S_x(M)} \sum_{i=1}^{m} |\phi(x; b)\psi(b; d_i) - \phi(x; b_i)\psi(b; d_i)| \, d\mu \\
\leq&\sum_{i=1}^{m} \psi(a; d_i) \int_{S_x(M)} |\phi(x; b) - \phi(x; b_i)| \, d\mu \\
\leq&\sum_{i:\psi(b;d_i)>0} \psi(b; d_i) \int_{S_x(M)} |\phi(x; b) - \phi(x; b_i)| \, d\mu \\
\leq&\varepsilon,
\end{aligned}$$

as for each i with $\psi(b; d_i) > 0$, $\int_{S_x(M)} |\phi(x; b) - \phi(x; b_i)| \, d\mu \leq \varepsilon$ by assumption. Thus also integrating over y, we find that $\int_{S_{xy}(M)} |\phi(x; y) - \sum_{i=1}^{m} \phi(x; b_i)\psi(x; d_i)| \, d\mu \otimes \nu \leq \varepsilon$, finishing the base case.

Now assume this works for all n-ary predicates, and consider $\phi(x_1, \dots, x_{n+1})$. Applying our proof to the repartitioned binary predicate $\phi(x_1, x_2, \dots; x_{n+1})$, we see that there is $\psi(x_{n+1}; z_{n+1})$, and some m_n such that for all generically stable $\mu_1, \dots, \mu_{n+1}$, there are

$(b_1, \ldots, b_{m_n})$ and $(d_1, \ldots, d_{m_n})$ such that

$$\int_{S_{x_1 \ldots x_{n+1}}(M)} \left| \phi(x_1, \ldots, x_{n+1}) - \sum_{i=1}^{m_n} \phi(x_1, \ldots, x_n; b_i)\psi(x_{n+1}; d_i) \right| \, d\mu_1 \otimes \cdots \otimes \mu_{n+1} \leq \frac{\varepsilon}{2}.$$

We now apply the induction hypothesis to each $\phi(x_1, \ldots, x_n; b_i)$, seeing that there is some $\theta(x_1, \ldots, x_n; z)$ that belongs to the tensor product of the algebras of definable predicates on the separate variables $x_1, \ldots, x_n$ such that for all generically stable $\mu_1, \ldots, \mu_n$, and every $b \in M^{x_{n+1}}$, there is some $c \in M^z$ such that

$$\int_{S_{x_1 \ldots x_n}(M)} |\phi(x_1, \ldots, x_n, b) - \theta(x_1, \ldots, x_n; c)| \, d\mu_1 \otimes \cdots \otimes \mu_n \leq \frac{\varepsilon}{2}.$$

Then for any $\mu_1, \ldots, \mu_{n+1}$, there are $(b_1, \ldots, b_n), (d_1, \ldots, d_n)$ as above, and for each b_i we choose $c_i \in M^z$ as above. Letting ω abbreviate $\mu_1 \otimes \cdots \otimes \mu_{n+1}$, we calculate

$$\begin{aligned}
&\int_{S_{x_1 \ldots x_{n+1}}(M)} \left| \phi(x_1, \ldots, x_{n+1}) - \sum_{i=1}^{m_n} \theta(x_1, \ldots, x_n; c_i)\psi(x_{n+1}; d_i) \right| \, d\omega \\
\leq &\frac{\varepsilon}{2} + \int_{S_{x_1 \ldots x_{n+1}}(M)} \left| \sum_{i=1}^{m_n} \phi(x_1, \ldots, x_n; b_i)\psi(x_{n+1}; d_i) - \sum_{i=1}^{m_n} \theta(x_1, \ldots, x_n; c_i)\psi(x_{n+1}; d_i) \right| \, d\omega \\
\leq &\frac{\varepsilon}{2} + \sum_{i=1}^{m_n} \int_{S_{x_1 \ldots x_{n+1}}(M)} \psi(x_{n+1}; d_i) |\phi(x_1, \ldots, x_n; b_i) - \theta(x_1, \ldots, x_n; c_i)| \, d\omega \\
\leq &\frac{\varepsilon}{2} + \frac{\varepsilon}{2} \int_{S_{x_1}(M)} \sum_{i=1}^{m_n} \psi(x_1; d_i) \, d\mu_1 \\
= &\varepsilon.
\end{aligned}$$

□

We now define the precise kind of regularity partition that NIP allows us to find, as well as homogeneous tuples, which will be essential for distal regularity.

Definition 4.4.5. A (ε, δ)-NIP regularity partition for $\phi(x_1, \ldots, x_n)$ with respect to measures $\mu_i \in \mathfrak{M}_{x_i}(M)$ is a grid partition defined by $\psi(x;z) = \prod_{i=1}^n \psi_i(x_i;z_i)$ over a set D such that there is a subset $D_0 \subset D$ with

$$\sum_{d \in D_0} \int_{S_x(M)} \psi(x;d)\, d\mu_1 \times \cdots \times \mu_n \leq \delta$$

and for each $d \in D \setminus D_0$, there are is a value r_d such that

$$\int_{S_x(M)} \psi(x;d)|\phi(x) - r_d|\, d\mu_1 \times \cdots \times \mu_n \leq \varepsilon \int_{S_x(M)} \psi(x;d)\, d\mu_1 \times \cdots \times \mu_n.$$

If $\phi(x_1, \ldots, x_n)$ is a definable predicate, $\varepsilon > 0$, and $A_i \subseteq M^{x_i}$ for each i, we say that $(A_1, \ldots, A_n)$ is $(\phi, \varepsilon)-$*homogeneous* when for all $a, a' \in A_1 \times \cdots \times A_n$, $|\phi(a) - \phi(a')| \leq \varepsilon$.

We also say that definable/constructible predicates $\psi_i(x_i)$ are $(\phi, \varepsilon)-$homogeneous when their supports are, or we may say this about their product $\prod_{i=1}^n \psi_i(x_1)$.

Lemma 4.4.6. *For any definable predicate $\theta(x_1, \ldots, x_n; y)$ of the form $\sum_{j=1}^m \prod_{i=1}^n \theta_{ij}(x_i; y_{ij})$, there is a predicate $\psi(x_1, \ldots, x_n; z) = \prod_{i=1}^n \psi_i(x_i; z_i)$ such that for any $b \in M^y$, ψ defines a grid partition of unity such that each piece is $(\theta(x_1, \ldots, x_n; b), \varepsilon)$-homogeneous, and the size of the partition is bounded by a function of m, n. We can choose ψ to be either definable or constructible.*

Proof. Let N be large enough that $m\left(\left(1 + \frac{2}{N}\right)^n - 1\right) \leq \frac{\varepsilon}{2}$. Let $(f_k : 0 \leq k \leq N)$ be a continuous partition of unity on $[0,1]$ such that the support of each f_k lies in the interval $\left(\frac{k-1}{N}, \frac{k+1}{N}\right)$. Then for any $\phi(x)$, $(f_k \circ \phi : 0 \leq k \leq N)$ is a partition of unity on M^x, such that on each piece, ϕ varies by at most $\frac{2}{N}$. By taking products of the partitioning functions, we can find a definable partition of unity P_i on each M^{x_i} such that each $\theta_{ij}(x_i)$ varies by at most $\frac{2}{N}$ on each piece, and set $P = \otimes_{i=1}^n P_i$. Then for each $\pi \in P$, the function $\theta(x_1, \ldots, x_n)$ varies from some value by at most $m\left(\left(1 + \frac{2}{N}\right)^n - 1\right) \leq \frac{\varepsilon}{2}$, and thus vary in total by at most ε, on the support of π. Each partition P_i is a refinement of m partitions of unity into $N+1$ pieces, so $|P_i| \leq (N+1)^m$, and thus $|P| \leq (N+1)^{mn}$.

If instead we desire a constructible partition, then we can let f_k be indicator functions of the intervals $\left[\frac{k}{N}, \frac{k+1}{N}\right)$. $\square$

Theorem 4.4.7 (See [CS21, Theorem 3.3])**.** *Assume T is NIP and $M \models T$. Let $\phi(x_1, \dots, x_n)$ be a definable predicate and let $\varepsilon, \delta > 0$. Then there is a predicate $\psi(x_1, \dots, x_n; z) = \prod_{i=1}^n \psi_i(x_i; z_i)$ such that for all measures $\mu_i \in \mathfrak{M}_{x_i}(M)$ with μ_i generically stable for $i < n$, ψ defines a (ε, δ)-NIP regularity partition of unity for $\phi(x_1, \dots, x_n)$ with respect to the μ_is, with the size of the partition depending only on $\phi, \varepsilon, \delta$. We can choose ψ to be either definable or constructible.*

Proof. By Lemma 4.4.4, we know that there is a definable predicate $\theta(x_1, \dots, x_n; z)$ of the form $\sum_{j=1}^m \prod_{i=1}^n \theta_{ij}(x_i; z_i)$ such that if $\mu_i \in \mathfrak{M}_{x_i}(M)$ are measures with μ_i generically stable for $i < n$ and $c \in M^w$, then there is $d \in M^z$ such that

$$\int_{S_{x_1 \dots x_n}(M)} |\phi(x_1, \dots, x_n; c) - \theta(x_1, \dots, x_n; d)| \, d\mu_1 \otimes \dots \otimes \mu_n \leq \delta^2.$$

We now apply Lemma 4.4.6 to θ, finding that there is some predicate ψ that for any $d \in M^z$, defines a grid partition of unity where each piece is $(\theta(x_1, \dots, x_n; d), \varepsilon)$-homogeneous. Then for any appropriate measures μ_i and $c \in M^w$, denoting $\omega = \mu_1 \otimes \dots \otimes \mu_n$, we find $d \in M^z$ as before, and let P be the grid partition of unity defined by ψ with homogeneous pieces. Now define $e : P \to [0,1]$ by

$$e(\pi) = \frac{\int_{S_x(M)} |\phi(x; c) - \theta(x; d)| \, \pi(x) \, d\omega}{\int_{S_x(M)} \pi(x) \, d\omega}.$$

Giving P the measure $\mu(\{\pi\}) = \int_{S_x(M)} \pi(x) \, d\omega$, we see that

$$\int_P e \, d\mu = \sum_{\pi \in P} \int_{S_x(M)} |\phi(x; c) - \theta(x; d)| \, \pi(x) \, d\omega \leq \delta^2,$$

so by Markov's inequality, the measure of all $\pi \in P$ with $\int_{S_x(M)} |\phi(x; c) - \theta(x; d)| \, \pi(x) \, d\omega > \delta$

is at most δ. If that set is $P_0 \subseteq P$, then $\sum_{\pi \in P_0} \int_{S_x(M)} \pi(x)\, d\omega \leq \delta$, and for each $\pi \in P \setminus P_0$, we find that $\int_{S_x(M)} |\phi(x;c) - \theta(x;d)|\, \pi(x)\, d\omega \leq \delta$.

Now suppose $\pi \in P \setminus P_0$. By (θ, ε)-homogeneity, there is some r_π be such that on the support of π, $|\theta(x;d) - r_\pi| \leq \frac{\varepsilon}{2}$, and thus $|\phi(x;c) - r_\pi| \dot{-} \frac{\varepsilon}{2} \leq |\phi(x;c) - \theta(x;d)|$, so $\int_{S_x(M)} |\phi(x;c) - r_\pi| \dot{-} \frac{\varepsilon}{2} \pi(x)\, d\omega \leq \delta$, and thus this partition is a (ε, δ)-NIP regularity partition. $\square$

4.4.2 Weak Orthogonality and Strong Erdős-Hajnal

In this subsection, we introduce continuous versions of the notions of regularity that characterize weakly orthogonal measures, and thus distality.

Definition 4.4.8. We say that a predicate $\phi(x_1, \ldots, x_n)$ has the *ε-strong Erdős-Hajnal property* (or ε-SEH) when there exists δ such that for any finite sets $A_i \subseteq M^{x_i}$, there are subsets $B_i \subseteq A_i$ such that $|B_i| \geq \delta |A_i|$ and $(B_1, \ldots, B_n)$ is (ϕ, ε)-homogeneous.

A predicate ϕ has the *definable* ε-SEH with respect to measures $\mu_1, \ldots, \mu_n$ when there are predicates $\psi_i(x_i; z_i)$ and $\delta > 0$ such that there are parameters $d_i \in M^{z_i}$ such that for each i, $\int_{S_{x_i}(M)} \psi_i(x_i; d_i)\, d\mu_i \geq \delta$ and the supports of $\psi_i(x_i; d_i)$ are (ϕ, ε)-homogeneous. If for some $\phi(x_1, \ldots, x_n; y)$, some $\varepsilon > 0$, and some class of tuples of Keisler measures, $\phi(x_1, \ldots, x_n; b)$ has the definable ε-SEH with respect to all tuples of measures in that class, and the same $\delta > 0$ and predicates $\psi_i(x_i; z_i)$ can be used in each case, then we say that ϕ has the *uniformly definable* ε-SEH with respect to that class of tuples of measures.

In classical logic, the definable strong Erdős-Hajnal property implies a Szemerédi-style regularity lemma, as in [CS18, Section 5], We will now define a real-valued version of this regularity property.

Definition 4.4.9. If P is a rectangular partition of unity on $M^{x_1 \ldots x_n}$, then we call P a

(ε, δ)-*distal regularity partition* for ϕ with respect to measures $\mu_i \in \mathfrak{M}_{x_i}(M)$ when

$$\sum_{\pi_1,\ldots,\pi_n} \prod_{i=1} \int_{S_{x_i}(M)} \pi_i(x_i)\, d\mu_i \leq \delta,$$

where the sum ranges over all $(\pi_1, \ldots, \pi_n)$ with $\prod_{i=1} \pi_i \in P$ that are not (ϕ, ε)-homogeneous.

In a series of lemmas, we will show that a fixed tuple of measures $\mu_1, \ldots, \mu_n$ is weakly orthogonal if any of several equivalent regularity properties hold. We will prove the implications of this equivalence with enough detail to later show that if any of these properties holds in a uniformly definable way across all generically stable Keisler measures, then the theory is distal.

Recall that by *a product measure of continuous localizations of* $\mu_1, \ldots, \mu_n$, we mean a measure ω such that there exist M-definable predicates $\theta_i(x_i)$ such that for all M-definable predicates $\phi_i(x_i)$,

$$\int_{S_{x_1\ldots x_n}(M)} \prod_{i=1}^{n} \phi_i(x_i)\, d\omega = \prod_{i=1}^{n} \int_{S_{x_i}(M)} \phi_i(x_i)\theta_i(x_i)\, d\mu_i.$$

Theorem 4.4.10. *Let $\mu_1 \in \mathfrak{M}_{x_1}(M), \ldots, \mu_n \in \mathfrak{M}_{x_n}(M)$. The following are equivalent:*

- *The measures $\mu_1, \ldots, \mu_n$ are weakly orthogonal.*
- *For each M-definable predicate $\phi(x_1, \ldots, x_n)$ and each $\varepsilon, \delta > 0$, there is some C such that ϕ admits a definable (ε, δ)-distal regularity partition*
- *For each M-definable predicate $\phi(x_1, \ldots, x_n)$ and each $\varepsilon, \delta > 0$, there is some C such that ϕ admits a constructible (ε, δ)-distal regularity partition*
- *For each M-definable predicate $\phi(x_1, \ldots, x_n)$ and each $\varepsilon > \gamma \geq 0$, there is some $\delta > 0$ such that for any product measure ω of continuous localizations of $\mu_1, \ldots, \mu_n$, if $\int_{S_{x_1\ldots x_n}(M)} \phi\, d\omega \geq \varepsilon$, then there are M-definable predicates $\psi_i(x_i)$ such that whenever $\psi_i(a_i) > 0$ for each i, $\phi(a_1, \ldots, a_n) \geq \gamma$, and $\int_{S_{x_i}(M)} \psi_i(x_i)\, d\mu_i \geq \delta$ for each i.*

- *For each M-definable predicate $\phi(x_1, \dots, x_n)$ and each $\varepsilon > 0$, ϕ has the definable ε-SEH with respect to any continuous localizations of $\mu_1, \dots, \mu_n$.*

Furthermore, if these hold, then the (ε, δ)-distal regularity partitions can be chosen to be grid partitions of size $O(\delta^{-C})$ for some constant C depending on $\phi, \varepsilon, \mu_1, \dots, \mu_n$.

Proof. We will show that weak orthogonality is equivalent to the existence of distal regularity partitions, and then we will relate definable distal regularity partitions to both strong Erdős-Hajnal statements. We start by restating the result of Corollary 4.2.21 about weakly orthogonal Keisler measures in terms of partitions of unity.

Lemma 4.4.11. *For $1 \leq i \leq n$, let $\mu_i \in \mathfrak{M}_{x_i}(M)$ be Keisler measures. Let $\phi(x_1, \dots, x_n)$ and $\psi^{\pm}(x_1, \dots, x_n)$ be M-definable predicates such that*

- *The predicates $\psi^{\pm}(x_1, \dots, x_n)$ are each of the form $\sum_{j=1}^{m} \prod_{i=1}^{n} \theta_{ij}^{\pm}(x_i)$*
- *For all $(x_1, \dots, x_n)$, $\psi^{-}(x_1, \dots, x_n) \leq \phi(x_1, \dots, x_n) \leq \psi^{+}(x_1, \dots, x_n)$.*
- *For any product measure ω of $\mu_1, \dots, \mu_n$, $\int_{S_{x_1 \dots x_n}(M)} (\psi^{+} - \psi^{-})\, d\omega \leq \varepsilon$.*

Then there is a grid partition of unity $P = \otimes_{i=1}^{n} P_i$ on $M^{x_1 \dots x_n}$, which can be chosen to either be definable or constructible, such that if for each tuple $\pi \in P$, we set $r_\pi^{-} = \inf_{x : \pi(x) > 0} \phi(x)$ and $r_\pi^{+} = \sup_{x : \pi(x) > 0} \phi(x)$, and then define

$$\chi^{\pm}(x_1, \dots, x_n) = \sum_{\pi \in P} r_\pi^{\pm} \pi(x_1, \dots, x_n),$$

then

- *For all $(x_1, \dots, x_n)$, $\chi^{-}(x_1, \dots, x_n) \leq \phi(x_1, \dots, x_n) \leq \chi^{+}(x_1, \dots, x_n)$.*
- *For any product measure ω of $\mu_1, \dots, \mu_n$, $\int_{S_{x_1 \dots x_n}(M)} (\chi^{+} - \chi^{-})\, d\omega \leq 2\varepsilon$.*

Furthermore, the definition of P depends only on the predicates ϕ, ψ^+, ψ^-, and not the parameters used in their definitions.

Proof. By Lemma 4.4.6, there are predicates $\pi^{\pm}(x_1, \ldots, x_n; z) = \prod_{i=1}^n \pi_i^{\pm}(x_i; z_i)$ such that $\pi^{\pm}$ defines a rectangular partition of unity such that each piece is $(\psi^{\pm}(x_1, \ldots, x_n), \frac{\varepsilon}{2})$-homogeneous. Thus if we let $\pi_i(x_i; z_i^+, z_i^-) = \pi_i^+(x_i; z_i^+)\pi_i^-(x_i; z_i^-)$, then $\pi(x_1, \ldots, x_n; z') = \prod_{i=1}^n \pi(x_i; z_i^+, z_i^-)$ defines a refinement of the two partitions of unity, so that each piece is $(\psi^+(x_1, \ldots, x_n), \frac{\varepsilon}{2})$-homogeneous and $(\psi^+(x_1, \ldots, x_n), \frac{\varepsilon}{2})$-homogeneous.

Then for each π, we let $r_\pi^+ = \sup \psi^+(x_1, \ldots, x_n)$ where the sup ranges over the support of π, and let $r_\pi^- = \inf \psi^-(x_1, \ldots, x_n)$. We then let $\chi^{\pm}(x_1, \ldots, x_n) = \sum_{\pi \in P} r_\pi^{\pm}\pi$. By symmetry, it suffices to show that $\psi^+ \leq \chi^+$ and for any product measure ω, $\int_{S_{x_1 \ldots x_n}(M)} (\chi^+ - \psi^+)\, d\omega \leq \frac{\varepsilon}{2}$. The integral fact will follow by showing that $\psi^+ \leq \chi^+ \leq \psi^+ + \frac{\varepsilon}{2}$. Let $a \in M^{x_1 \ldots x_n}$. Then for every $\pi \in P$ with $\pi(a) > 0$, ψ^+ varies by at most $\frac{\varepsilon}{2}$ on the set containing a on which r_π^+ is a supremum, so we have $\psi^+(a) \leq r_\pi^+ \leq \psi^+(a) + \frac{\varepsilon}{2}$. As $\chi^+(a)$ is a convex combination of such numbers, $\psi^+(a) \leq \chi^+(a) \leq \psi^+(a) + \frac{\varepsilon}{2}$. $\square$

Lemma 4.4.12. *Let $x_1, \ldots, x_n$ be variable tuples, and let $\mu_i \in \mathfrak{M}_{x_i}(M)$ for each i. Then the measures μ_i are weakly orthogonal if and only if for every M-definable predicate $\phi(x_1, \ldots, x_n)$ and every $\delta > 0, \varepsilon > 0$, there exists a (ε, δ)-distal regularity partition P for ϕ with respect to the measures μ_i.*

Proof. First assume that μ_i are weakly orthogonal. By Corollary 4.2.21 put together with Lemma 4.4.11, we can find a rectangular partition of unity P such that if we let

$$r_\pi^+ = \sup_{x:\pi(x)>0} \phi(x),$$

$$r_\pi^- = \inf_{x:\pi(x)>0} \phi(x),$$

$$\chi^{\pm} = \sum_{\pi \in P} r_\pi^{\pm}\pi,$$

then $\chi^-(x) \leq \phi(x) \leq \chi^+(x)$ and for any product measure ω of $\mu_1, \ldots, \mu_n$, we can bound $\int_{S_x(M)}(\chi^+ - \chi^-)\, d\omega \leq \delta\varepsilon$.

We clearly see for some $\pi = \prod_{i=1}^n \pi_i \in P$, the tuple $(\pi_1, \ldots, \pi_n)$ is (ϕ, ε)-homogeneous if and only if $r_\pi^+ - r_\pi^- \leq \varepsilon$.

We calculate that $\int_{S_{x_1 \ldots x_n}(M)}(\chi^+ - \chi^-)\, d\omega = \sum_{\pi \in P}(r_\pi^+ - r_\pi^-) \int_{S_{x_1 \ldots x_n}(M)} \pi\, d\omega \leq \delta\varepsilon$, so placing a measure on the finite set P by giving π the measure $\int_{S_{x_1 \ldots x_n}(M)} \pi\, d\omega$ and applying Markov's inequality, we see that the measure of the non-homogeneous predicates π is at most δ, so this is a $(\varepsilon, \delta)-$distal regularity partition.

On the other hand, if P is a $(\varepsilon, \delta)-$distal regularity partition, then as before, set the notations $r_\pi^+ = \sup_{x:\pi(x)>0} \phi(x)$, $r_\pi^- = \inf_{x:\pi(x)>0} \phi(x)$ and $\chi^\pm = \sum_{\pi \in P} r_\pi^\pm \pi$. As before, on the homogeneous pieces, $r_\pi^+ - r_\pi^- \leq \varepsilon$ and for any product measure ω,

$$\int_{S_{x_1 \ldots x_n}(M)}(\chi^+ - \chi^-)\, d\omega = \sum_{\pi \in P}(r_\pi^+ - r_\pi^-) \int_{S_{x_1 \ldots x_n}(M)} \pi\, d\omega,$$

where the sum over the homogeneous pieces is at most ε, and the sum over the non-homogeneous pieces is at most δ, so the $\int_{S_{x_1 \ldots x_n}(M)}(\chi^+ - \chi^-)\, d\omega \leq \delta + \varepsilon$. As the integrals of $\chi^\pm$ do not depend on the choice of ω, this means that for two different product measures ω_1, ω_2, we have $|\int_{S_{x_1 \ldots x_n}(M)} \phi\, d\omega_1 - \int_{S_{x_1 \ldots x_n}(M)} \phi\, d\omega_2| \leq \delta + \varepsilon$. Thus if such partitions exist for $\delta, \varepsilon > 0$ arbitrarily small, $\omega_1 = \omega_2$ and the measures μ_i are weakly orthogonal. $\square$

We now show that strong Erdős-Hajnal behavior on all continous localizations implies a distal regularity partition in a uniform way.

Lemma 4.4.13. *Let $\phi(x_1, \ldots, x_n)$ be an M-definable predicate, let $\varepsilon, \delta > 0$, and let $\mu_i \in \mathfrak{M}_{x_i}(M)$ be a Keisler measure for each i. Suppose that for any Borel localizations $\nu_1, \ldots, \nu_n$ of $\mu_1, \ldots, \mu_n$, there exist M-definable predicates $\psi_i(x_i; d_i)$ such that the supports of $\psi_i(x_i)$ are (ϕ, ε)-homogeneous and for each i, $\int_{S_{x_i}(M)} \psi_i(x_i; d_i)\, d\nu_i \geq \delta$. Then for any $\gamma > 0$, there exists a (ε, γ)-distal regularity partition for ϕ with respect to $\mu_1, \ldots, \mu_n$ definable over D, with*

$|D| = O(\gamma^{-C})$ for some C depending only on δ. Furthermore, if the predicates $\psi_i(x_i; z_i)$ can be chosen uniformly for all continuous localizations $\nu_1, \ldots, \nu_n$ of $\mu_1, \ldots, \mu_n$, then the distal regularity partition can be defined by a predicate which is a continuous combination of the ψ_is depending only on δ, γ.

Proof. First, we note that if $\int_{S_{x_i}(M)} \psi_i(x_i; d_i)\, d\nu_i \geq \delta$, then $\nu_i\left(\psi_i(x_i; d_i) \geq \frac{\delta}{2}\right) \geq \frac{\delta}{2}$.

We will actually find a set $D \subset M^{z_1 \ldots z_n}$ such that for each $(d_1, \ldots, d_n) \in D$, the supports $\psi_i(x_i; d_i)$ are (ϕ, ε)-homogeneous, and

$$\mu_1 \times \mu_n \left(\bigcup_{d_1, \ldots, d_n} \psi_i(x_i; d_i) \geq \frac{\delta}{2} \right) \geq 1 - \gamma.$$

Once we have found this, we use Lemma 4.4.2 to partition each M^{x_i} into a partition of unity P_i such that on the support of each piece, for each $d_i \in D_i$, either $\psi_i(x_i; d_i) > 0$ or $\psi_i(x_i; d_i) < \frac{\delta}{2}$. Then if $\pi_i \in P_i$ for each i, either there is some $(d_1, \ldots, d_n) \in D$ such that the support of each $\pi_i(x_i)$ is contained in the support of $\psi_i(x_i; d_i)$, and this tuple of supports is (ϕ, ε)-homogeneous, or the product of the supports of π_i are disjoint from the set of measure $1 - \gamma$ mentioned before. Thus the integrals of the non-(ϕ, ε)-homogeneous pieces add up to at most γ.

We will construct D_m with $\mu_1 \times \cdots \times \mu_n \left(\bigcup_{d_1, \ldots, d_n} \left[\psi_i(x_i; d_i) \geq \frac{\delta}{2}\right] \right) \geq 1 - \left(1 - \left(\frac{\delta}{2}\right)^n\right)^m$ for all m recursively, and iterate until $\left(1 - \left(\frac{\delta}{2}\right)^n\right)^m \leq \gamma$. We simultaneously construct a rectangular constructible partition P_m on $M^{x_1 \ldots x_n}$, such that $\bigcup_{d_1, \ldots, d_n} \left[\psi_i(x_i; d_i) \geq \frac{\delta}{2}\right]$ is a union of pieces of P_m. If X_m is the union of all pieces of P_m contained in $\bigcup_{d_1, \ldots, d_n} \left[\psi_i(x_i; d_i) \geq \frac{\delta}{2}\right]$, we will make sure that at each stage, $\mu_1 \times \cdots \times \mu_n(X_m) \geq 1 - \left(1 - \left(\frac{\delta}{2}\right)^n\right)^m$. At each stage, we will ensure that $|D_m|, |P_m| \leq (n+1)^m$.

For $m = 0$, we may use $D = \emptyset$, and let P_0 be the trivial 1-piece partition. Here it is possible that $X_0 = \emptyset$. Assume for induction that we have D_m and P_m such that $\mu_1 \times \cdots \times \mu_n(X_m) \geq 1 - \left(1 - \left(\frac{\delta}{2}\right)^n\right)^m$. Then to form D_{m+1} and P_{m+1}, we will replace each piece

$A_1 \times \cdots \times A_n$ of P_m into at most $n+1$ pieces, and add at most one element to D for each such piece. Let $A_1 \times \cdots \times A_n \in P_m$. If $A_1 \times \cdots \times A_n \subseteq \bigcup_{d_1,\ldots,d_n} \psi_i(x_i; d_i) \geq \frac{\delta}{2}$, then it is already in X_m, and we can leave this piece in P_{m+1}. This way we ensure that $X_m \subseteq X_{m+1}$. If $\prod_{i=1}^n \mu_i(A_i) = 0$, then we will still leave this piece, as it does not affect the measure of X_{m+1}. Otherwise, we can find $d_1, \ldots, d_n$ such that the supports of $\psi_i(x_i; d_i)$ are (ϕ, ε)-homogeneous and if ν_i is the localization of μ_i to A_i, then $\nu_i\left(\psi_i(x_i; d_i) \geq \frac{\delta}{2}\right) \geq \frac{\delta}{2}$. Thus if we add $(d_1, \ldots, d_n)$ to D and replace the piece $A_1 \times A_n$ with the $n+1$ pieces $\prod_{i \leq j}\left(A_i \cap \left[\psi_i(x_i; d_i) \geq \frac{\delta}{2}\right]\right) \times \prod_{i > j}\left(A_i \setminus \left[\psi_i(x_i; d_i) \geq \frac{\delta}{2}\right]\right)$, we find that the one piece $\prod_{i=1}^n \left(A_i \cap \left[\psi_i(x_i; d_i) \geq \frac{\delta}{2}\right]\right)$ which definitely contributes to X_{m+1} has total measure at least $\left(\frac{\delta}{2}\right)^n \prod_{i=1}^n \mu_i(A_i)$. Thus if we do this for all pieces of P_m, the total measure of $M^{x_1 \ldots x_n} \setminus X_{m+1}$ decreases by a factor of at least $\left(\frac{\delta}{2}\right)^n$, meaning that $\mu_1 \times \cdots \times \mu_n(X_{m+1}) \geq 1 - \left(1 - \left(\frac{\delta}{2}\right)^n\right)^{m+1}$. We also find that $|P_{m+1}| \leq (n+1)|P_m| \leq (n+1)^{m+1}$, and $|D_{m+1}| \leq |D_m| + |P_m| \leq 2(n+1)^m \leq (n+1)^{m+1}$.

Now if we choose $M = \left\lceil \frac{\log \gamma}{\log(1-\delta^n)} \right\rceil$, we find that $\mu_1 \times \cdots \times \mu_n(X_M) \geq 1 - \gamma$ as desired. If $C = -\frac{\log(n+1)}{\log(1-\delta^n)}$, then $(n+1)^M \leq (n+1)\gamma^{-C}$, so the number of pieces is $O(\gamma^{-C})$. □

From a distal regularity partition, we can derive a statement about integrals of predicates which is analogous to the density version of the strong Erdős-Hajnal property in [CS18, Section 4].

Lemma 4.4.14. *Let $\mu_i \in \mathfrak{M}_{x_i}(M)$ for $1 \leq i \leq n$, let $\phi(x_1, \ldots, x_n)$ be an M-definable predicate, and let $(\varepsilon, \delta), \psi_1(x_1; z_1), \ldots, \psi_n(x_n; z_n)$ be such that $\prod_{i=1}^n \psi_i(x_i; z_i)$ defines a (ε, δ)-distal regularity partition for ϕ with respect to $\mu_1, \ldots, \mu_n$ of size K.*

Suppose that $\alpha, \beta \geq 0$ are such that $\alpha > \beta + \delta + \varepsilon$. Then if ω is a measure extending $\mu_1 \times \cdots \times \mu_n$ such that $\int_{S_{x_1 \ldots x_n}(M)} \phi \, d\omega \geq \alpha$, then there are some $d_i \in M^{z_i}$ such that if $a = (a_1, \ldots, a_n)$ satisfies $\prod_{i=1}^n \psi_i(a_i; d_i) > 0$, then $\phi(a) \geq \beta$ and for each i,

$$\int_{S_{x_i}(M)} \psi_i(x_i; d_i) \, d\mu_i \geq \frac{\alpha - \beta - \delta - \varepsilon}{K} > 0.$$

Proof. Let $\psi(x;z) = \prod_{i=1}^n \psi_i(x_i;z_i)$, and let D be the set of parameters such that $(\psi(x;d) : d \in D)$ is the distal regularity partition. We will break D into three disjoint sets A, B, C. Let A be the set of all d such that the support of $\psi(x;d)$ is (ϕ, ε)-homogeneous and on that support, $\phi(x) \geq \gamma$. Let B be the set of all other d such that the support of $\psi(x;d)$ is (ϕ, ε)-homogeneous, and let C be the set of all d such that the support is non-homogeneous.

Then by the partition of unity assumption,

$$\sum_{d \in A} \int_{S_x(M)} \phi(x)\psi(x;d)\, d\omega + \sum_{d \in B} \int_{S_x(M)} \phi(x)\psi(x;d)\, d\omega + \sum_{d \in C} \int_{S_x(M)} \phi(x)\psi(x;d)\, d\omega$$
$$= \int_{S_x(M)} \phi(x)\, d\omega \geq \alpha.$$

However, by the assumption of homogeneity, on the support of $\psi(x;d)$ for any $d \in B$, $\phi(x) \leq \beta + \varepsilon$, so

$$\sum_{d \in B} \int_{S_x(M)} \phi(x)\psi(x;d)\, d\omega \leq \beta + \varepsilon,$$

and by the distal regularity assumption,

$$\sum_{d \in C} \int_{S_x(M)} \phi(x)\psi(x;d)\, d\omega \leq \delta,$$

so

$$\sum_{d \in A} \int_{S_x(M)} \phi(x)\psi(x;d)\, d\omega \geq \alpha - \beta - \delta - \varepsilon.$$

As this is the sum over at most K cells, we find that at least one of the terms of this sum is at least $\frac{\alpha-\beta-\delta-\varepsilon}{K}$, so for the larger integral,

$$\int_{S_x(M)} \psi(x;d)\, d\omega = \prod_{i=1}^{N} \int_{S_{x_i}(M)} \psi_i(x_i;d_i)\, d\mu_i \geq \frac{\alpha - \beta - \delta - \varepsilon}{K}.$$

As each term in the product is at most 1, they are all at least $\frac{\alpha-\beta-\delta-\varepsilon}{K}$. □

Going full circle, we can use this density property to imply the Strong Erdős-Hajnal property.

Lemma 4.4.15. *Let* $\mu_i \in \mathfrak{M}_{x_i}(M)$ *for* $1 \leq i \leq n$, *let* $\phi(x_1, \ldots, x_n)$ *be an* M*-definable predicate, let* $s \in \mathbb{N}, \delta > 0, \psi_1(x_1; z_1), \ldots, \psi_n(x_n; z_n)$. *Assume that for all* $0 \leq j \leq s$, *if* ω *is a measure extending* $\mu_1 \times \cdots \times \mu_n$, *and* $\int_{S_{x_1 \ldots x_n}(M)} \frac{1}{s} \dot{-} \left|\phi - \frac{j}{s}\right| \, d\omega \geq \frac{1}{s(s+1)}$, *then there are some* $d_i \in M^{z_i}$ *such that if* $a = (a_1, \ldots, a_n)$ *satisfies* $\prod_{i=1}^n \psi_i(a_i; d_i) > 0$, *then* $\left|\phi - \frac{j}{s}\right| \leq \frac{1}{s}$ *and for each* i,

$$\int_{S_{x_i}(M)} \psi_i(x_i; d_i) \, d\mu_i \geq \delta > 0.$$

Then there are some $d_i \in M^{z_i}$ *such that the supports of* $\psi_i(x_i; d_i)$ *are* $\left(\phi, \frac{2}{s}\right)$*-homogeneous and*

$$\int_{S_{x_i}(M)} \psi_i(x_i; d_i) \, d\mu_i \geq \delta.$$

Proof. For any $r \in [0, 1]$, $\sum_{j=0}^s \frac{1}{s} \dot{-} \left|r - \frac{j}{s}\right| = \frac{1}{s}$. Thus for some $0 \leq j \leq s$,

$$\int_{S_{x_1 \ldots x_n}(M)} \frac{1}{s} \dot{-} \left|\phi - \frac{j}{s}\right| \, d\omega \geq \frac{1}{s(s+1)},$$

and we find $d_i \in M^{z_i}$ such that and for each i, $\int_{S_{x_i}(M)} \psi_i(x_i; d_i) \, d\mu_i \geq \delta > 0$, and on the support of $\prod_{i=1}^n \psi_i(a_i; d_i)$, $\left|\phi - \frac{j}{s}\right| \leq \frac{1}{s}$, so the supports of $\psi_i(x_i; d_i)$ are $\left(\phi, \frac{2}{s}\right)$-homogeneous. $\square$

We now finish the proof of Theorem 4.4.10 by observing that if $\mu_1, \ldots, \mu_n$ are weakly orthogonal, then by Lemma 4.2.22, any continuous localizations are weakly orthogonal, so for any $\phi(x_1, \ldots, x_n)$, distal regularity partitions exist, and by Lemma 4.4.14, the density version of strong Erdős-Hajnal holds for these measures, and thus by Lemma 4.4.15, strong Erdős-Hajnal holds for these measures. If we assume that strong Erdős-Hajnal holds for any continuous localizations of $\mu_1, \ldots, \mu_n$, then by Lemma 4.4.13, a distal regularity partition exists, so the measures are weakly orthogonal. This shows that all of the properties are equivalent to weak orthogonality. Also, Lemma 4.4.13 produces grid partitions of polynomial size. □

4.5 Keisler Measures in Distal Theories

In classical logic, a theory is distal if and only if all generically stable measures are smooth. We prove that this still holds in continuous logic, and show that it is enough to check that all generically stable measures are weakly orthogonal.

Theorem 4.5.1. *The following are equivalent:*

- *The theory T is distal*
- *Every generically stable measure is smooth*
- *All pairs of generically stable measures are weakly orthogonal.*

Proof. First we show that in distal theories, generically stable measures are smooth.

Lemma 4.5.2 (Generalizing [Sim15, Prop. 9.26])**.** *Assume T is distal. Then all generically stable measures are smooth.*

Proof. Assume T is distal, and let $\mu \in \mathfrak{M}_x(\mathcal{U})$ be a generically stable measure, invariant over a small model M. To show that $\mu|_M$ is smooth, fix $M \preceq N$, a predicate $\phi(x)$ with parameters in N, and $\varepsilon > 0$. By [And23b, Theorem 3.5.9], both ϕ and $1 - \phi$ admit strong honest definitions, and thus strong* honest definitions by [And23b, Lemma 3.5.11]. Thus there is an extension $(N, P_M) \preceq (N', P_{M'})$ and a predicate $\psi^+(x)$ (a strong* honest definition for ϕ over M) with parameters in M' such that for $a \in M$, $\psi^+(a) = \phi(a)$ and for all a' in N', $\phi(a') \leq \psi^+(a')$. By applying the same result to $1 - \phi$ and then subtracting from 1, we can find ϕ^-, also with parameters in M', such that for $a \in M$, $\psi^-(a) = \phi(a)$, and for all a' in N', $\phi(a') \geq \psi^-(a')$. To refer to the parameters more easily, we now write $\psi^+(x; d_+)$ and $\psi^-(x; d_-)$, where $d_+, d_- \in M'^z$.

We see that $\psi^+(x;d_+) - \phi(x)$ and $\phi(x) - \psi^-(x;d_-)$, nonnegative everywhere, are both 0 at all tuples in M. Thus by Lemma 4.2.11 and the fact that μ is approximately realized in M, $\int_{S_x(\mathcal{U})} \psi^+(x;d_+) - \phi(x)\, d\mu$ and $\int_{S_x(\mathcal{U})} \phi(x) - \psi^-(x;d_-)\, d\mu$ are both 0. As μ is definable, the function $\int_{S_x(\mathcal{U})} \psi^+(x;z) - \phi(x)\, d\mu$ is continuous from $S_z(N)$ to $\mathbb{R}$. Thus there is some some basic open neighborhood of $\mathrm{tp}(d_+/N)$ such that $\int_{S_x(\mathcal{U})} \psi^+(x;z) - \phi(x)\, d\mu < \varepsilon$ and $\sup_x (\phi(x) - \psi^+(x;z)) < \varepsilon$ on that neighborhood. We may assume that this neighborhood is defined by $\theta(z) < \delta$, where $\theta(z)$ is some formula with parameters in N such that $N' \vDash \theta(d_+) < \delta$. By elementarity of the extension $(N, P_M) \preceq (N', P_{M'})$, we know that there is some (possibly infinite) tuple d'_+ in M such that $N \vDash \theta(d'_+) < \delta$, and thus $\int_{S_x(\mathcal{U})} \psi^+(x;d'_+) - \phi(x)\, d\mu < \varepsilon$ and $\sup_x \left(\phi(x) - \psi^+(x;d'_+)\right) < \varepsilon$. Similarly there exists a tuple d'^- in M such that $\int_{S_x(\mathcal{U})} \psi^-(x;d'^-) - \phi(x)\, d\mu < \varepsilon$ and $\sup_x(\psi^-(x;d'^-) - \phi(x)) < \varepsilon$. Combining these, we see that $\int_{S_x(\mathcal{U})} \psi^+(x;d'_+)\, d\mu - \int_{S_x(\mathcal{U})} \psi^-(x;d'^-)\, d\mu < 2\varepsilon$. As these formulas have parameters in M, and we can bound $\phi(x)$ above and below with $\phi^-(x;d'^-)(x) - \varepsilon < \phi(x) < \phi^+(x;d'_+)(x) - \varepsilon$, we see that for any measure ν extending $\mu|_M$, we have

$$\int_{S_x(M)} \psi^-(x;d'^-)\, d\mu|_M - \varepsilon < \int_{S_x(\mathcal{U})} \phi(x)\, d\nu < \int_{S_x(M)} \psi^+(x;d'_+)\, d\mu|_M + \varepsilon$$

limits the value of $\int_{S_x(\mathcal{U})} \phi(x)\, d\nu$ to an interval of width at most 4ε depending only on $\mu|_M$.

As ε was arbitrary, we see that ν is determined by $\mu|_M$, so μ is smooth. $\square$

Now as smooth measures are weakly orthogonal to all measures, it suffices to show that if all generically stable measures are smooth, then the theory is distal.

Lemma 4.5.3. *Let $I = (a_t : t \in [0,1])$ be an indiscernible segment. If for some model M containing I, $\mu_I \in \mathfrak{M}_x(M)$ is weakly orthogonal to itself, then I is distal.*

Proof. Assume that I is not distal. Then there exist points $0 < t_1 < t_2 < 1$ and b_1, b_2 such that the sequences $I_{[0,t_i)} + b_i + I_{(t_i,1]}$ defined by replacing a_{t_i} with b_i are indiscernible for both $i = 1, 2$, but the sequence $I_{[0,t_1)} + b_1 + I_{(t_1,t_2)} + b_2 + I_{(t_2,1]}$ defined by making both replacements is not indiscernible. By [And23b, Lemma 3.5.4], we may assume that $\mathrm{tp}(a_{t_i}/M) = \mathrm{tp}(b_i/M) = \lim(I_{[0,t_i)}/M)$ for $i = 1, 2$, where M is some small model containing I.

By the non-indiscernibility assumption, there is some formula $\phi(y_1, x_1, y_2, x_2, y_3)$ and c_1, c_2, c_3 finite subtuples of $I_{[0,t_1)}, I_{(t_1,t_2)}, I_{(t_2,1]}$ respectively such that $\phi(c_1, a_{t_1}, c_2, a_{t_2}, c_3) \neq \phi(c_1, b_1, c_2, b_2, c_3)$. Assume that $\phi(c_1, a_{t_1}, c_2, a_{t_2}, c_3) = 0$ while $\phi(c_1, b_1, c_2, b_2, c_3) = \varepsilon > 0$. Let u_1 be the maximum index such that $a_{u_1} \in c_1$, let v_1 be the minimum index such that $a_{v_2} \in c_2$, let u_2 be the maximum index such that $a_{u_2} \in c_2$, and let v_2 be the minimum index such that $a_{v_2} \in c_3$. Then $0 \leq u_1 < t_1 < v_1 \leq u_2 < t_2 < v_2 \leq 1$.

If $t'_i \in (u_i, v_i)$ for each i, then the partial type $\lim(I_{[0,t'_1)}/M) \times \lim(I_{[0,t'_2)}/M)$ is consistent with $\phi(c_1, x, c_2, y, c_3) = 0$, because there are realizations of these limit types that could replace $a_{t'_1}$ and $a_{t'_2}$ while preserving the indiscernibility of the sequence. We will show that $\lim(I_{[0,t'_1)}/M) \times \lim(I_{[0,t'_2)}/M)$ is also consistent with $\phi(c_1, x, c_2, x, c_3) = \varepsilon$. Let $\tau : [0, 1] \to [0, 1]$ is an order-preserving map that fixes all points in $[0, u_1] \cup [v_1, u_2] \cup [v_2, 1]$, but $\tau(t_1) = t'_1$ and $\tau(t_2) = t'_2$. Then by the homogeneity of $\mathcal{U}$, there is an automorphism $\sigma \in \mathrm{Aut}(\mathcal{U})$ such that for all $t \in [0, 1]$, $\sigma(a_t) = a_{\tau(t)}$. We see then that $\phi(c_1, \sigma(b_1), c_2, \sigma(b_2), c_3) = \varepsilon$, and that replacing $a_{t'_1}$ with $\sigma(b_1)$ or $a_{t'_2}$ with $\sigma(b_2)$ leaves I indiscernible. Thus by [And23b, Lemma 3.5.4], there are b'_1, b'_2 with $\mathrm{tp}(b'_i/M) = \lim(I_{[0,t'_i)}/M)$ for $i = 1, 2$ but $\phi(c_1, b'_1, c_2, b'_2, c_3) = \varepsilon$.

This tells us that if J and K are the indiscernible segments obtained by linearly reindexing $I_{[u_1,v_1]}$ and $I_{[u_2,v_2]}$ respectively, we find that $\mu_J, \mu_K \in \mathfrak{M}_x(M)$ cannot be weakly orthogonal. Otherwise, by Corollary 4.2.21, there is an M-definable predicate $\psi(x, y) = \sum_{i=1}^{m} \theta_i(x)\theta'_i(y)$ such that $M \vDash \phi(c_1, x, c_2, y, c_3) \leq \psi(x, y)$, but also for any ω extending $\mu_J \times \mu_K$, $\int_{S_{xy}(M)} \psi(x, y)\, d\omega < \varepsilon$. This cannot be true, as for any $t'_1 \in [u_1, v_1], t'_2 \in [u_2, v_2]$, we have $\psi(a_{t'_1}, a_{t'_2}) \geq \varepsilon$, and for any ω extending $\mu_J \times \mu_K$, we have

$$\int_{S_{xy}(M)} \psi(x, y)\, d\omega = \frac{1}{(v_1 - u_1)(v_2 - u_2)} \int_{u_1}^{v_1} \int_{u_2}^{v_2} \psi(a_t, a'_t)\, dt'\, dt \geq \varepsilon.$$

However, if μ_J and μ_K are not weakly orthogonal, then μ_I is not weakly orthogonal with itself. We see this by a proof analogous to that of Lemma 4.2.22, as $\mu_I = (v_1 - u_1)\mu_J + (1 - (v_1 - u_1))\mu_{[0,1]\setminus J} = (v_2 - u_2)\mu_K + (1 - (v_2 - u_2))\mu_{[0,1]\setminus K}$, and we see that for any ω extending $\mu_J \times \mu_K$, $\mu_I \otimes \mu_I + (v_1 - u_1)(v_2 - u_2)(\omega - \mu_J \otimes \mu_K)$ will also be a Keisler measure extending $\mu_I \times \mu_I$, which will differ from $\mu_I \otimes \mu_I$ if we choose $\omega \neq \mu_J \otimes \mu_K$. $\square$

As average measures for indiscernible segments are generically stable by Lemma 4.3.17, this completes the proof of Theorem 4.5.1. $\square$

4.5.1 Regularity by way of Weak Orthogonality

We now generalize the results from [CS18] about the Strong Erdős-Hajnal property and regularity in distal structures. First we will use the approach from [Sim16] to prove a regularity lemma nonconstructively using weakly orthogonal measures and ultraproducts, which we will prove equivalent to distality. Then in the next subsection we will show the same results using the explicit combinatorial approach from [CS18].

By Theorem 4.5.1, we know that a theory is distal if and only if all sequences of measures $\mu_1, \dots, \mu_n$ with μ_i generically stable for $i < n$ are weakly orthogonal, and thus by Theorem 4.4.10, a theory is distal if and only for each such tuples of measures and each predicate $\phi(x_1, \dots, x_n)$, one of the regularity properties from that theorem applies to ϕ over $\mu_1, \dots, \mu_n$. Now we will show that in fact, if the theory is distal, all of those properties hold in uniformly definable ways.

Lemma 4.5.4. *Assume T is distal. Then for each definable predicate $\phi(x_1, \dots, x_n; y)$ and each $\varepsilon > 0$, there is a finite set Ψ of definable predicates such that each $\psi(x_1, \dots, x_n; z) \in \Delta$ can be expressed as a sum of predicates of the form $\prod_{i=1}^n \psi_i(x_i; z_i)$, and if $M \vDash T$, $\mu_i \in \mathfrak{M}_{x_i}(M)$ are Keisler measures, with μ_i generically stable for $i < n$, and $b \in M^y$, then there are $\psi^-, \psi^+ \in \Delta, d_-, d_+ \in M^z$ such that if we write $x = (x_1, \dots, x_n)$, $\psi^-(x; d_-) \leq \phi(x; y) \leq \psi^+(x; d_+)$ and $\int_{S_x(M)} \psi^+(x; d_+) - \psi^-(x; d_-)\, d\mu_1 \times \cdots \times \mu_n \leq \varepsilon$.*

Proof. Fix $\phi(x_1, \ldots, x_n; y)$ and $\varepsilon > 0$. It suffices to show that for some finite set Ψ, and any model M, appropriate measures μ_i, and $b \in M^y$, there are $\psi^-, \psi^+ \in \Psi$ and $d_-, d_+ \in M^z$ such that $\sup_x \psi^-(x; d_-) \dot{-} \phi(x; y) \leq \frac{\varepsilon}{3}, \sup_x \phi(x; y) \dot{-} \psi^+(x; d_+) \leq \frac{\varepsilon}{3}$ and $\int_{S_x(M)} \psi^+(x; d_+) - \psi^-(x; d_-)\, d\mu_1 \times \cdots \times \mu_n \leq \frac{\varepsilon}{3}$. If so, then we may simply subtract $\frac{\varepsilon}{3}$ from ψ^- and add $\frac{\varepsilon}{3}$ to ψ^+.

Suppose that no such finite set Ψ works. Let Σ be the set of all possible definable predicates that can be expressed as finite sums of the form $\prod_{i=1}^n \psi_i(x_i; z_i)$. Let I be the set of finite subsets of Σ, for every finite subset $\Delta \in I$, let $S_\Delta = \{\Delta' \in I : \Delta \subseteq \Delta'\}$, and let $F = \{S \subseteq I : \exists \Delta \in I, S_\Delta \subseteq S\}$. This is the standard filter used in the ultrafilter proof of the compactness theorem, so there exists an ultrafilter U extending it.

For each finite $\Delta \subset \Sigma$, by our contradiction assumption, there are $M, \mu_1, \ldots, \mu_n, b$ such that for all $\psi^-, \psi^+ \in \Delta, d_-, d_+ \in M^z$,

$$\max\Bigg(\sup_x \left(\psi^-(x; d_-) \dot{-} \phi(x; b)\right), \sup_x \left(\phi(x; b) \dot{-} \psi^+(x; d_+)\right), \\ \int_{S_x(M)} \psi^+(x; d_+) - \psi^-(x; d_-)\, d\mu_1 \times \cdots \times \mu_n\Bigg) > \frac{\varepsilon}{3}.$$

Then we let $\tilde{M}$ be the ultraproduct of all these M with the ultrafilter U, and let $\tilde{\mu}_i$ be the ultralimits of the measures, with $\tilde{b}$ the ultraproduct of the parameters. By Lemma 4.3.13, for $i < n$, the measure $\tilde{\mu}_i$ is generically stable and thus smooth, so by Corollary 4.2.25, all of these measures are weakly orthogonal. Thus by Lemma 4.2.22, there are actually some ψ^-, ψ^+ and $\tilde{d}_-, \tilde{d}_+ \in \tilde{M}^z$ such that $\psi^-(x; d_-) \leq \phi(x; y) \leq \psi^+(x; d_+)$ and $\int_{S_x(M)} \psi^+(x; d_+) - \psi^-(x; d_-)\, d\mu_1 \times \cdots \times \mu_n < \frac{\varepsilon}{3}$. Thus on a U-large set of models M, we have

$$\max\Bigg(\sup_x \psi^-(x; d_-) \dot{-} \phi(x; y), \sup_x \phi(x; y) \dot{-} \psi^+(x; d_+), \\ \int_{S_x(M)} \psi^+(x; d_+) - \psi^-(x; d_-)\, d\mu_1 \times \cdots \times \mu_n\Bigg) < \frac{\varepsilon}{3}.$$

This contradicts our assumption, which made sure that on the U-large set of Δ containing ψ^-, ψ^+, this quantity was greater than $\frac{\varepsilon}{3}$. $\square$

We can use Lemma 4.5.4 to make the definability and constructibility in Theorem 4.4.10 uniform. We state these consequences separately as a distal regularity lemma and strong Erdős-Hajnal properties. We note that also by Theorem 4.4.10, any of these properties implies weak orthogonality of all generically stable measures, and thus by Theorem 4.5.1, distality.

Theorem 4.5.5. *Assume T is distal. For each definable predicate $\phi(x_1, \ldots, x_n; y)$ and $\varepsilon > 0$, there exist predicates $\psi_i(x_i; z_i)$, which can be chosen to be either definable or constructible, and a constant C such that if $\mu_1 \in \mathfrak{M}_{x_1}(M), \ldots, \mu_n \in \mathfrak{M}_{x_n}(M)$ are such that for $i < n$, μ_i is generically stable, $b \in M^y$, and $\delta > 0$, the following all hold: The predicate $\prod_{i=1}^n \psi_i(x_i; z_i)$ defines a (ε, δ)-distal regularity grid partition for $\phi(x_1, \ldots, x_n; b)$ of size $O(\delta^{-C})$.*

Proof. The lemmas in the proof of Theorem 4.4.10 all preserve the uniformity of predicates. Thus by starting with Lemma 4.5.4, we see that one of a finite set of predicates can be used to define distal regularity partitions, which we may assume is a single predicate by standard coding tricks. $\square$

In the case where $|x_1| = \cdots = |x_n|$ and all of the measures are equal, we can find a common refinement of the partitions of unity on each piece, and deal with a single partition.

Corollary 4.5.6. *Assume T is distal. For each definable predicate $\phi(x_1, \ldots, x_n; y)$ with $|x_1| = \cdots = |x_n| = |x|$, and $\varepsilon > 0$, there exists a predicate $\psi(x; z)$, which can be chosen to either be definable or constructible, and $\delta > 0$ such that if $\mu \in \mathfrak{M}_x(M)$ is generically stable, and $b \in M^y$, then ψ defines a partition P such that $\otimes_{i=1}^n P$ is a (ε, δ)-distal regularity partition for $\phi(x_1, \ldots, x_n; b)$, such that $|P| = O(\delta^{-C})$.*

Finally, we state the characterization of distality in terms of the definable strong Erdős-Hajnal property, generalizing [CS18, Theorem 3.1] and [CS18, Corollary 4.6]. This follows by applying the equivalences in the proof of Theorem 4.4.10 to Theorem 4.5.5.

Corollary 4.5.7. *A theory T is distal if and only if each definable predicate $\phi(x_1, \ldots, x_n; y)$ has the unformly definable ε-strong Erdős-Hajnal property with respect to all Keisler measures $\mu_1 \in \mathfrak{M}_{x_1}(M), \ldots, \mu_n \in \mathfrak{M}_{x_n}(M)$ for every $\varepsilon > 0$.*

Specifically, there exist definable predicates $\psi_i(x_i; z_i)$ and $\delta > 0$ such that if the measures $\mu_1 \in \mathfrak{M}_{x_1}(M), \ldots, \mu_n \in \mathfrak{M}_{x_n}(M)$ are such that for $i < n$, μ_i is generically stable, and $b \in M^y$, then for any product measure ω of $\mu_1, \ldots, \mu_n$, there are $d_i \in M^{z_i}$ such that $\psi_i(x_i; d_i)$ are $(\phi(x; b), \varepsilon)$-homogeneous and $\int_{S_{x_i}(M)} \psi_i(x_i; d_i)\, d\mu_i \geq \delta$ for each i.

Furthermore, for any $\varepsilon > \gamma \geq 0$, there are $\psi_i(x_i; z_i)$ and $\delta > 0$ such that if $\mu_1 \in \mathfrak{M}_{x_1}(M), \ldots, \mu_n \in \mathfrak{M}_{x_n}(M)$ are such that for $i < n$, μ_i is generically stable, $b \in M^y$, and ω is a product measure of $\mu_1, \ldots, \mu_n$ such that $\int_{S_{x_1 \ldots x_n}(M)} \phi\, d\omega \geq \varepsilon$, then there are $d_i \in M^{z_i}$ such that $\phi(a_1, \ldots, a_n; b) \geq \gamma$ whenever $\psi_i(a_i; d_i) > 0$ for each i, and $\int_{S_{x_i}(M)} \psi_i(x_i; d_i)\, d\mu_i \geq \delta$ for each i.

4.5.2 Distal Cutting Lemma

We now show how to find the predicates defining strong Erdős-Hajnal properties and distal regularity partitions more explicitly, using techniques that are also useful for distal combinatorics.

Definition 4.5.8. Let $\phi(x; y)$ be a definable predicate, and let $\nu \in \mathfrak{M}_x(M)$ be a generically stable measure. Then we say that a predicate $\psi(x; z)$ defines a (ε, δ)-*cutting of weight* γ for $\phi(x; y)$ with respect to ν when there is a finite set $D \subseteq M^z$ such that $\inf_x \sum_{d \in D} \psi(x; d) \geq \gamma$, and for each $d \in D$,

$$\nu\left(y : \operatorname{osc}(\phi(x; y), \{a : \psi(a; d) > 0\}) > \varepsilon\right) \leq \delta.$$

A (ε, δ)-cutting *of size at most* N consists of $\psi(x; z)$ and a particular valid choice of D

with $|D| \leq N$.

Lemma 4.5.9 (Generalizes [CS18, Claim 3.5] and [CGS20, Theorem 3.2])**.** *If M is distal, and $\phi(x;y)$ is a definable predicate, then for every $\varepsilon, \delta > 0$, there exists a $\gamma > 0$ and a predicate $\psi(x;z)$ that defines a (ε, δ)-cutting of size at most $O_{\phi,\varepsilon}(\delta^{-1} \ln \delta^{-1})$ and weight at least γ with respect to any generically stable measure $\nu \in \mathfrak{M}_y(M)$.*

Proof. Let $M, \phi(x;y), \varepsilon, \delta$ be as above. Let $\theta(x;z)$ be a strong honest definition for $\phi(x;y)$. Then we define $\chi(y;z) = \sup_x (\phi(x;y) - \theta(x;z)) \dot{-} \inf_x (\phi(x;y) + \theta(x;z))$.

Let $C = \mathrm{vc}_{\varepsilon/4}(\chi(y;z))$. By Theorem 4.3.11, there is a δ-net B for the fuzzy set system $(1-\chi)^{M^z}_{1-\varepsilon/2,1}$ with respect to ν, with $B = O\left(C\varepsilon^{-1} \ln \varepsilon^{-1}\right)$. This means that if d is such that $\chi(b;d) = 0$ for all $b \in B$, then $\nu\left(\chi(y;d) > \frac{\varepsilon}{2}\right) < \delta$.

Because θ is a strong honest definition, for every $a \in M^x$, there is some $d \in B^z$ such that $\theta(a;d) = 0$ and for all $a' \in M^x, b \in B$, $\theta(a';b) \geq |\phi(a;b) - \phi(a';b)|$. Thus also for all $b \in B, a' \in M^x$, $\phi(a';b) - \theta(a';d) \leq \phi(a;d) \leq \phi(a';b) + \theta(a';d)$, so $\chi(b;d) = 0$. Let D be the set of all $d \in B^z$ with $\chi(b;d) = 0$ for all $b \in B$, and recall then that for each $d \in D$, $\nu\left(\chi(y;d) > \frac{\varepsilon}{2}\right) < \delta$. We also find that for any $b \in M^y, d \in D$, and any $a, a' \in M^x$, we have $|\phi(a;b) - \phi(a';b)| \leq \chi(b;d) + \theta(a;d) + \theta(a';d)$. Now let k be such that there exists a definable predicate $\theta'(x;y_1,\dots,y_k)$ such that for all d, if $(d_1,\dots,d_k)$ is an initial segment of the tuple d, then $|\theta'(x;d_1,\dots,d_k) - \theta(x;d)| \leq \frac{\varepsilon}{8}$. We find that then there is a set $D_k \subseteq D$ of size at most $|B|^k$ such that for each $d \in D$, there is $(d_1,\dots,d_k) \in D_0$ an initial segment of d, so for all $a \in M^x$, there is $d \in D_k$ with $\theta(a;d) \leq \frac{\varepsilon}{8}$. Thus also $\inf_x \sum_{d \in D_k} \left(\frac{\varepsilon}{4} \dot{-} \theta(x;d)\right) \geq \frac{\varepsilon}{4}$, so we let $\psi(x;z) = \frac{\varepsilon}{4} \dot{-} \theta(x;d)$ and let $\gamma = \frac{\varepsilon}{4}$. If $d \in D_k, a, a' \in M^x$ are such that $\psi(a;d), \psi(a';d) > 0$, then $\theta(a;d), \theta(a';d) < \frac{\varepsilon}{4}$, and for all b outside a set of ν-measure at most δ, $\chi(b;d) \leq \frac{\varepsilon}{2}$, so $|\phi(a;b) - \phi(a';b)| \leq \chi(b;d) + \theta(a;d) + \theta(a';d) \leq \varepsilon$. $\square$

We can now use a cutting to prove a version of uniformly definable strong Erdős-Hajnal, and from it distal regularity, in two variables.

Lemma 4.5.10. *Let $\phi(x;y;w)$ be a definable predicate, and let $\varepsilon > 0$. Then for any $0 < \beta < \frac{1}{50\varepsilon^{-1}+5}$, there are $0 < \alpha < 1$ and definable predicates $\psi_1(x;z_1)$, $\psi_2(x;z_2)$ such that for any Keisler measure $\mu \in \mathfrak{M}_x(M)$, any generically stable measure $\nu \in \mathfrak{M}_y(M)$, and any $c \in M^w$, there are $d_1 \in M^{z_1}, d_2 \in M^{z_2}$ such that $\int_{S_x(M)} \psi_1(x;d_1)\, d\mu \geq \alpha$, $\int_{S_y(M)} \psi_2(y;d_2)\, d\nu \geq \beta$, and the pair $\psi_1(x;d_1), \psi_2(y;d_2)$ is $(\phi(x;y;c),\varepsilon)$-homogeneous.*

Proof. We will prove this for some M-definable $\phi(x;y) = \phi(x;y;c)$. As the formulas $\psi_1(x;z_1)$ and $\phi_2(y;z_2)$ will be constructed from a particular choice of strong honest definition for $\phi(x;y)$, it will suffice to show that there is some formula $\theta(x;z;w)$ such that for any $c \in M^z$, $\theta(x;z;c)$ is a strong honest definition for $\phi(x;y;c)$. To do this, we find a strong honest definition for $\phi(x;y,w)$, calling this $\theta(x;z,w)$, where z is a tuple of copies of y, and we have set all copies of w equal.

Let $s = \left\lceil \frac{10}{\varepsilon} \right\rceil$, and let $\delta = 1 - 5(s+1)\beta$, so that $\delta > 0$ but also $\beta = \frac{1-\delta}{5(s+1)}$.

As in the proof of Lemma 4.5.9, let $\theta(x;z)$ be a strong honest definition for $\phi(x;y)$, define $\psi^+(y;z) = \sup_x(\phi(x;y) - \theta(x;z)), \psi^-(y;z) = \inf_x(\phi(x;y) + \theta(x;z))$, and $\chi(y;z) = \psi^+(y;z) - \psi^-(y;z)$. Recall that there is some $\gamma > 0$ such that for any choice of ν, there is a finite set $D \in M^z$ of size at most $O_{\phi,\varepsilon}(\delta^{-1}\ln\delta^{-1})$ such that for each $d \in D$, $\nu(\chi(y;d) > \frac{1}{s}) < \delta$ and $\inf_x \sum_{d\in D}\left(\frac{\varepsilon}{4} \dot{-} \theta(x;d)\right) \geq \frac{\varepsilon}{4}$.

If we let $\alpha > 0$ be such that $\frac{\gamma}{|D|} \geq \alpha$, then there is always some $d \in D$ such that $\int_{S_x(M)} \psi(x;d)\, d\mu \geq \alpha$. We can thus let $\psi_1(x;d_1) = \psi(x;d)$.

We now let f_i be defined by $f_i(t) = 1 - |st - i|$, so that $(f_0, \ldots, f_s)$ is a partition of unity on $[0,1]$ and the support of f_i is $(\frac{i-1}{s}, \frac{i+1}{s})$. Thus $(f_i(\psi^+(y;d))f_j(\psi^-(y;d)) : 0 \leq i,j \leq s)$ forms a partition of unity on $S_y(M)$. If b is such that $f_i(\psi^+(b;d))f_j(\psi^-(b;d)) > 0$, then $i \geq j-1$, and also, $\frac{i-j-2}{s} \leq \chi(b;d) = \psi^+(b;d) - \psi^-(b;d) \leq \frac{i-j+2}{s}$. Thus if $\chi(b;d) \leq \frac{1}{s}$, we find that $i - j \leq 3$. Thus on all such b, $\sum_{i,j:-1\leq i-j\leq 3} f_i(\psi^+(b;d))f_j(\psi^-(b;d)) = 1$. The measure of such b is at least $1-\delta$, so $\sum_{i,j:-1\leq i-j\leq 3}\int_{S_y(M)} f_i(\psi^+(y;d))f_j(\psi^-(y;d))\, d\nu \geq 1-\delta$, and thus for some i,j with $-1 \leq i-j \leq 3$, $\int_{S_y(M)} f_i(\psi^+(y;d))f_j(\psi^-(y;d))\, d\nu \geq \frac{1-\delta}{5(s+1)} = \beta$, so we can let $\psi_2(y; d_2) = f_i(\psi^+(y; d))f_j(\psi^-(y; d))$, using standard coding tricks to account for the finitely many choices of i, j.

We now check homogeneity. For all b in the support of that $f_i(\psi^+(y; d))f_j(\psi^-(y; d))$, and all a in the support of $\frac{\varepsilon}{4} \dot{-} \theta(x; d)$, we have that $\frac{j-1}{s} - \frac{\varepsilon}{4} \leq \phi(a; b) \leq \frac{i+1}{s} + \frac{\varepsilon}{4}$, and this interval is of width at most $\frac{5}{s} + \frac{\varepsilon}{2} \leq \varepsilon$. Thus the pair $\left(\frac{\varepsilon}{4} \dot{-} \theta(x; d), f_i(\psi^+(y; d))f_j(\psi^-(y; d))\right)$ is (ϕ, ε)-homogeneous. □

Fixing some β and setting $\delta = \min(\alpha, \beta)$, we get an actual definable strong Erdős-Hajnal statement.

Corollary 4.5.11. *Let $\phi(x; y; w)$ be a definable predicate, and let $\varepsilon > 0$. There are $\delta > 0$ and definable predicates $\psi_1(x; z_1)$, $\psi_2(x; z_2)$ such that for any Keisler measure $\mu \in \mathfrak{M}_x(M)$, any generically stable measure $\nu \in \mathfrak{M}_y(M)$, and any $c \in M^w$, there are $d_1 \in M^{z_1}, d_2 \in M^{z_2}$ such that $\int_{S_x(M)} \psi_1(x; d_1)\, d\mu \geq \delta$, $\int_{S_y(M)} \psi_2(y; d_2)\, d\nu \geq \delta$, and the pair $\psi_1(x; d_1), \psi_2(y; d_2)$ is (ϕ, ε)-homogeneous.*

We now use Lemmas 4.4.13 and 4.4.14 to show that if the integral of ϕ is large enough, the value of the predicate is positive on the whole pair. This allows us to induct in dimension, and find an alternate proof of Theorem 4.5.5.

Theorem 4.5.12. *For each definable predicate $\phi(x_1, \dots, x_n; y)$ and $\varepsilon > \gamma \geq 0$, there exist definable predicates $\psi_i(x_i; z_i)$ and $\delta > 0$ such that if $\mu_1 \in \mathfrak{M}_{x_1}(M), \dots, \mu_n \in \mathfrak{M}_{x_n}(M)$ are such that for $i < n$, μ_i is generically stable, and $b \in M^y$, for any product measure ω of $\mu_1, \dots, \mu_n$, if $\int_{S_{x_1 \dots x_n}(M)} \phi\, d\omega \geq \varepsilon$, then there are $d_i \in M^{z_i}$ such that $\phi(a_1, \dots, a_n; b) \geq \gamma$ whenever $\psi_i(a_i; d_i) > 0$ for each i, and $\int_{S_{x_i}(M)} \psi_i(x_i; d_i)\, d\mu_i \geq \delta$ for each i.*

Proof. For a base case, we start with Corollary 4.5.11, and then applying Lemmas 4.4.13 and 4.4.14, recalling that all localizations of generically stable measures are generically stable by Corollary 4.2.26.

Now assume that this holds all predicates with variables partitioned in n pieces, and consider $\phi(x_1, \ldots, x_n, x_{n+1}; y)$. We repartition it as $\phi(x_1, \ldots, x_n; x_{n+1}; y)$, and apply the base case to this binary predicate, getting some $\psi(x_1, \ldots, x_n; z), \psi_{n+1}(x_{n+1}; z_{n+1})$ such that for any measures $\mu \in \mathfrak{M}_{x_1 \ldots x_n}(M), \mu_{n+1} \in \mathfrak{M}_{x_{n+1}}(M)$ with μ generically stable, any product measure ω of μ, μ_{n+1}, and any $c \in M^y$, if $\int_{S_{x_1 \ldots x_{n+1}}(M)} \phi(x_1, \ldots, x_{n+1}; c)\, d\omega \geq \varepsilon$, then there are $d \in M^z, d_{n+1} \in M^{z_{n+1}}$ such that $\phi(a_1, \ldots, a_{n+1}; c) \geq \gamma$ whenever $\psi(a_1, \ldots, a_n; d) > 0$ and $\psi_{n+1}(a_{n+1}; d_{n+1}) > 0$, and as far as integrals, $\int_{S_{x_1 \ldots x_n}(M)} \psi(x_1, \ldots, x_n; d)\, d\mu \geq \delta_0$ and $\int_{S_{x_{n+1}}(M)} \psi_{n+1}(x_{n+1}; d_{n+1})\, d\mu_{n+1} \geq \delta_0$.

Now we can apply the induction hypothesis to $\psi(x_1, \ldots, x_n; z)$, giving us predicates $\psi_i(x_i; z_i)$ for $1 \leq i \leq n$ and some $\delta_1 > 0$ such that such that for any measures $\mu_i \in \mathfrak{M}_{x_i}(M)$ with each μ_i generically stable, any product measure ω of the μ_is, and any $d \in M^z$, if $\int_{S_{x_1 \ldots x_n}(M)} \psi(x_1, \ldots, x_n; d)\, d\omega \geq \delta_0$ there are $d_i \in M^{z_i}$ such that $\psi(a_1, \ldots, a_n; d) > 0$ whenever $\psi_i(a_i; d_i) > 0$ for each i, and for each i, $\int_{S_{x_i}(M)} \psi_i(x_i; d_i)\, d\mu_i \geq \delta_1$.

We now let $\delta = \min(\delta_0, \delta_1)$. For any $c \in M^y$, generically stable measures $\mu_i \in \mathfrak{M}_{x_i}(M)$ for $1 \leq i \leq n$, and measure $\mu_{n+1} \in \mathfrak{M}_{x_{n+1}}(M)$, and any product measure ω of the μ_is, we let μ be the restriction of ω to the variables $x_1 \ldots x_n$. As μ is a product measure of the $\mu_1, \ldots, \mu_n$, and these measures are smooth, it is $\mu_1 \otimes \cdots \otimes \mu_n$, which is itself smooth. Thus there are d, d_{n+1} such that on the support of $\psi(x_1, \ldots, x_n; d)\psi_{n+1}(x_{n+1}; d_{n+1})$, $\phi(x_1, \ldots, x_n; c) \geq \gamma$, $\int_{S_{x_{n+1}}(M)} \psi_{n+1}(x_{n+1}; d_{n+1})\, d\mu_{n+1} \geq \delta_0 \geq \delta$, and $\int_{S_{x_1 \ldots x_n}(M)} \psi(x_1, \ldots, x_n; d)\, d\mu \geq \delta_0$. From this last integral, we see that there are d_i for $1 \leq i \leq n$ such that on the support of $\prod_{i=1}^{n} \psi_i(x_i; d_i)$, $\psi(x_1, \ldots, x_n; d) > 0$, so on the support of $\prod_{i=1}^{n+1} \psi_i(x_i; d_i)$, $\phi(x_1, \ldots, x_n; c) \geq \gamma$. Also, for each i, $\int_{S_{x_i}(M)} \psi_i(x_i; d_i)\, d\mu_i \geq \delta_1 \geq \delta$. $\square$

Now by the equivalences in the proof of Theorem 4.4.10, this gives us another proof of Theorem 4.5.5.

4.5.3 Equipartitions

In classical logic, by [CS18, Corollary 5.14] and [Sim16, Proposition 3.3], distal regularity partitions can be chosen to be *equipartitions*, where the measures of each piece are approximately equal. In the case of $[0, 1]$-valued partitions of unity, it is trivial to split a partition of unity into pieces of approximately equal integral. However, it is not so easy to repartition a partition into sets of approximately equal measure, and this will require *uniform cutting of generically stable measures*. In this subsection, we check that we can modify our results about constructible partitions to work with equipartitions in a uniformly constructible way.

Lemma 4.5.13. *In a distal structure M, if $\mu \in \mathfrak{M}_x(M)$ is a generically stable measure and $p \in S_x(M)$ a type with $\mu(\{p\}) > 0$, then p is realized in M.*

Proof. By Theorem 4.5.2, μ is smooth, so by Corollary 4.2.26, the localization measure $\mu_{\{p\}} = \delta_p$ is smooth as well. A type that is smooth over M as a measure is realized in M, because any nonrealized type has multiple realizations, each of which would be a valid extension. □

Lemma 4.5.14. *Any distal structure M* uniformly cuts generically stable measures. *That is, for every definable predicate $\phi(x; y)$ and $\varepsilon > 0$, there is a definable predicate $\chi(x; z)$ such that if $\mu \in \mathfrak{M}_x(M)$ is a measure such that for all $a \in M^x$, $\mu(\{a\}) = 0$, and $0 \leq r \leq \mu(\phi(M))$, then there exists $c \in M^z$ with $|\mu(\phi(M) \cap \chi(M; c)) - r| \leq \varepsilon$.*

Proof. It suffices to show this for the trivial predicate $\phi(x) = 0$, because we can simply replace μ with its localization to $\phi(M)$, and replace r with $\frac{r}{\mu(\phi(M))}$. This works unless $\mu(\phi(M)) = 0$, when this is still trivial.

Now fix μ. Let $p \in S_x(M)$ be a type. Assume for contradiction that $\mu(\{p\}) > 0$. Then by Lemma 4.5.13, p is realized, contradicting our assumption on μ, so $\mu(\{p\}) = 0$, and μ is atomless. As μ is also regular, if $p \in S_x(M)$, then there must be an open set $U \subseteq S_x(M)$ containing p such that $\mu(U) < \varepsilon$. We can express U as $\psi(x) < \delta$ for some M-definable predicate $\psi(x)$ and $\delta \in [0,1]$. As $\psi(p) < \delta$, there is some $\psi(p) < \delta' < \delta$, so if we let U_p be the open set defined by $\psi(x) < \delta'$, and let F_p be the closed set defined by $\psi(x) \leq \delta'$, we find that $p \in U_p \subseteq F_p$ and $\mu(F_p) \leq \mu(U) \leq \varepsilon$.

By compactness, $S_x(M)$ can be covered with finitely many open sets U_p. In fact, there is some K where in can be covered with at most K many open sets U_p, where the sets F_p are uniformly definable as $\chi(M;c)$ for various parameters $c \in M^z$. We show this by contradiction. For each finite set F of pairs $(K, \chi(x;z))$, find a generically stable measure μ_F such that this fails for each $(K,\chi) \in F$. By taking an ultraproduct of these counterexamples according to an appropriate ultrafilter, as in the proof of Lemma 4.5.4, we find a generically stable measure μ such that this fails for every K and every definable predicate $\chi_1(x;z)$. This gives a contradiction, as for every μ, there is some finite cover of open sets U_p contained in closed sets F_p, and by the standard coding tricks, a single formula $\chi_1(x;z)$ can be used for each F_p in the finite cover.

We can then find a formula $\chi(x;z)$ such that for any $k \leq K$ and any $c_1, \dots, c_k$, there is some c such that $\chi(M;c) = \bigcup_{i \leq k} \chi_1(M;c_i)$. We can cover $S_x(M)$ with closed sets $\chi_1(M;c_1), \dots, \chi_1(M;c_K)$, each of measure at most ε, and assume that $c_1, \dots, c_k$ form a minimal subset such that $\mu(\bigcup_{i=1}^k \chi(M;c_k)) \geq r$. By minimality, we have that $|\mu(\bigcup_{i=1}^k \chi(M;c_k)) - r| \leq \varepsilon$. □

Lemma 4.5.15. *Any distal structure M* uniformly cuts finite sets. *That is, for every definable predicate $\phi(x;y)$ and $\varepsilon > 0$, there is a definable predicate $\chi(x;z)$ such that for any sufficiently large finite set $A \subseteq M^x$, any $b \in M^y$, and any $0 \leq m \leq |\phi(A;b)|$, either $|\phi(A;b) = 0|$ or there is some $c \in M^z$ such that*

$$\left| \frac{|\phi(A;b) \cap \chi(A;c)|}{|\phi(A;b)|} - \frac{m}{|\phi(A;b)|} \right| \leq \varepsilon.$$

Proof. It is enough to show this for $\phi(x;y)$ which is uniformly 0. Specifically, we will show that for all $\varepsilon > 0$, there is a definable predicate $\chi(x;z)$ such that for any sufficiently large finite set $A \subseteq M^x$, any $b \in M^y$, and any $r \in [0,1]$, there is some $c \in M^z$ such that

$$\left|\frac{|\chi(A;c)|}{|A|} - r\right| \leq \varepsilon.$$

We can then apply this with $\phi(A;b)$ in place of A, and $r = \frac{m}{|\phi(A;b)|}$.

Assume for contradiction that this does not hold. Let $(A_n, r_n) : n \in \mathbb{N}$ be a sequence of counterexamples, with $|A_n| \geq n$ for each n. That is, for any predicate $\chi(x;z)$, any $c \in M^z$ and any $n \in \mathbb{N}$,

$$\left|\frac{|\chi(A_n;c)|}{|A_n|} - r_n\right| > \varepsilon.$$

Now let μ_n be the uniform measure on A_n for each n. Fix a nonprincipal ultrafilter on $\mathbb{N}$, and let μ be the ultraproduct of the μ_ns, and $r \in [0,1]$ be the ultralimit of the r_ns. Then μ is a generically stable measure with $\mu(\{c\}) = 0$ for each c. Thus Lemma 4.5.14 applies. Let $\chi(x;z)$ be as given by that lemma, but with $\frac{\varepsilon}{2}$ substituted for ε. Then there exists some sequence $(c_n : n \in \mathbb{N})$ such that if c is the element of the ultraproduct representing that sequence, $|\mu(\chi(M;c)) - r| \leq \frac{\varepsilon}{2}$, and thus also on a large set of indices n, $|\mu(\chi(M;c_n)) - r_n| < \varepsilon$. However, for each such n,

$$|\mu(\chi(A_n;c_n)) - r_n| = \left|\frac{|\chi(A_n;c_n)|}{|A_n|} - r_n\right|,$$

contradicting the choice of (A_n, r_n). $\square$

Having seen that distal structures uniformly cut finite sets, we can make the partition in Corollary 4.5.6 a uniformly constructible equipartition, as in [CS18, Corollary 5.14].

Corollary 4.5.16. *Let $\phi(x_1, \ldots, x_k)$ be a definable predicate where each x_i is a copy of the same variable tuple x, and fix $\varepsilon > 0$. Then there is a constructible predicate $\psi(x;z)$ and some $C > 0$ such that the following holds:*

For any generically stable measure $\mu \in \mathfrak{M}_x(M)$ such that $\mu(\{a\}) = 0$ for all $a \in M^x$, and *any $\gamma, \delta > 0$, ψ defines a constructible (ε, δ)-distal regularity partition P of M^x of size at most $\mathcal{O}(\delta^{-C})$, such that each cell in P is uniformly (in terms of $\phi, \varepsilon, \delta, \gamma$) constructible over a set of parameters of size $\mathcal{O}(\delta^{-C})$, such that for any two sets $A, B \in P$, $|\mu(A) - \mu(B)| \leq \gamma$.*

Proof. We start with $\psi(x; z)$ and $C > 0$ as given by Theorem 4.5.5, with $\frac{\varepsilon}{2}$ playing the role of ε. (This ψ and C will not be the final ψ and C.) Fix μ, γ, δ.

We can use the same repartitioning argument from the proof of [CS18, Corollary 5.14] to form an equipartition, using the predicate χ from Lemma 4.5.14 to cut the measure μ. The resulting equipartition will consist of boolean combinations of pieces from the previous partition and χ-zerosets, and thus with the usual coding tricks, are uniformly constructible. $\square$

CHAPTER 5

Examples of Distal Metric Structures

5.1 Introduction

This chapter, joint work with Itaï Ben Yaacov, provides examples of distal metric structures, and contrasts them with notable non-examples observed by James Hanson. We also give a statement, in Section 5.2 of the strong Erdös-Hajnal property for metric structures. This property of distal structures, unlike distality, passes to reducts, and we will provide an example of a stable metric structure that still has this property.

In Section 5.3, we examine certain metric valued fields. These structures are constructed by taking a field with valuation in $\mathbb{R}_{\geq 0}$, and incorporating the valuation metric into the metric structure. In [Ben14], theories of algebraically and real closed metric valued fields are developed. In Section 5.3, we show, using the indiscernible sequence definition, that real closed metric valued fields are distal. We also show that algebraically closed metric valued fields are interpretable in real closed metric valued fields, from which we conclude that these have the strong Erdös-Hajnal property, although they are stable and thus not distal.

Section 5.4 explores a fundamentally different distal metric theory, which we call *dual linear continua*. Models of this theory consist of the set of functions from some linear continuum (such as the linear order $[0,1]$) to $[0,1]$ which are continuous, nondecreasing, and surjective, with a particular structure placed upon them. In the case of the linear continuum $[0,1]$, the automorphism group of this structure is the group of increasing homeomorphisms from $[0,1]$ to itself. This structure had been studied before in [Ben18] and in [Iba16], where it had been shown to be NIP but decidedly not stable. We show that in fact, the structure is distal, both by studying its indiscernible sequences and by constructing explicit distal cell decompositions.

Finally, in Section 5.5, we examine some metric structures which are NIP but not distal, suggested by James Hanson. These include any metric structure expanding a Banach space, and in particular the Keisler randomization of any (metric) structure. This shows that unlike stability [BK09] or NIP [Ben09], distality is *not* preserved by taking randomizations.

5.2 Strong Erdős-Hajnal

In addition to considering examples of distal metric structures, we will identify interesting reducts of distal metric structures. Even if these reducts are no longer distal, they will retain properties such as the *strong Erdős-Hajnal property.* In [CS18], it was shown that distality is equivalent to the *definable* strong Erdős-Hajnal property, which implies the strong Erdős-Hajnal property for all of its reducts. This characterization of distality was extended to metric structures in [And23c], and we will now describe the strong Erdős-Hajnal property for reducts of distal metric structures. We define homogeneity for sets and definable predicates as in [And23c]:

Definition 5.2.1. For $i = 1, \dots, n$, let $A_i \subseteq M^{x_i}$, let $\phi(x_1, \dots, x_n)$ be a definable predicate (possibly with parameters) and let $\varepsilon > 0$. Then we say that $(A_i : 1 \leq i \leq n)$ is (ϕ, ε)-*homogeneous* when for all $(a_i : i \in I), (a'_i : i \in I) \in A_1 \times \cdots \times A_n$, $|\phi(a_1, \dots, a_n) - \phi(a'_1, \dots, a'_n)| \leq \varepsilon$.

If for $1 \leq i \leq n$, $\psi_i(x_i)$ are definable predicates (possibly with parameters), we say that $(\psi_i(x_i) : 1 \leq i \leq n)$ are (ϕ, ε)-homogeneous when the supports $\psi_i(x_i) > 0$ are.

Fact 5.2.2 ([And23c, Corollary 4.5.7])**.** *A theory T of continuous logic is distal if and only if every definable predicate $\phi(x_1, \dots, x_n; y)$ has the definable strong Erdős-Hajnal property:*

For every $\varepsilon > 0$, there exist definable predicates $\psi_i(x_i; z_i)$ and $\delta > 0$ such that if $\mu_1 \in \mathfrak{M}_{x_1}(M), \ldots, \mu_n \in \mathfrak{M}_{x_n}(M)$ are such that for $i < n$, μ_i is generically stable, and $b \in M^y$, then for any product measure ω of $\mu_1, \ldots, \mu_n$, there are $d_i \in M^{z_i}$ such that $\psi_i(x_i; d_i)$ are $(\phi(x; b), \varepsilon)$-homogeneous and $\int_{S_{x_i}(M)} \psi_i(x_i; d_i)\, d\mu_i \geq \delta$ for each i.

As all counting measures are generically stable, we can deduce the following:

Lemma 5.2.3. *Assume T is a distal theory in continuous logic. If $\phi(x_1, \ldots, x_n)$ is a definable predicate with parameters, then ϕ has the strong Erdős-Hajnal property: Then for every $\varepsilon > 0$, there is some $\delta > 0$ such that if $A_1, \ldots, A_n$ are finite subsets of $M^{x_1}, \ldots, M^{x_n}$ respectively, then there are $B_1, \ldots, B_n$ with $B_i \subseteq A_i$ and $|B_i| \geq \delta |A_i|$ such that $(B_i : 1 \leq i \leq n)$ is (ϕ, ε)-homogeneous.*

Lemma 5.2.4. *The strong Erdős-Hajnal property is closed under continuous combinations: if for $1 \leq j \leq n$, definable predicates $\phi_j(x) = \phi_j(x_1, \ldots, x_m)$ have the strong Erdős-Hajnal property, and $u : [0, 1]^n \to [0, 1]$ is continuous, then $u(\phi_1(x), \ldots, \phi_n(x))$ also has the strong Erdős-Hajnal property.*

Proof. Fix $\varepsilon > 0$, and $A_i \subseteq M^{x_i}$ finite for each $1 \leq i \leq m$. By continuity, let $\delta > 0$ be such that if $a, b \in [0, 1]^n$ have $\max_i |a_i - b_i| \leq \delta$ in the sup metric, then $|u(a) - u(b)| \leq \varepsilon$. By the strong Erdős-Hajnal property of $\phi_1, \ldots, \phi_n$, there is some $\gamma > 0$ such that for each $1 \leq j \leq n$, if $B_i \subseteq M^{x_i}$ are finite, there are $C_i \subseteq B_i$ with $|C_i| \geq \gamma |B_i|$ such that $(C_1, \ldots, C_m)$ are (ϕ_j, δ)-homogeneous. Thus we can set $A_i^0 = A_i$, and recursively define A_i^j such that $A_i^j \subseteq A_i^{j-1}$, $|A_i^j| \subseteq \gamma |A_i^{j-1}|$, and $(A_1^j, \ldots, A_m^j)$ is (ϕ_j, δ)-homogeneous. Then $(A_1^n, \ldots, A_m^n)$ will be (ϕ_j, δ)-homogeneous for all $1 \leq j \leq n$, and thus also $(u(\phi_1, \ldots, \phi_n), \varepsilon)$-homogeneous. Also, $|A_i^n| \geq \delta^n |A_i|$, and δ did not depend on the choice of A_i. $\square$

Lemma 5.2.5. *The strong Erdős-Hajnal property is closed under uniform limits: if for $j \in \mathbb{N}$, definable predicates $\phi_j(x) = \phi_j(x_1, \ldots, x_m)$ have the strong Erdős-Hajnal property and converge uniformly to $\phi(x)$, then $\phi(x)$ also has the strong Erdős-Hajnal property.*

Proof. Fix $\varepsilon > 0$. Let N be large enough that $\sup_x |\phi_N(x) - \phi(x)| \leq \frac{\varepsilon}{3}$. Then let $\delta > 0$ be such that if $B_i \subseteq M^{x_i}$ are finite, there are $C_i \subseteq B_i$ with $|C_i| \geq \delta |B_i|$ such that $(C_1, \ldots, C_m)$ are $\left(\phi_N, \frac{\varepsilon}{3}\right)$-homogeneous. Then if we fix $A_i \subseteq M^{x_i}$ finite for each $1 \leq i \leq m$, there are $B_i \subseteq A_i$ for each i with $|B_i| \geq |A_i|$ and for all $a, b \in B_1 \times \cdots \times B_m$, we have $|\phi(a) - \phi(b)| \leq |\phi(a) - \phi_N(a)| + |\phi_N(a) - \phi_N(b)| + |\phi_N(b) - \phi(B)| \leq \varepsilon$. □

These lemmas show that in a quantifier-elimination language, to determine if all definable predicates in a structure have the strong Erdős-Hajnal property, it suffices to check for atomic formulas.

We can also reduce checking the ε-strong Erdős-Hajnal property for all ε to a simpler criterion.

Lemma 5.2.6. *A definable predicate $\phi(x_1, \ldots, x_n)$ has the strong Erdős-Hajnal property if and only if for all $0 \leq r < s \leq 1$, there is some $\delta > 0$ such that such that if $A_1, \ldots, A_n$ are finite subsets of $M^{x_1}, \ldots, M^{x_n}$ respectively, then there are $B_1, \ldots, B_n$ with $B_i \subseteq A_i$ and $|B_i| \geq \delta |A_i|$ such that either for all $b \in B_1 \times \cdots \times B_n$, $\phi(b) < s$, or for all $b \in B_1 \times \cdots \times B_n$, $\phi(b) > r$.*

Proof. Suppose ϕ has the strong Erdős-Hajnal property, fix $0 \leq r < s \leq 1$, and let $0 < \varepsilon < s - r$. Then we can find $B_1, \ldots, B_n$ of adequate size that are ε-homogeneous, implying that either $\phi(b) > r$ or $\phi(b) < s$ is true for all $b \in B_1 \times \cdots \times B_n$.

Conversely, assume this new condition holds. We will prove for each n that ϕ has the $\frac{1}{n}$-strong Erdős-Hajnal property. By taking a finite minimum, we can find $\delta > 0$ such that for all $0 \leq i < n$, given $A_1, \ldots, A_n$, there are $B_i \subseteq A_i$ and $|B_i| \geq \delta |A_i|$ such that either for all $b \in B_1 \times \cdots \times B_n$, $\phi(b) < \frac{i+1}{n}$, or for all $b \in B_1 \times \cdots \times B_n$, $\phi(b) > \frac{i}{n}$. Then by a recursive application of this property for each $r = \frac{i}{n}, \frac{i+1}{n}$, we can find $B_i \subseteq A_i$ with $|B_i| \subseteq \delta^n |A_i|$ that satisfy this property for each (r, s) simultaneously. Thus there must be some i such that $b \in B_1 \times \cdots \times B_n$, $\frac{i}{n} \leq \phi(b) \leq \frac{i+1}{n}$, so $B_1 \times \cdots \times B_n$ is $\left(\phi, \frac{1}{n}\right)$-homogeneous. It would suffice to reduce the size of the sets only $\log n$ times by a binary search method, improving the constants if necessary. $\square$

Before trying to determine which metric structures have the strong Erdős-Hajnal property for all definable predicates, it makes sense to ask whether the metric has this property. This is true for ultrametrics.

Lemma 5.2.7. *Let (X, d) be a bounded ultrametric space. The metric $d(x, y)$ has the strong Erdős-Hajnal property.*

Proof. Fix $0 \leq r < 1$, and let $A, B \subseteq X$ be finite. We will show that there are $A_0 \subseteq A, B_0 \subseteq B$ with $|A_0| \leq \frac{1}{3}|A|$ and $|B_0| \leq \frac{1}{3}|B|$ such that either for all $(a, b) \in A_0 \times B_0$, $d(a, b) \leq r$, or for all $(a, b) \in A_0 \times B_0$, $d(a, b) > r$.

By the ultrametric criterion, $A \cup B$ can be covered with disjoint closed r-balls. Thus let $A = A_1 \cup \cdots \cup A_n$ and $B = B_1 \cup \cdots \cup B_n$, where $A_1 \cup B_1$ is contained in a closed r-ball, but for $i \neq j$, if $u \in A_i \cup B_i$ and $v \in A_j \cup B_j$, then $d(u, v) > r$. If there is some i with $|A_i| \geq \frac{1}{3}|A|$ and $|B_i| \geq \frac{1}{3}|B|$, then we can let $A_0 = A_i$ and $B_0 = B_i$. Let $S \subseteq \{1, \ldots, n\}$ be the set of all i such that $\frac{|A_i|}{|A|} \geq \frac{|B_i|}{|B|}$. We can see that $\sum_{i \in S} \frac{|A_i|}{|A|} \geq \sum_{i \in S} \frac{|B_i|}{|B|} = 1 - \sum_{i \notin S} \frac{|B_i|}{|B|}$, from which we can deduce that either $\sum_{i \in S} \frac{|A_i|}{|A|} \geq \frac{1}{2}$ or $\sum_{i \notin S} \frac{|B_i|}{|B|} \geq \frac{1}{2}$. Without loss of generality, assume the former. In this case, choose a minimal set $S' \subseteq S$ with $\sum_{i \in S'} \frac{|A_i|}{|A|} \geq \frac{1}{3}$. By minimality, for any one $i' \in S'$, $\sum_{i \in S', i \neq i'} \frac{|A_i|}{|A|} < \frac{1}{3}$, and thus $\sum_{i \in S', i \neq i'} \frac{|B_i|}{|B|} \leq \frac{1}{3}$. By assumption, either $|A_{i'}| < \frac{1}{3}|A|$ and $|B_{i'}| < \frac{1}{3}|B|$. As $i' \in S$, meaning $\frac{|A_{i'}|}{|A|} \geq \frac{|B_{i'}|}{|B|}$, we can deduce that $\frac{|B_{i'}|}{|B|} < \frac{1}{3}$, so $\sum_{i \in S'} \frac{|B_i|}{|B|} \leq \frac{2}{3}$, and $\sum_{i \notin S'} \frac{|B_i|}{|B|} \geq \frac{1}{3}$. Thus we can let $A_0 = \cup_{i \in S'} A_i$ and let $B_0 = \cup_{i \notin S'} B_i$, and get $|A_0| \geq \frac{1}{3}|A|$ and $|B_0| \geq \frac{1}{3}|B|$. If $a \in A_0$ and $b \in B$, then there are $i \in S'$ and $j \notin S'$ with $a \in A_i$ and $b \in B_j$, so as $i \neq j$, $d(a, b) > r$. $\square$

5.3 Valued Fields

In [Ben14], Ben Yaacov set up a framework for studying fields with $(\mathbb{R}_{\geq 0}, *)$-valued valuations as metric structures. More specifically, the metric structures are projective spaces over such fields.

Definition 5.3.1. Given a field K, let $K\mathbb{P}^n$ denote the n-dimensional projective space over K, whose elements we write in homogeneous coordinates as $[x_0 : x_1 : \cdots : x_n]$, which we will generally assume satisfy $\max_i |x_i| = 1$.

Let $\mathcal{L}_{\mathbb{P}^1}$ be the language considering of the constant symbol ∞ and, for each $n \in \mathbb{N}$ and each polynomial $P(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, x_1, \ldots, x_n]$, a relation symbol $||P(\bar{x})||$.

Given a field K with a multiplicative valuation $|\cdot|$ taking values in $\mathbb{R}_{\geq 0}$, we interpret $K\mathbb{P}^1$ as an $\mathcal{L}_{\mathbb{P}^1}$-structure as follows, using homogeneous coordinates:

$$d([a : a^*], [b : b^*]) = |ab^* - a^*b|$$

$$\infty = [1 : 0]$$

$$||P([a_1 : a_1^*], \ldots, [a_n : a_n^*])|| = |P^h(a_1, \ldots, a_n, a_1^*, \ldots, a_n^*)|,$$

where P^h is the *homogenization* of P.

Fact 5.3.2 ([Ben14, Theorem 1.8]). *There is a theory MVF in the language $\mathcal{L}_{\mathbb{P}^1}$, whose models are (up to isomorphism) exactly the projective lines of valued fields with complete valuation.*

We also can consider a language with more sorts, to encompass all projective spaces over K in one structure:

Definition 5.3.3. Let $\mathcal{L}_{\mathbb{P}}$ be the language with sorts $(P^n : n \in \mathbb{N})$ with the following symbols:

- For each m, n, a function $\otimes : \mathbb{P}^m \times \mathbb{P}^n \to \mathbb{P}^{n+m+nm}$

- For each $A \in SL_{n+1}(\mathbb{Z})$, a function $A : \mathbb{P}^n \to \mathbb{P}^n$
- For each n, a predicate symbol $||\cdot||$ on $\mathbb{P}^n$.

Given any field K with a multiplicative valuation $|\cdot|$ taking values in $\mathbb{R}_{\geq 0}$, we construct an $\mathcal{L}_{\mathbb{P}}$-structure $K\mathbb{P}$ by interpreting $\mathbb{P}^n$ as $K\mathbb{P}^n$. We interpret the $\otimes$ symbols as Segre embeddings, interpret the special linear transformation symbols with their natural action on K^{n+1}, each of which respects the quotient relation that defines $K\mathbb{P}^n$. We can then define the other symbols by

$$||[a_0 : \cdots : a_n]|| = |a_0|$$
$$d(a,b) = \max_{i<j} |a_i b_j - a_j b_i|.$$

Fact 5.3.4. *The $\mathcal{L}_{\mathbb{P}}$-structure $K\mathbb{P}$ and $\mathcal{L}_{\mathbb{P}^1}$-structure $K\mathbb{P}^1$ induced by a valued field K are biinterpretable.*

The theory MVF admits a natural algebraically closed completion:

Fact 5.3.5 ([Ben14, Lemma 2.2])**.** *There is a $\mathcal{L}_{\mathbb{P}^1}$-theory $ACMVF$ whose models are precisely the projective lines over algebraically closed fields with nontrivial complete valuations.*

To define the theory of real closed metric valued fields, we extend the language:

Definition 5.3.6. Extend $\mathcal{L}_{\mathbb{P}^1}$ to the language $\mathcal{L}_{o\mathbb{P}^1}$ by adding a for each such polynomial $P(x_1, \dots, x_n)$ an extra symbol $\langle P(\bar{x})\rangle$, which we interpret as

$$\langle P(\bar{x})\rangle = d(P(\bar{x}), \mathrm{Sq}),$$

where Sq is the (closed in any metric valued field) set of squares.

In a real closed ordered field, Sq is also the set of nonnegative elements, so we can naturally think of this as encoding a linear ordering. This gives rise to the languages $RCMVF$ and $ORCMVF$ of (ordered) real closed metric valued fields:

Fact 5.3.7 ([Ben14, Proposition 3.6, Theorem 3.11]). *There are a $\mathcal{L}_{\mathbb{P}^1}$-theory $RCMVF$ and a $\mathcal{L}_{o\mathbb{P}^1}$-theory $ORCMVF$ such that the models of $RCMVF$ are exactly the projective lines of real closed fields with complete non-trivial valuations, and models of $ORCMVF$ are exactly the projective lines of such fields where the extra predicate is the distance predicate to the set of nonnegative elements.*

Furthermore, any model of $RCMVF$ admits a unique expansion modelling $ORCMVF$. In this expansion, the extra predicate is the distance predicate to the set of nonnegative elements.

We can now show that these theories are distal.

Theorem 5.3.8. *RCMVF is distal.*

Proof. By Theorem 3.5.22, it suffices to check that if $(a_i : i \in \mathbb{Q}) + b + (c_j : j \in \mathbb{Q})$ is an indiscernible sequence, with $(a_i : i \in \mathbb{Q}) + (c_j : j \in \mathbb{Q})$ indiscernible over a singleton d, then $(a_i : i \in \mathbb{Q}) + b + (c_j : j \in \mathbb{Q})$ is indiscernible over d also.

Let $i_0 < \cdots < i_{n-1} \in \mathbb{Q}$ and $i_{n+1} < \cdots < i_{2n} \in \mathbb{Q}$. We will show that for all $i_n > i_{n-1}$, and all $\varphi(x; y_0, \ldots, y_{2n})$, $\varphi(d; a_{i_0}, \ldots, a_{i_n}, c_{i_{n+1}}, \ldots, c_{i_{2n}}) = \varphi(d; a_{i_0}, \ldots, a_{i_{n-1}}, b, c_{i_{n+1}}, \ldots, c_{i_{2n}})$.

By quantifier elimination, it suffices to show that if $\varphi(x; y_0, \ldots, y_{2n})$ is an atomic $\mathcal{L}_{o\mathbf{P}^1}$-formula of either the form $\|P(x; \bar{y})\|$ or $\langle P(x; \bar{y})\rangle$, then $\vDash \varphi(d; \bar{a}) = \varphi(d; \bar{a}')$, whenever $\bar{a}$ and $\bar{a}'$ are increasing sequences of length $2n+1$ in $(a_i : i \in \mathbb{Q}) + b + (c_j : j \in \mathbb{Q})$. As in the proof of [Ben14, Theorem 3.12], we find that $\varphi(x; \bar{y})$ is a continuous combination of things of the form $|x - f(\bar{y})|$ and $\langle x - f(\bar{y})\rangle$, where f is a partial $\emptyset$-definable function. Thus it will suffice to show the desired result for φ of those forms. Given y, let $f_0(y) = f(a_{i_0}, \ldots, a_{i_{n-1}}, y, c_{i_{n+1}}, \ldots, c_{i_{2n}})$. We wish to show that $|d - f_0(y)|$ and $\langle d - f_0(y)\rangle$ are constant on the indiscernible sequence $I = (a_i : i > i_{n-1}) + b + (c_i : i < i_{n+1}$. The sequence $f_0(y) : y \in I$ will itself be indiscernible, and thus monotone, and $f_0(y) : y \in I \setminus \{b\}$ is indiscernible over d, so $|d - f_0(y)|$ and $\langle d - f_0(y)\rangle$ are constant over $I \setminus \{b\}$.As for any values r, s, the set of y such that $|d - f_0(y)| = r$ and

$\langle d - f_0(y)\rangle = s$ is order-convex, we see that $|d - f_0(y)|$ and $\langle d - f_0(y)\rangle$ must also be constant on all of I as desired. $\square$

It is also possible to interpret $\mathrm{ACMVF}_{(0,0)}$ in RCMVF, and thus show the (not definable) strong Erdős-Hajnal property for that stable theory. In general, if K is a metric valued field, it is complete and thus Henselian, so if L/K is a finite-degree field extension, and thus L is a metric valued field with the unique valuation extending the valuation on K. We claim that L is interpretable in K, and as a consequence, $\mathrm{ACMVF}_{(0,0)}$ is interpretable in RCMVF.

Theorem 5.3.9. *Let K be a metric valued field, and let L be a finite extension of K, with the unique valuation extending that of K. Then L is interpretable in K.*

Proof. Let $d = [L : K]$, and let $\alpha \in L$ be the root of a monic degree-d polynomial in $K[X]$ such that $L = K(\alpha)$.

Roughly speaking, we will represent an element $W \in L\mathbb{P}^1$ with two elements of $K\mathbb{P}^d$, spelling out W and W^{-1} in the α-basis. From the construction of $K\mathbb{P}$, we see that for any homogeneous polynomial $P \in \mathbb{Z}[X_0, \dots, X_d]$, $|P(X_0, \dots, X_d)|$, evaluated at a representative where $\bigcap_{i=0}^{d} |X_i| = 1$, is a definable predicate on $K\mathbb{P}^d$ without parameters. If instead $P \in K[X_0, \dots, X_d]$, this will be definable with parameters.

Our interpretation will use the element $[X_0 : \dots : X_d] \in K\mathbb{P}^n$ to represent $[\sum_{i=0} X_i \alpha^i : X_d] \in L\mathbb{P}^1$. This is well-defined, and it is surjective because any $[Y : 1]$ can be represented by some $[X_0 : \dots : X_{d-1} : 1]$, and the single point at infinity $[1 : 0]$ can be represented by $[1 : 0 : \dots : 0]$. We wish to show that for any polynomial $P \in \mathbb{Z}[Y_1, \dots, Y_n, Y_1^*, \dots, Y_n^*]$, the predicate $||P||$, evaluated at $[X_{01} : \dots : X_{d1}], \dots, [X_{0n} : \dots : X_{dn}]$, is a definable predicate. To do this, we will first show that for any polynomial $Q \in \mathbb{Z}[Y_1, \dots, Y_n, Z_1, \dots, Z_n]$, homogeneous in each pair (Y_i, Z_i), the function

$$\left| Q\left(\sum_{i=0} X_{i1}\alpha^i, \dots, \sum_{i=0} X_{in}\alpha^i, X_{d1}, \dots, X_{dn}\right)\right|$$

is a definable predicate. Then if P^h is the homogenization of P, we can evaluate $||P||$ by calculating

$$\left| P^h \left(\sum_{i=0} X_{i1}\alpha^i, \ldots, \sum_{i=0} X_{in}\alpha^i, X_{d1}, \ldots, X_{dn} \right) \right|$$

and then correcting for the max norm $|\sum_{i=0} X_{ij}\alpha^i| \vee |X_{dj}|$ for each j, by dividing by the appropriate power of $|\sum_{i=0} X_{ij}\alpha^i| \vee |X_{dj}|$, which is itself a nowhere-zero definable predicate, as it is the maximum of the valuations of two homogeneous polynomials, namely $|Y_j|$ and $|Y_j^*|$.

For all $x \in L$, we can understand the valuation $|x|$ in terms of the norm $|N_{L/K}(x)| = |x|^d$, as $|N_{L/K}(x)| = |\prod_{i=1}^d x_i| = |x|^d$, where $\{x_1, \ldots, x_d\}$ are the conjugates of x under the d automorphisms of L/K, each of which has $|x_i| = |x|$ by Henselianity of the complete field K. The norm $N_{L/K}\left(\sum_{i=0}^{d-1} X_i\alpha^i\right)$ can be defined as a determinant, and in particular is a homogeneous degree d polynomial in $K[X_0, \ldots, X_{d-1}]$, so if $Q_0, \ldots, Q_{d-1} \in K[X_{ij} : 0 \le i \le d, 1 \le j \le n]$ are polynomials homogeneous in each tuple $(X_{0i}, \ldots, X_{di})$ of the same multidegree (or zero), then $N_{L/K}\left(\sum_{i=0}^{d-1} Q_i\alpha^i\right)$ is itself a homogeneous polynomial in $K[X_{ij} : 0 \le i \le d, 1 \le j \le n]$, so $\left|\sum_{i=0}^{d-1} Q_i\alpha^i\right| = \left|N_{L/K}\left(\sum_{i=0}^{d-1} Q_i\alpha^i\right)\right|^{1/d}$ will be a definable predicate. For each $Q \in \mathbb{Z}[Y_1, \ldots, Y_n, Z_1, \ldots, Z_n]$ is homogeneous in each pair (Y_i, Z_i), with $d_i = \deg_{Y_i}(Q) + \deg_{Z_i}(Q)$, then we can express

$$Q \left(\sum_{i=0} X_{i1}\alpha^i, \ldots, \sum_{i=0} X_{in}\alpha^i, X_{d1}, \ldots, X_{dn} \right) = \sum_{i=0}^{d-1} Q_i\alpha^i,$$

where each Q_i is homogeneous in each tuple $(X_{0i}, \ldots, X_{di})$ with the same multidegree $d_i = \sum_{j=0}^d \deg_{X_{ji}} Q$, unless it is zero. Thus $|Q| = \left|\sum_{i=0}^{d-1} Q_i\alpha^i\right|$ is definable. $\square$

5.4 Dual Linear Continua

In [Ben18], Ben Yaacov analyzes an $\aleph_0$-categorical metric structure whose homeomorphism group is $\mathrm{Hom}^+([0,1])$, the group of increasing homeomorphisms of $[0,1]$ under the topology of uniform convergence. We call models of the theory of this structure *dual linear continua*, because we will show that they are in correspondence with linear continua with endpoints, which are characterized by the following definition and fact.

Definition 5.4.1. A *linear continuum* is a dense linear ordering with the least upper bound property.

Fact 5.4.2. *A linear order is connected in the order topology if and only if it is a linear continuum, and it is connected and compact if and only if it is a linear continuum with endpoints.*

Definition 5.4.3. Given a linear ordering L, let M_L be the set of functions $f : L \to [0,1]$ such that

- f is nondecreasing,
- f is continuous with respect to the order topology on L,
- $\inf_x f(x) = 0$,
- $\sup_x f(x) = 1$.

We give M_L the sup metric.

In [Ben18], $M_{[0,1]}$ is given additional structure, which makes its automorphism group $\mathrm{Hom}^+([0,1])$. If $f \in \mathrm{Hom}^+([0,1])$, then f acts on $M_{[0,1]}$ by composition, sending $g \in M_{[0,1]}$ to $g \circ f^{-1} \in M_{[0,1]}$. In order to describe the analogous structure on M_L for other linear orders, we will first describe the type spaces of this structure.

Lemma 5.4.4. *The type* $\operatorname{tp}(f_1, \ldots, f_n)$ *is determined exactly by the image of the function* $(f_1, \ldots, f_n) : [0,1] \to [0,1]^n$.

Proof. In [Ben18], it is shown that the type of $(f_1, \ldots, f_n)$ is determined by the function $(g_1, \ldots, g_n)$ such that $\frac{1}{n}\sum_{i=1}^n f_i = \mathrm{id}$ and $(g_1, \ldots, g_n) \circ \left(\frac{1}{n}\sum_{i=1}^n f_i\right) = (f_1, \ldots, f_n)$. This correspondence is a homomorphism between the space of such function tuples and the space of types. The function $(g_1, \ldots, g_n)$ has the same image as $(f_1, \ldots, f_n)$, so the type of a tuple determines its image.

We now consider two tuples $(f_1, \ldots, f_n)$ and $(g_1, \ldots, g_n)$ with the same image, and show that they have the same type. We may assume that $\frac{1}{n}\sum_{i=1}^n f_i = \frac{1}{n}\sum_{i=1}^n g_i = \mathrm{id}$, and show that the tuples are equal. For any $t \in [0,1]$, there is some $t' \in [0,1]$ such that $(f_1(t), \ldots, f_n(t)) = (g_1(t'), \ldots, g_n(t'))$. However, $\sum_{i=1}^n f_i(t) = nt$ and $\sum_{i=1}^n f_i(t') = nt'$, so $t = t'$, and the tuples are equal.

$\square$

Given that correspondence, if $p \in S_n(\emptyset)$ is a type in this theory, let $\operatorname{im}(p)$ be the image of any realization of p in $(M_{[0,1]})^n$. (Such a realization exists because $M_{[0,1]}$ is $\aleph_0$-categorical and thus $\aleph_0$-saturated.) We will use this characterization to understand the topology and metric on the type space, but first, some simple topological lemmas. (Recall that while in general, the metric on a type space does not induce the topology, it does in the $\aleph_0$-categorical case.)

Definition 5.4.5. If I is a set, give $[0,1]^I$ the product order defined by $(x_i : i \in I) \leq (y_i : i \in I)$. A *chain from 0 to 1* in $[0,1]^I$ is a set $C \subseteq [0,1]^I$ which is a chain in the product order and contains the constant tuples with values 0 and 1.

Lemma 5.4.6. *Let* I *be a set, and let* $C \subseteq [0,1]^I$ *be a chain from 0 to 1. Then the subset topology on* C *is the order topology, and* C *is compact.*

Proof. First, we note that for each $i \in I, r \in [0,1]$, there is some $f \in C$ such that $f(i) = r$. If not, then we may partition C with the two disjoint open sets $\{f \in C : f(i) < r\}$ and

$\{f \in C : f(i) > r\}$, contradicting connectedness.

To show the topologies agree, it suffices, without loss of generality, to show that for $f \in C$, the closed interval $[0, f] \subseteq C$ is closed in the subset topology, and that for any $r \in [0, 1]$ and $i \in I$, the set $\{f \in C : f(i) \leq r\}$ is closed in the order topology.

By definition, the closed interval $[0, f]$ is the set $\bigcap_{i \in I}\{g \in C : g(i) \leq f(i)\}$, which is closed in the subset topology.

Meanwhile, $\{f \in C : f(i) \leq r\} = \bigcap_{g \in C: g(i) > r}[0, g]$. For each $g \in C$ with $g(i) > r$, it follows that $\{f \in C : f(i) \leq r\} \subseteq [0, g]$ because C is linearly ordered. Also, for each $s \in (r, 1]$, there is some $g \in C$ with $g(i) = s$, so $\bigcap_{g \in C: g(i) > r}[0, g] \subseteq \{f \in C : f(i) \leq r\}$.

In the order topology, by Fact 5.4.2, connectedness and endpoints imply compactness. □

Lemma 5.4.7. *Let I be a countable set. Then if $C \subseteq [0, 1]^I$ is a connected chain from 0 to 1, then there are continuous, surjective, nondecreasing functions $(f_i : i \in I)$ such that $\mathrm{im}(f_i : i \in I) = C$.*

Proof. By taking a bijection, we may assume that I is an initial segment of $\mathbb{N}$. If $I = \{0, \ldots, n\}$, we let C' be the set of all $c \in [0, 1]^{\mathbb{N}}$ such that $c \restriction_I \in C$. This is clearly also a chain from 0 to 1, which is connected because it is the image of C under a continuous map that just duplicates coordinates. If C' is compact, then C is also the image of C' under a continuous map that deletes coordinates, so C is compact. If there are continuous, surjective, nondecreasing functions $(f_i : i \in \mathbb{N})$ such that $\mathrm{im}(f_i : i \in I) = C'$, then $(f_i : i \leq n)$ will suffice for C, so we may assume that $I = \mathbb{N}$.

Define $g : C \to [0, 1]$ by $g(c) = \sum_i c(i)2^{-i}$. This is a strictly increasing continuous function, which attains values 0 and 1. Because it is defined on a connected set, its image is connected, so g is surjective. Because C is a chain and g is strictly increasing, g is also injective, so it is a homeomorphism as its domain and codomain are compact Hausdorff. Thus we can let each f_i be the ith coordinate map of g^{-1}. These are continuous, surjective, and nondecreasing, and $(f_i : i \in I) = g^{-1}$, whose image is C. □

We now characterize the type spaces.

Lemma 5.4.8. *The map* im *is an isometry between the type space $S_n(\emptyset)$ in the theory of $M_{[0,1]}$, and the set of all connected chains from 0 to 1 in $[0,1]^n$, given the Hausdorff metric as compact subsets of $[0,1]^n$, itself given the sup metric.*

Proof. First we confirm that the image of im is what we claim, and then we will show that im, as a function to the set of compact subsets of $[0,1]^n$ with the Hausdorff metric, is an isometry. This is an injective continuous map between compact Hausdorff spaces, so it is a homeomorphism onto its image, and the rest of the lemma will follow from these two claims.

Clearly if $(f_1, \dots, f_n) \in M^n$, then $\mathrm{im}(f_1, \dots, f_n)$ is a connected chain from 0 to 1. If $P \subseteq [0,1]^n$ is a connected chain from 0 to 1, then by Lemma 5.4.7, there is some $(f_1, \dots, f_n) : [0,1] \to [0,1]^n$ with P as its image, and each f_i continuous, surjective, and nondecreasing. Thus $P = \mathrm{im}(f_1, \dots, f_n)$ and $f_1, \dots, f_n \in M^n_{[0,1]}$, so such sets are exactly the images of n-types over M.

Now we check that the metric coincides with the metric on types. Let $p, q \in S_n(\emptyset)$. First we show that $d(\mathrm{im}(p), \mathrm{im}(q)) \leq d(p, q)$. As

$$d(\mathrm{im}(p), \mathrm{im}(q)) = \max\left(\sup_{x \in \mathrm{p}} d(x, \mathrm{im}(q)), \sup_{y \in \mathrm{im}(q)} d(y, \mathrm{im}(p))\right)$$

and $d(p, q) = \inf_{\bar{f}, \bar{g} : \mathrm{tp}(\bar{f}) = p, \mathrm{tp}(\bar{g}) = q} d(\bar{f}, \bar{g})$, it suffices to show, without loss of generality, that for each $\bar{f}, \bar{g}$ such that $\mathrm{tp}(\bar{f}) = p, \mathrm{tp}(\bar{g}) = q$, and each $x \in \mathrm{im}(p)$, $d(x, \mathrm{im}(q)) \leq d(\bar{f}, \bar{g})$. Let t be such that $x = (\bar{f})(t)$. Then

$$d(x, \mathrm{im}(q)) \leq d(\bar{f}(t), \bar{g}(t)) \leq d(\bar{f}, \bar{g}).$$

It now suffices to show that there exist $\bar{f}, \bar{g}$ with $\mathrm{tp}(\bar{f}) = p, \mathrm{tp}(\bar{g}) = q$ such that $d(\bar{f}, \bar{g}) \leq d(\mathrm{im}(p), \mathrm{im}(q))$. Let $\bar{f}^*, \bar{g}^*$ be such that $\mathrm{tp}(\bar{f}^*) = p, \mathrm{tp}(\bar{g}^*) = q$, and $\frac{1}{n}\sum_{i=1}^n f_i^* = \frac{1}{n}\sum_{i=1}^n g_i^* = \mathrm{id}$. We will show that there exists a connected chain $C \subseteq [0,1]^2$ containing $(0,0)$ and $(1,1)$, such that for all $(t, t') \in C$, $d(\bar{f}^*(t), \bar{g}^*(t')) \leq d(\mathrm{im}(p), \mathrm{im}(q))$. By Lemma 5.4.7, there are continuous, surjective, nondecreasing functions f', g' such that $\mathrm{im}((f', g')) = C$. Then let $\bar{f} = \bar{f}^* \circ f'$ and $\bar{g} = \bar{g}^* \circ g'$. We know that $\mathrm{tp}(\bar{f}) = \mathrm{tp}(\bar{f}^*) = p$ and $\mathrm{tp}(\bar{g}) = \mathrm{tp}(\bar{g}^*) = q$, and we know that for each t, $(f'(t), g'(t)) \in C$, so $d(\bar{f}(t), \bar{g}(t)) \leq d(\mathrm{im}(p), \mathrm{im}(q))$, so these $\bar{f}, \bar{g}$ will suffice.

To construct the chain C, first assume without loss of generality that $d(\mathrm{im}(p), \mathrm{im}(q)) = \sup_{x \in \mathrm{im}(p)} d(x, \mathrm{im}(q))$. Then for all $t \in [0, 1]$, let $Y_t = \{t' : d(f^*(t), g^*(t')) \leq d(\mathrm{im}(p), \mathrm{im}(q))\}$. This is a closed interval in $\mathrm{im}(q)$. It will always be nonempty by the assumption that $d(\mathrm{im}(p), \mathrm{im}(q)) = \sup_{x \in \mathrm{im}(p)} d(x, \mathrm{im}(q))$. For each t, let $y_t = \min Y_t$. Then $t \mapsto y_t$ is a nondecreasing function from $[0, 1] \to [0, 1]$, so it is piecewise continuous with countably many discontinuities. Thus filling in these discontinuities with countably many vertical intervals turns the graph of this function into a path from $(0, 0)$ to $(1, 1)$, which we call C. It now suffices to show that for all $(t, t') \in C$, $d(\bar{f}^*(t), \bar{g}^*(t')) \leq d(\mathrm{im}(p), \mathrm{im}(q))$, that is, $t' \in Y_t$. If $t' = y_t$, this follows by definition, so we may assume that (t, t') lies on one of the vertical segments, so $\lim_{s \to t^-} y_s \leq t' \leq \lim_{s \to t^+} y_s$. Because $\{(t, t') : t' \in Y_t\}$ is closed, we find that $\lim_{s \to t^-}(s, y_s)$ and $\lim_{s \to t^+}(s, y_s)$ are both points of $\{(t, t') : t' \in Y_t\}$. Thus t' lies between two points in Y_t, which is an interval, so $t' \in Y_t$. □

Now we can define the structure on M_L for any linear continuum with endpoints L, by defining the type of any tuple in the type spaces $S_n(\emptyset)$ of the structure $M_{[0,1]}$, which from this point on we simply call S_n.

Definition 5.4.9. Given a linear continuum with endpoints L and $f_1, \ldots, f_n \in M_L$, let $\mathrm{tp}(f_1, \ldots, f_n)$ be the unique type p with $\mathrm{im}(p) = \mathrm{im}(f_1, \ldots, f_n)$.

For this to genuinely define a structure on M_L, it suffices to check that for each C-Lipschitz n-ary definable relation, which is interpreted as some $h : S_n \to [0, 1]$, that $h \circ \mathrm{tp} : M_L^n \to [0, 1]$ is also C-Lipschitz. This follows from tp being a contraction. By the proof of Lemma 5.4.8, for any $p, q \in S_n$, $d(p, q) = \inf_{\bar{f}, \bar{g} \in M_L : \mathrm{tp}(\bar{f}) = p, \mathrm{tp}(\bar{g}) = q} d(\mathrm{tp}(\bar{f}), \mathrm{tp}(\bar{g}))$, so this is indeed a contraction.

As the finite-dimensional type spaces coincide with those of $M_{[0,1]}$, these structures are all elementarily equivalent. We now characterize arbitrary type spaces over models of this theory.

Lemma 5.4.10. *Let x be a possibly infinite variable tuple. The type space S_x consists of all connected chains from 0 to 1 in* $[0,1]^x$.

Proof. The type space S_x is the topological inverse limit of all S_y where y is a finite subtuple of x. Thinking of each S_y as the set of connected chains from 0 to 1 in $[0,1]^y$, the restriction maps $S_y \to S_z$ for $z \subseteq y$ are given by restricting variables of chains. This means that S_x is homeomorphic to the inverse image of these spaces S_y as a subset of $[0,1]^x$, and it suffices to determine which sets in $[0,1]^x$ have connected chains as each finite projection. Such sets are exactly inverse limits of directed systems of connected chains from 0 to 1 in finite-dimensional spaces - that is, those sets whose projections to finite-dimensional spaces are connected chains from 0 to 1. It is clear that such projections are chains from 0 to 1 if and only if the original set is a chain from 0 to 1, and that the projections of a connected set are all connected. It suffices now to show that a set X whose finite-dimensional projections are connected chains from 0 to 1 is connected. Such a set X is an inverse limit of closed sets by Lemma 5.4.7, so it is itself closed, and thus compact. Thus if X is disconnected, there are basis open sets A, B such that $X \cap A, X \cap B$ partition X. However, using the standard basis of the topology, this means there is a finite subtuple y of x and open sets $A, B \subseteq [0,1]^y$ such that A, B partition $\pi(X)$, where π projects $[0,1]^x$ onto $[0,1]^y$. This contradicts the connectedness of $\pi(X)$. $\square$

We can now fully characterize models of $\mathrm{Th}(M_{[0,1]})$.

Theorem 5.4.11. *If $M \equiv M_{[0,1]}$, then M is isomorphic to M_L for some linear continuum with endpoints L.*

Proof. Let $p \in S_M$ be the type of M enumerated as a tuple, and let $L = \mathrm{im}(p)$. By Lemma 5.4.10, this is a connected chain from 0 to 1, and is thus a linear continuum with endpoints. We now define $f : M \to M_L$. If $m \in M, x = (x_m : m \in M) \in L$, then we define $f(m)(x) = x_m$.

Let $m_1, \dots, m_n \in M$. We wish to show that $\mathrm{tp}(m_1, \dots, m_n) = \mathrm{tp}(f(m_1), \dots, f(m_n))$, by showing the images of the types are equal. We know that $\mathrm{im}(\mathrm{tp}(m_1, \dots, m_n))$ is just the projection of $L = \mathrm{im}(p)$ onto the coordinates $(m_1, \dots, m_n)$, coinciding precisely with $\mathrm{im}(\mathrm{tp}(f(m_1), \dots, f(m_n))) = \mathrm{im}(f(m_1), \dots, f(m_n))$. As f preserves types, it is also an isometry.

Let $g \in M_L$, fix $n \in \mathbb{N}$, and then let $x_1 < \dots < x_n \in L$ be such that for each i, $g(x_i) = \frac{i}{n+1}$ for each $1 \leq i \leq n$. Then for each $1 \leq i < n$, there is some m_i with $f(m_i)(x_i) < f(m_i)(x_{i+1})$. As the sequence $\frac{1}{n-1} \sum_{i=1}^{n-1} f(m_i)(x_j)$ for $1 \leq j \leq n$ is strictly increasing, there is a continuous monotone bijection $\theta : [0,1] \to [0,1]$ such that for each $1 \leq j < n$, $\theta\left(\frac{1}{n-1} \sum_{i=1}^{n-1} f(m_i)(x_j)\right) = \frac{j}{n+1}$, and also $\theta(0) = 0$ and $\theta(1) = 1$.

We claim that the function that sends $g_1, \dots, g_{n-1} \in M_{[0,1]}$ to the function given by $t \mapsto \theta\left(\frac{1}{n-1} \sum_{i=1}^{n-1} g_i(t)\right)$ is definable. To show this, we first show that the predicate taking the average of $n-1$ elements of $M_{[0,1]}$ is definable, and then check that composition with θ is definable. As shown in [Ben10a], it suffices to show that these functions are type-definable. The graph of the average function, viewed as a subset of the type space, consists of all connected chains from 0 to 1 through $[0,1]^n$ contained in the closed subset $x_n = \frac{1}{n-1} \sum_{i=1}^{n-1} x_i$. The set of connected chains from 0 to 1 contained in a closed subset is closed, as the type space topology on chains is given by the Vietoris topology. Thus the graph of the average function is closed, as is the graph of composition with θ, consisting of all chains residing in the (closed) graph of θ.

By the definability of this function, as $M \equiv M_{[0,1]}$, there is some $m \in M$ such that for each $(a_1, \dots, a_n; b) \in (m_1, \dots, m_n; m)$, $b = \theta\left(\frac{1}{n-1} \sum_{i=1}^{n-1} a_i\right)$. Thus for each $1 \leq j \leq n$, $f(m)(x_j) = \theta\left(\frac{1}{n-1} \sum_{i=1}^{n-1} f(m_i)(x_j)\right) = \frac{j}{n+1} = g(x_j)$. Thus for all $x \in L$, $f(m)(x)$ and $g(x)$

lie in a common interval $\left[\frac{j}{n+1}, \frac{j+1}{n+1}\right]$, and thus $d(f(m), g) \leq \frac{1}{n+1}$. Because f is an isometry and M is complete, this shows us that f is a bijection. As f also preserves types, it is an isomorphism. $\square$

5.4.1 Indiscernibles

The structure $M_{[0,1]}$ is known to be NIP and in a certain precise way, purely unstable.[Iba16, Corollary 4.17] We will study its indiscernible sequences, and show that it is distal, in analogy to the $\aleph_0$-categorical structure $(\mathbb{Q}, <)$, whose automorphism group is $\mathrm{Hom}^+(\mathbb{Q} \cap [0, 1])$.

We now show that all types in the theory of $M_{[0,1]}$ are determined by types on two variables.

Lemma 5.4.12. *If $p \in S_n$, then $p(x_1, \dots, x_n)$ is implied by $\bigcup_{i<j} p \restriction_{x_i x_j}$, where $p \restriction_{x_i x_j}$ is the restriction of p to the variable tuple $x_i x_j$.*

In general, if $p \in S_n$ and $(a_1, \dots, a_n)$ is such that for each $i < j$, $(a_i, a_j) \in \mathrm{im}(p \restriction_{x_i x_j})$, then $(a_1, \dots, a_n) \in \mathrm{im}(p)$.

Proof. Let $p \in S_n(\emptyset)$ be a type with $(a_i, a_j) \in \mathrm{im}(p \restriction_{x_i x_j})$ for each $i < j$. Let $(f_1, \dots, f_n)$ be a realization of p in $M^n_{[0,1]}$. Then consider the n closed intervals $f_i^{-1}(\{a_i\})$. Because $(a_i, a_j) \in \mathrm{im}(p \restriction_{x_i x_j})$, the intervals $f_i^{-1}(\{a_i\})$ and $f_j^{-1}(\{a_j\})$ must nontrivially intersect. Closed real intervals have the 2-Helly property - any family of intervals that intersect pairwise has a nontrivial intersection, so $(a_1, \dots, a_n)$ must be in the image of $(f_1, \dots, f_n)$.

If p, q are types such that for each $i < j$, $p \restriction_{x_i x_j} = q \restriction_{x_i x_j}$, then for each $(a_1, \dots, a_n) \in \mathrm{im}(p)$, we know that for each $i < j$, $(a_i, a_j) \in \mathrm{im}(q \restriction_{x_i x_j})$, so $(a_1, \dots, a_n) \in \mathrm{im}(q)$. Thus $p = q$, and these types are determined by their restrictions to two variables. $\square$

We now analyze (possibly finite) indiscernible sequences in structures elementarily equivalent to $M_{[0,1]}$. Whenever $M \equiv M_{[0,1]}$, by Theorem 5.4.11, we may assume that $M = M_L$ for some linear continuum L.

Lemma 5.4.13. *Let L be a linear continuum with endpoints and let $(f_i : i \in I)$ be an indiscernible sequence in one variable in M_L, and let $(a_i : i \in I) \in \operatorname{im}(f_i : i \in I)$. Then $(a_i : i \in I)$ is either nondecreasing or nonincreasing.*

Proof. It suffices to show that if $(f, g, h) \in M_L^3$ is indiscernible, and there is some $t_0 \in L$ such that $(f, g, h)(t_0) = (a, b, c)$, then $a \leq b \leq c$ or $a \geq b \geq c$.

If $b = c$, this is trivial, so we may assume without loss of generality that $b < c$. By indiscernibility, there also exists t_1 such that $(f, g)(t_1) = (b, c)$. Then because $g(t_0) = b < c = g(t_1)$, we know that $t_0 < t_1$, and thus $a = f(t_0) \leq f(t_1) = b$. □

Lemma 5.4.14. *Let L be a linear continuum with endpoints and let $(f_i : i \in I)$ be an indiscernible sequence in one variable in M_L of length at least 3, and let $i < j$ be elements of I, with $(a, b) \in \operatorname{im}(f_i, f_j)$. Then either $(a, a) \in \operatorname{im}(f_i, f_j)$ or $(b, b) \in \operatorname{im}(f_i, f_j)$.*

Proof. Let $(f, g, h) \in M_L^3$ be indiscernible, with $(a, b) \in \operatorname{im}(f, g)$. Without loss of generality, assume $a < b$. There must be some c with $(a, b, c) \in \operatorname{im}(f, g, h)$, and by Lemma 5.4.13, $b \leq c$. If $b = c$, then $(b, b) \in \operatorname{im}(g, h) = \operatorname{im}(f, g)$, and we are done. Otherwise, $b < c$.

Let $t_0, t_1, t_2 \in L$ be such that$(f, h)(t_0) = (a, b)$, $(g, h)(t_1) = (a, c)$, $(f, g, h)(t_2) = (a, b, c)$. By monotonicity of h, $t_0 < t_1$, and by monotonicity of g, $t_1 < t_2$. By monotonicity of f, then, $f(t_1) = a$, so $(a, a) \in \operatorname{im}(f, g)$.

As the desired property is true for length-3 indiscernible sequences in M_L, it is also true for all longer indiscernibles. □

Lemma 5.4.15. *Let L be a linear continuum with endpoints. Any indiscernible sequence in one variable in M_L is distal.*

Proof. By Lemma 3.5.3, it suffices to show that if $(f_i : i \in I)$ is a sequence of elements in M_L, and $i_0 < i_1$ are such that (i_0, i_1) is infinite and removing either f_{i_0} or f_{i_1} makes the sequence indiscernible, then $(f_i : i \in I)$ is indiscernible. To do this, we show that

$\mathrm{tp}(f_{i_0}, f_{i_1}) = \mathrm{tp}(f_i, f_j)$ for all other $i < j$. This will even apply for sequences of finite length - if $f_1, \ldots, f_5 \in M$ are such that (f_1, f_2, f_3, f_5) and (f_1, f_3, f_4, f_5) are indiscernible, then $\mathrm{tp}(f_2, f_4) = \mathrm{tp}(f_1, f_3)$, and this will imply indiscernibility for any infinite sequence containing these elements and satisfying the above properties.

Let $(a, b) \in \mathrm{im}(f_1, f_3)$. By Lemma 5.4.14, either $(a, a) \in \mathrm{im}(f_1, f_3)$ or $(b, b) \in \mathrm{im}(f_1, f_3)$. Without loss of generality, we may assume the former case. By Lemma 5.4.12 and indiscernibility, we see that $(a, a, b) \in \mathrm{im}(f_1, f_3, f_4)$, so there is some t with $(f_1, f_3, f_4)(t) = (a, a, b)$, and by Lemma 5.4.13, we have that $f_2(t) = a$ as well, so $(a, b) \in \mathrm{im}(f_2, f_4)$, so $\mathrm{tp}(f_2, f_4) = \mathrm{tp}(f_1, f_3)$ as desired. □

We now show that the interaction between tuples in any M_L can be coded by their averages.

Lemma 5.4.16. *Let L be a linear continuum with endpoints and let $\bar{f} = (f_1, \ldots, f_n)$ and $\bar{g} = (g_1, \ldots, g_n)$ be tuples in M_L. Define $\hat{f} = \frac{1}{n}\sum_{i=1}^n f_n$ and $\hat{g} = \frac{1}{n}\sum_{i=1}^n g_n$. Then $\mathrm{tp}(\bar{f}, \bar{g})$ is determined by $\mathrm{tp}(\bar{f}), \mathrm{tp}(\bar{g}), \mathrm{tp}(\hat{f}, \hat{g})$.*

Proof. Clearly $\mathrm{im}(\bar{f}, \hat{f})$ is determined by $\mathrm{im}(\bar{f})$. By the monotonicity of $\bar{f}$ and the surjectivity of $\hat{f}$, for any $a \in L$, there is exactly one $\bar{a} \in L^n$ such that $(\bar{a}, a) \in \mathrm{im}(\bar{f}, \hat{f})$. Thus $\mathrm{im}(\bar{f}, \bar{g})$ consists of all $(\bar{a}, \bar{b})$ such that if $\hat{a}, \hat{b}$ are the averages of $\bar{a}, \bar{b}$, then $\bar{a} \in \mathrm{im}(\bar{f}), \bar{b} \in \mathrm{im}(\bar{g})$, and $(\hat{a}, \hat{b}) \in \mathrm{im}(\hat{f}, \hat{g})$. □

Proposition 5.4.17. *Then the structure $M_{[0,1]}$ is distal.*

Proof. Let $(f_i : i \in I)$ be an indiscernible sequence in an elementary extension of $M_{[0,1]}$, which we may assume is M_L for some linear continuum L with endpoints. As in the proof of Lemma 5.4.15 but with longer tuples, we will just show that if $(f_1, \ldots, f_5) \in (M_L^n)^5$ is such that (f_1, f_2, f_3, f_5) and (f_1, f_3, f_4, f_5) are indiscernible, then the $\mathrm{tp}(f_2, f_4) = \mathrm{tp}(f_1, f_3)$.

Clearly for each $1 \leq i, j \leq 5$, $\mathrm{tp}(f_i) = \mathrm{tp}(f_j)$. For $1 \leq i \leq 5$, let $\hat{f}_i \in M_L$ be the pointwise average of the tuple f_i. By indiscernibility of these subsequences, we can deduce that $(\hat{f}_1, \ldots, \hat{f}_5)$ is indiscernible, so by the proof of Lemma 5.4.15, $\mathrm{tp}(\hat{f}_2, \hat{f}_4) = \mathrm{tp}(\hat{f}_1, \hat{f}_3)$. By Lemma 5.4.16, this constrains the types of the tuples enough that $\mathrm{tp}(f_2, f_4) = \mathrm{tp}(f_1, f_3)$. $\square$

We can say more about indiscernibles.

Theorem 5.4.18. *If $p \in S_2(\emptyset)$, then p is the type (f_i, f_j) with $i < j$ in some infinite indiscernible sequence $(f_i : i \in I)$ if and only if for all $(a, b) \in \mathrm{im}(p)$, either $(a, a) \in \mathrm{im}(p)$ or $(b, b) \in \mathrm{im}(p)$.*

Proof. Lemma 5.4.14 tells us that the type of any pair in an indiscernible has this property. Now assume that $p \in S_2(\emptyset)$ is such that for all $(a, b) \in \mathrm{im}(p)$, either $(a, a) \in \mathrm{im}(p)$ or $(b, b) \in \mathrm{im}(p)$. We will show that for any n, there are $f_1, \ldots, f_n \in M_{[0,1]}$ with $\mathrm{tp}(f_i, f_j) = p$ for all $i < j$, so by compactness, in some elementary extension, there is an infinite sequence $(f_i : i \in I)$ such that for all $i < j$, $\mathrm{tp}(f_i, f_j) = p$. By Lemma 5.4.12, $(f_i : i \in I)$ is indiscernible.

As a consequence of [Ben18, Theorem 3.2], there are $f, g \in M_{[0,1]}$ such that $\mathrm{tp}(f, g) = p$ and $\frac{f(t)+g(t)}{2} = t$ is the identity. Thus we may partition $[0, 1]$ into three disjoint sets, A_-, A_0, A_+, on which $f - g$ is respectively negative, 0, and positive, and note that A_-, A_+ are open while A_+ is closed. Thus also A_- and A_+ each consist of a countable number of open interval connected components.

We will define our functions f_i on A_0 and connected components of A_+, A_- separately, and we will define them so that $\frac{1}{n} \sum_{i=1}^{n} f_i(t) = t$ for all t. For each $1 \leq i \leq n$, if $t \in A_0$, then we let $f_i(t) = t$. Now let (a, b) be a connected component of A_+, and we will define $f_1, \ldots, f_n$ on $[a, b]$. As $f - g$ is positive on (a, b), we see that $a = g(a) \leq g\left(\frac{a+b}{2}\right) < f\left(\frac{a+b}{2}\right) \leq f(b) = b$. At least one of $f\left(\frac{a+b}{2}\right), g\left(\frac{a+b}{2}\right)$ is in A_0, and both are in $[a, b]$, so it must be either a or b. However, these numbers add to $a + b$, so they must be b and a respectively. By continuity and monotonicity, we see that the other values of $(c, d) \in \mathrm{im}(p)$ with $\frac{c+d}{2} \in [a, b]$ are exactly the points of the form $(t, a), (b, t)$ for $t \in [a, b]$. Thus the values of $(f_1, \ldots, f_n)$ on $[a, b]$ should all be of the form $(b, \ldots, b, t, a, \ldots, a)$ for $t \in [a, b]$.

It will thus suffice to define $f_1, \dots, f_n$ on $[a, b]$ such that

- for all i, f_i is continuous and monotone on $[a, b]$,
- for all i, $f_i(a) = a$ and $f_i(b) = b$,
- for all $i < j$, $t \in [a, b]$, either $f_i(t) = b$ or $f_j(t) = a$.

We define our functions on $[a, b]$ by breaking up $[a, b]$ into n subintervals of the form $\left[\frac{ia+(n-i)b}{n}, \frac{(i+1)a+(n-i-1)b}{n}\right]$, where $f_i(t) = a$ on $\left[a, \frac{ia+(n-i)b}{n}\right]$, $f_i(t)$ increases from a to b linearly on $\left[\frac{ia+(n-i)b}{n}, \frac{(i+1)a+(n-i-1)b}{n}\right]$, and $f_i(t) = b$ on $\left[\frac{(i+1)a+(n-i-1)b}{n}, b\right]$. If $i < j$, we see that for all $t \in [a, b]$, either $f_i(t) = b$ or $f_j(t) = a$, so we are done.

If instead (a, b) is a connected component of A_-, the functions can be defined similarly. As we have defined continuous, monotone functions on closed intervals covering $[0, 1]$ in a way that endpoints agree and any pair (f_i, f_j) with $i < j$ only takes values in $\mathrm{im}(p)$, we are done. $\square$

5.4.2 Another Language

We now propose a new language for this structure. Because by Lemma 5.4.12, all types are determined by their restrictions to pairs of variables, it suffices to choose predicate symbols that generate all definable predicates on two variables. By Stone-Weierstrass, it suffices to find a set of definable predicates on two variables that separates points on the type space S_2. For this, we may take the family $\{\phi_\alpha(x, y) : \alpha \in [0, 1] \cap \mathbb{Q}\}$, where when $(a, b) \in \mathrm{im}(\mathrm{tp}(f, g))$ is the unique point such that $a + b = \alpha$, $\phi_\alpha(f, g) = a$. It is clear that each of these is 1-Lipschitz, so define $\mathcal{L}$ to be the language consisting only of 1-Lipschitz binary predicates $\phi_\alpha(x, y)$ for $\alpha \in [0, 1] \cap \mathbb{Q}$. Because the image of a type, and thus the type itself, is determined entirely by the value of these atomic predicates, $\mathrm{Th}(M_{[0,1]})$ eliminates quantifiers in $\mathcal{L}$.

We can also axiomatize $\mathrm{Th}(M_{[0,1]})$ fairly easily in this language. For simplicity, we extend the language by quantifier-free definitions to include $\phi_\alpha(x, y)$ for $\alpha \in [0, 1]$ by taking uniform limits.

Lemma 5.4.19. *The theory of $M_{[0,1]}$ is axiomatized by the following theory, which we describe with equations and inequalities for clarity:*

$$\begin{aligned}
&\{\phi_\alpha(x,y) + \phi_\alpha(y,x) = \alpha : \alpha \in [0,1]\} \\
\cup&\{\phi_\alpha(x,y) \leq \phi_\beta(x,y) : 0 \leq \alpha < \beta \leq 1\} \\
\cup&\{\inf_{x_1,\ldots,x_n} \bigvee_{0\leq k\leq m, i\neq j} |\phi_{c_k(i)+c_k(j)}(x_i,x_j) - c_k(i)| = 0 : c_0,\ldots,c_m \in [0,1]^n \textit{is a finite chain}\}
\end{aligned}$$

Proof. It suffices to require that for each pair of variables x, y, the set $\{(\phi_\alpha(x,y), \phi_\alpha(y,x)) : \alpha \in [0,1]\}$ forms a valid type in S_2, and to require that every type in each S_n is realized, which they are in all structures as the theory is separably categorical. First, we require that $\phi_\alpha(x,y) + \phi_\alpha(y,x) = \alpha$ with an axiom for each α. To check that $\{(\phi_\alpha(x,y), \phi_\alpha(y,x)) : \alpha \in [0,1]\}$ is a chain from 0 to 1, we add axioms ensuring that $\phi_\alpha(x,y) \leq \phi_\beta(x,y)$ for each $\alpha < \beta$. These also imply that $\{(\phi_\alpha(x,y), \phi_\alpha(y,x)) : \alpha \in [0,1]\}$ is connected, as the function $\alpha \mapsto (\phi_\alpha(x,y), \phi_\alpha(y,x))$ is 1-Lipschitz and thus continuous.

Now to ensure that each type is realized. For each connected chain $C \subseteq [0,1]^n$ from 0 to 1, and each $m \in \mathbb{N}$, let $c_0, \ldots, c_m$ be the points on C such that at c_i, $\sum_{i=1}^n x_i = \frac{in}{m}$. Then let $(a_1, \ldots, a_n)$ be such that $\{c_0, \ldots, c_m\} \subseteq \mathrm{im}(a_1, \ldots, a_n)$. For each $c \in \mathrm{im}(a_1, \ldots, a_n)$, there is some i with $c_i \leq c \leq c_{i+1}$, so $d(c, C) \leq d(c, c_i) \leq \frac{n}{m}$. Thus if

$$\bigvee_{0\leq k\leq m, i\neq j} |\phi_{c_k(i)+c_k(j)}(x_i,x_j) - c_k(i)| = 0$$

at a particular $(a_1, \ldots, a_n)$, we find that $d(\mathrm{tp}(a_1, \ldots, a_n), C) \leq \frac{n}{m}$ and if

$$\inf_{x_1,\ldots,x_n} \bigvee_{0\leq k\leq m, i\neq j} |\phi_{c_k(i)+c_k(j)}(x_i,x_j) - c_k(i)| = 0$$

for $(c_0, \ldots, c_m)$ for all m, the type with image C is realized. As $\{c_0, \ldots, c_m\}$ could be any chain from 0 to 1, and in fact this predicate will still be 0 for any finite chain, we simply require this for all finite chains. $\square$

We now consider distal cell decompositions in this language.

Definition 5.4.20. If $\phi(x; y)$ is a definable predicate, and Ψ is a finite set of definable predicates of the form $\psi(x; y_1, \ldots, y_k)$, then Ψ *weakly defines* a *ε-distal cell decomposition* over M for $\phi(x; y)$ when for every finite $B \subseteq M^y$ with $|B| \geq 2$ and every $a \in M^x$, there are $\psi \in \Psi$ and $b_1, \ldots, b_k \in M^x$ such that $\psi(a; b_1, \ldots, b_k) > 0$ and for all $a' \in M^x$, $\psi(a'; b_1, \ldots, b_k) > 0$ implies $|\phi(a; b) - \phi(a'; b)| \leq \varepsilon$ for all $b \in B$.

Theorem 5.4.21. *Each $\phi_\alpha(x; y)$ admits a ε-distal cell decomposition over $M_{[0,1]}$ for each $\varepsilon > 0$, which we construct explicitly.*

Proof. Let $B \subseteq M_{[0,1]}$ be finite with $|B| \geq 2$.

For each $0 \leq i \leq n$, let F_{i-} be a continuous function with support $\left[0, \frac{i}{n}\right)$, and let F_{i+} be a continuous function with support $\left(\frac{i}{n}, 1\right]$. We will show for each $0 < i < n$, there are there are some $\psi_{i-}(x), \psi_{i+}(x)$, with $\psi_{i-}(x)$ of the form either 1 or $F_{i-}(\phi_\alpha(x; b_-))$, and $\psi_{i+}(x)$ either of the form 1 or $F_{i+}(\phi_\alpha(x; b_+))$ with $b_-, b_+ \in B$, such that $\psi_{i\pm}(a) > 0$, while $\psi_{i-}(a') > 0$ implies $\phi_\alpha(a'; b) < \frac{i}{n}$ for each $b \in B$ with $\phi_\alpha(a'; b) < \frac{i}{n}$, and $\psi_{i+}(a') > 0$ implies $\phi_\alpha(a'; b) > \frac{i}{n}$ for each $b \in B$ with $\phi_\alpha(a'; b) > \frac{i}{n}$. Once we know this, we can let $\psi(x) = \bigwedge_{i=0}^{n} (\psi_{i-}(x) \wedge \psi_{i+}(x))$. Then $\psi(x)$ will be of the form $\psi(x; b_1, \ldots, b_k)$, where ψ is one of a finite set Ψ of formulas, and $b_1, \ldots, b_k$. This set Ψ weakly defines a $\frac{2}{n}$-distal cell decomposition, because $\psi(a) > 0$, and for every $b \in B$, there is some i such that $\frac{i}{n} < \phi_\alpha(a; b) < \frac{i+2}{n}$, so $\psi(a'; b_1, \ldots, b_k) > 0$ implies $\frac{i}{n} < \phi_\alpha(a'; b) < \frac{i+2}{n}$, and thus $|\phi_\alpha(a; b) - \phi_\alpha(a'; b)| \leq \frac{2}{n}$.

By symmetry, it suffices to construct $\psi_{i-}(x)$. Let $b_- \in B$ maximize $\sup(b_-^{-1}(\{\alpha - \frac{i}{n}\}))$ under the constraint that $\phi_\alpha(a; b) < \frac{i}{n}$. If there is not some b_- satisfying this constraint, then we simply let $\psi_{i-}(x) = 1$, the rest of the requirements are trivial. If it does exist, then we let $\psi_{i-}(a) = F_{i-}(\phi_\alpha(a; b_-))$, and by construction, $F_{i-}(\phi_\alpha(a; b_-)) > 0$.

If $a' \in M^x, b \in B$, we claim that $\phi_\alpha(a';b) < \frac{i}{n}$ if and only if the interval $b^{-1}((\alpha - \frac{i}{n}, 1])$ intersects the interval $a'^{-1}([0, \frac{i}{n}))$. Let t_α be such that $a'(t_\alpha) + b(t_\alpha) = \alpha$. If $\phi_\alpha(a';b) < \frac{i}{n}$, then these intervals intersect at t_α. If these intervals intersect at some t, then we know that $a'(t) < \frac{i}{n}$ and $b(t) > \alpha - \frac{i}{n}$. If $a'(t) + b(t) < \alpha$, then $t < t_\alpha$, and thus $b(t_\alpha) \geq b(t) > \alpha - \frac{i}{n}$, so $\phi_\alpha(a';b) < \frac{i}{n}$, and similarly if $a'(t) + b(t) > \alpha$, then $t > t_\alpha$, so $\phi_\alpha(a';b) = a'(t_\alpha) < \frac{i}{n}$. If $a'(t) + b(t) = \alpha$, then $\phi_\alpha(a';b) = a'(t) < \frac{i}{n}$.

Now assume $\psi_{i-}(a') > 0$, $\phi_\alpha(a;b) < \frac{i}{n}$, - we wish to show that $\phi_\alpha(a';b) < \frac{i}{n}$. Because $\psi_{i-}(a') > 0$, we see that $b_-^{-1}((\alpha - \frac{i}{n}, 1])$ intersects $a'^{-1}([0, \frac{i}{n}))$, and by the definition of b_-, because $\phi_\alpha(a;b) < \frac{i}{n}$, we know that $b_-^{-1}((\alpha - \frac{i}{n}, 1]) \subseteq b^{-1}((\alpha - \frac{i}{n}, 1])$. Thus $b^{-1}((\alpha - \frac{i}{n}, 1])$ intersects $a'^{-1}([0, \frac{i}{n}))$ and $\phi_\alpha(a';b) < \frac{i}{n}$. $\square$

5.4.3 Nondiscreteness

Dual linear continua provide the best example of distal metric structures that are truly different from distal discrete structures. There are several possible criteria for determining whether a metric structure is non-discrete, and [Han23] compares several of these. Of these, the strongest, there denoted as $\star$, is defined as follows:

Definition 5.4.22. A metric structure has the property $\star$ when for any small partial type $\Sigma(x)$ (in finitely many variables), the metric space of realizations of $\Sigma(x)$ in the monster has a bounded number of connected components.

Theorem 5.4.23. *Dual linear continua have property $\star$.*

Proof. By Theorem 5.4.11, the monster model is isomorphic to M_L for some linear continuum with endpoints L - we shall assume that it is indeed M_L. By [Han23, Theorem 3.1], to check $\star$ it suffices to check that the space of realizations of complete types over small models are connected. Thus let $M \preceq M_L$ be a small model and let $p(x_1, \ldots, x_n)$ be a complete M-type. There are unique functions $h_1, \ldots, h_n : [0,1] \to [0,1]$ such that for any realization $(f_1, \ldots, f_n)$ of p, with $f = \frac{1}{n}\sum_{i=1}^n f_i$, we have $f_i = h_i \circ f$. There is also a unique complete M-type $q(x)$

of averages of realizations of p. We see that $f \mapsto (h_1, \ldots, h_n) \circ f$ is a function from the space of realizations of q to the space of realizations of p, and is continuous with respect to the sup metric, so it suffices to show that the space of realizations of q, in one variable, is connected. In fact, we will show that it is convex, and thus path-connected.

Suppose $f, g \in M_L$ are both realizations of q. It suffices to show that for $\lambda \in [0, 1]$, $\mathrm{tp}((1-\lambda)f + \lambda g/M) = q$. In fact, by Lemma 5.4.12, it suffices to check that for each $a \in M$, $\mathrm{tp}((1-\lambda)f + \lambda g, a)$, or equivalently $\mathrm{im}((1-\lambda)f + \lambda g, a)$, does not depend on λ. For each $c \in [0, 1]$, both f and g obtain the same closed interval of values on the preimage $a^{-1}(\{c\})$. Thus for any $t \in a^{-1}(\{c\})$, we have that $(1-\lambda)f(t) + \lambda g(t)$ is also in this interval, so $((1-\lambda)f(t) + \lambda g(t), a(t)) \in \mathrm{im}((f, a))$, implying that $\mathrm{im}((1-\lambda)f + \lambda g, a) = \mathrm{im}((f, a))$ for all λ. $\square$

5.5 Nonexamples

In discrete logic, there is an open question as to which NIP structures admit distal expansions. The Strong Erdős-Hajnal property is one requirement for admitting a distal expansion, and we have shown that this is still required in continuous logic, but little else is known. In continuous logic, however, we can see a wide class of NIP structures which cannot admit distal expansions for a seemingly different reason: Banach structures. We thank James Hanson for pointing this out.

Definition 5.5.1. A *Banach structure* is an expansion of a Banach space, viewed as a metric structure. The theory of a Banach structure is called a *Banach theory.*

It will be easy to see that many of these are not distal, because of the following fact:

Fact 5.5.2 ([Han20b, Corollary 6.10])**.** *Every Banach theory with infinite dimensional models has an infinite indiscernible set in some model.*

Corollary 5.5.3. *No Banach theory with infinite dimensional models is distal.*

Proof. This is true because no distal structure has an infinite indiscernible set, by the same proof as in discrete logic:

If it did, we could partition such a set into two infinite subsets I, J and an extra element, d. Then by the indiscernibilty of the overall set, $I + J$ is indiscernible over d, and thus by distality, $I + d + J$ is indiscernible over d, implying that every element of $I + d + J$ satisfies $x = d$. This clearly contradicts the set being infinite. $\square$

One particularly interesting class of Banach structures is randomizations. If T is a metric theory, then the *randomization theory* T^R of T can be constructed in a few ways, each of which captures the idea that a model of T^R consists of random variables valued in models of T. The construction in [Ben13] adds to the sorts of T an extra sort, consisting of an algebra of random variables, which is an L_1-space, and thus is a Banach structure. While the randomization of a stable theory is stable ([Ben13, Theorem 4.9]) and the randomization of an NIP theory is NIP ([Ben09, Theorem 5.3]), we see that the same is not true of distality, as the randomization of *any* structure is not distal. Restricting to the original sorts of T will not change this, as the random variable sort is interpretable from the induced structure on the other sorts.

REFERENCES

[AB24] Aaron Anderson and Itaï Ben Yaacov. "Examples and Nonexamples of Distal Metric Structures." Forthcoming, 2024.

[ABC97] Noga Alon, Shai Ben-David, Nicolo Cesa-Bianchi, and David Haussler. "Scale-sensitive dimensions, uniform convergence, and learnability." *Journal of the ACM (JACM)*, **44**(4):615–631, 1997.

[ACG22] Matthias Aschenbrenner, Artem Chernikov, Allen Gehret, and Martin Ziegler. "Distality in valued fields and related structures." *Transactions of the American Mathematical Society*, **375**(7):4641–4710, 2022.

[ADH16] Matthias Aschenbrenner, Alf Dolich, Deirdre Haskell, Dugald Macpherson, and Sergei Starchenko. "Vapnik-Chervonenkis density in some theories without the independence property, I." *Transactions of the American Mathematical Society*, **368**(8):5889–5949, 2016.

[AHH22] Noga Alon, Steve Hanneke, Ron Holzman, and Shay Moran. "A theory of PAC learnability of partial concept classes." In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 658–671. IEEE, 2022.

[AK92] Noga Alon and Daniel J Kleitman. "Piercing convex sets and the Hadwiger-Debrunner (p, q)-problem." *Advances in Mathematics*, **96**(1):103–112, 1992.

[And23a] Aaron Anderson. "Combinatorial Bounds in Distal Structures." *The Journal of Symbolic Logic*, pp. 1–29, 2023.

[And23b] Aaron Anderson. "Fuzzy VC Combinatorics and Distality in Continuous Logic." *arXiv preprint arXiv:2310.04393*, 2023.

[And23c] Aaron Anderson. "Generically Stable Measures and Distal Regularity in Continuous Logic." *arXiv preprint arXiv:2310.06787*, 2023.

[Bar13] Salvador Barone. "Some quantitative results in real algebraic geometry." *Thesis*, 2013.

[Bas09] Saugata Basu. "Combinatorial complexity in o-minimal geometry." *Proceedings of the London Mathematical Society*, **100**(2):405–428, 08 2009.

[BBH08] Itaï Ben Yaacov, Alexander Berenstein, C. Ward Henson, and Alexander Usvyatsov. "Model theory for metric structures." *London Mathematical Society Lecture Note Series*, **350**:315, 2008.

[Ben09] Itaï Ben Yaacov. "Continuous and random Vapnik-Chervonenkis classes." *Israel Journal of Mathematics*, **173**(1):309–333, September 2009.

[Ben10a] Itaï Ben Yaacov. "Definability of groups in $\aleph_0$-stable metric structures." *The Journal of Symbolic Logic*, **75**(3):817–840, 2010.

[Ben10b] Itaï Ben Yaacov. "Stability and stable groups in continuous logic." *The Journal of Symbolic Logic*, **75**(3):1111–1136, 2010.

[Ben13] Itaï Ben Yaacov. "On theories of random variables." *Israel Journal of Mathematics*, **194**:957–1012, 2013.

[Ben14] Itaï Ben Yaacov. "Model theoretic properties of metric valued fields." *The Journal of Symbolic Logic*, **79**(3):655–675, 2014.

[Ben18] Itaï Ben Yaacov. "On a Roelcke-precompact Polish group that cannot act transitively on a complete metric space." *Israel Journal of Mathematics*, **224**(1):105–132, 2018.

[BK09] Itaï Ben Yaacov and H. Jerome Keisler. "Randomizations of models as metric structures." *Confluentes Mathematici*, **1**(02):197–223, 2009.

[BKS24] Martin Bays, Itay Kaplan, and Pierre Simon. "Density of compressible types and some consequences." *Journal of the European Mathematical Society*, pp. 1–43, 2024.

[BKW10] Martin Bays, Jonathan Kirby, and A. J. Wilkie. "A Schanuel property for exponentially transcendental powers." *Bulletin of The London Mathematical Society*, **42**:917–922, 2010.

[Bob17] Anton Bobkov. "Computations of Vapnik-Chervonenkis Density in Various Model-Theoretic Structures." *UCLA Electronic Theses and Dissertations*, 2017.

[Bou07] Nicolas Bourbaki. *Fonctions d'une variable réelle*. Springer Berlin Heidelberg, 2007.

[BPW00] Oleg Belegradek, Ya'acov Peterzil, and Frank Wagner. "Quasi-O-Minimal Structures." *The Journal of Symbolic Logic*, **65**(3):1115–1132, 2000.

[BS16] Saugata Basu and Martín Sombra. "Polynomial Partitioning on Varieties of Codimension Two and Point-Hypersurface Incidences in Four Dimensions." *Discrete & Computational Geometry*, **55**, 01 2016.

[BT16] Itaï Ben Yaacov and Todor Tsankov. "Weakly almost periodic functions, model-theoretic stability, and minimality of topological groups." *Transactions of the American Mathematical Society*, **368**(11):8267–8294, 2016.

[BU10] Itaï Ben Yaacov and Alexander Usvyatsov. "Continuous first order logic and local stability." *Transactions of the American Mathematical Society*, **362**(10):5213–5259, 2010.

[CCP24] Nicolas Chavarria, Gabriel Conant, and Anand Pillay. "Continuous stable regularity." *Journal of the London Mathematical Society*, **109**(1):e12822, 2024.

[CEG89] Bernard Chazelle, Herbert Edelsbrunner, Leonidas J. Guibas, and Micha Sharir. "A singly-exponential stratification scheme for real semi-algebraic varieties and its applications." In Giorgio Ausiello, Mariangiola Dezani-Ciancaglini, and Simonetta Ronchi Della Rocca, editors, *Automata, Languages and Programming*, pp. 179–193, Berlin, Heidelberg, 1989. Springer Berlin Heidelberg.

[CEG91] Bernard Chazelle, Herbert Edelsbrunner, Leonidas J. Guibas, and Micha Sharir. "A singly exponential stratification scheme for real semi-algebraic varieties and its applications." *Theoretical Computer Science*, **84**(1):77–105, 1991.

[CG20] Gabriel Conant and Kyle Gannon. "Remarks on generic stability in independent theories." *Annals of Pure and Applied Logic*, **171**(2):102736, 2020.

[CG21] Gabriel Conant and Kyle Gannon. "Associativity of the Morley product of invariant measures in NIP theories." *The Journal of Symbolic Logic*, **86**(3):1293–1300, 2021.

[CGH23a] Gabriel Conant, Kyle Gannon, and James Hanson. "Keisler measures in the wild." *Model Theory*, **2**(1):1–67, 2023.

[CGH23b] Gabriel Conant, Kyle Gannon, and James E Hanson. "Generic stability, randomizations, and NIP formulas." *arXiv preprint arXiv:2308.01801*, 2023.

[CGS20] Artem Chernikov, David Galvin, and Sergei Starchenko. "Cutting lemma and Zarankiewicz's problem in distal structures." *Selecta Mathematica*, **26**:1–27, 2020.

[CS89] Kenneth L. Clarkson and Peter W. Shor. "Applications of random sampling in computational geometry, II." *Discrete & Computational Geometry*, **4**(5):387–421, 1989.

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

[CS15] Artem Chernikov and Pierre Simon. "Externally definable sets and dependent pairs II." *Transactions of the American Mathematical Society*, **367**(7):5217–5235, 2015.

[CS18] Artem Chernikov and Sergei Starchenko. "Regularity lemma for distal structures." *Journal of the European Mathematical Society*, **20**(10):2437–2466, 2018.

[CS20] Artem Chernikov and Sergei Starchenko. "Model-Theoretic Elekes-Szabó in the strongly minimal case." *Journal of Mathematical Logic*, 07 2020.

[CS21] Artem Chernikov and Sergei Starchenko. "Definable regularity lemmas for NIP hypergraphs." *The Quarterly Journal of Mathematics*, **72**(4):1401–1433, 2021.

[CT20] Artem Chernikov and Henry Towsner. "Hypergraph regularity and higher arity VC-dimension." *arXiv preprint arXiv:2010.00726*, 2020.

[CV71] Alexey Chervonenkis and Vladimir Vapnik. "Theory of uniform convergence of frequencies of events to their probabilities and problems of search for an optimal solution from empirical data(Average risk minimization based on empirical data, showing relationship of problem to uniform convergence of averages toward expectation value)." *Automation and Remote Control*, **32**:207–217, 1971.

[DD88] Jan Denef and Lou van den Dries. "p-adic and Real Subanalytic Sets." *Annals of Mathematics*, **128**(1):79–138, 1988.

[DHM99] Lou van den Dries, Deirdre Haskell, and Dugald MacPherson. "One-Dimensional p-Adic Subanalytic Sets." *Journal of the London Mathematical Society*, **59**(1):1–20, 1999.

[EK21] Shlomo Eshel and Itay Kaplan. "On uniform definability of types over finite sets for NIP formulas." *Journal of Mathematical Logic*, **21**(03):2150015, 2021.

[FPS17] Jacob Fox, János Pach, Adam Sheffer, Andrew Suk, and Joshua Zahl. "A semi-algebraic version of Zarankiewicz's problem." *Journal of the European Mathematical Society (EMS Publishing)*, **19**(6), 2017.

[Gan20] Kyle Gannon. *Approximation theorems for Keisler measures.* University of Notre Dame, 2020.

[GK15] Larry Guth and Nets Katz. "On the Erdős distinct distances problem in the plane." *Annals of Mathematics*, **181**(1):155–190, 2015.

[GKP99] Janusz Gwoździewicz, Krzysztof Kurdyka, and Adam Parusiński. "On the Number of Solutions of an Algebraic Equation on the Curve $y = e^x + \sin x, x > 0$, and a Consequence for O-Minimal Structures." *Proceedings of the American Mathematical Society*, **127**(4):1057–1064, 1999.

[Han20a] James Hanson. *Approximation theorems for Keisler measures.* University of Wisconsin, Madison, 2020.

[Han20b] James Hanson. "Strongly minimal sets and categoricity in continuous logic." *arXiv preprint arXiv:2011.00610*, 2020.

[Han23] James Hanson. "A simple continuous theory." *arXiv preprint arXiv:2306.14324*, 2023.

[HK23] Yatir Halevi and Itay Kaplan. "Saturated models for the working model theorist." *Bulletin of Symbolic Logic*, **29**(2):163–169, 2023.

[HP94] Ehud Hrushovski and Anand Pillay. "Groups definable in local fields and pseudo-finite fields." *Israel Journal of Mathematics*, **85**(1-3):203–262, February 1994.

[HP11] Ehud Hrushovski and Anand Pillay. "On NIP and invariant measures." *Journal of the European Mathematical Society (EMS Publishing)*, **13**(4), 2011.

[HPS13] Ehud Hrushovski, Anand Pillay, and Pierre Simon. "Generically stable and smooth measures in NIP theories." *Transactions of the American Mathematical Society*, **365**(5):2341–2366, 2013.

[Iba16] Tomás Ibarlucía. "The dynamical hierarchy for Roelcke precompact Polish groups." *Israel Journal of Mathematics*, **215**(2):965–1009, 2016.

[Kei87] H Jerome Keisler. "Measures and forking." *Annals of Pure and Applied Logic*, **34**(2):119–169, 1987.

[Kha22] Karim Khanaki. "Generic stability and modes of convergence." *arXiv preprint arXiv:2204.03910*, 2022.

[Kol04] Vladlen Koltun. "Almost Tight Upper Bounds for Vertical Decompositions in Four Dimensions." *J. ACM*, **51**(5):699–730, September 2004.

[KP22] Krzysztof Krupiński and Adrián Portillo. "On stable quotients." *Notre Dame Journal of Formal Logic*, **63**(3):373–394, 2022.

[Lee12] Eva Leenknegt. "Cell decomposition for semiaffine structures on p-adic fields." *Journal of Logic and Analysis*, **4**, 08 2012.

[Lee14] Eva Leenknegt. "Reducts of p-adically closed fields." *Archive for Mathematical Logic*, **53**(3):285–306, 2014.

[LP93] James Loveys and Ya'acov Peterzil. "Linear o-Minimal Structures." *Israel Journal of Mathematics*, **81**:1–30, 1993.

[LS07] László Lovász and Balázs Szegedy. "Szemerédi's lemma for the analyst." *GAFA Geometric And Functional Analysis*, **17**:252–270, 2007.

[LS10] László Lovász and Balázs Szegedy. "Regularity partitions and the topology of graphons." *An Irregular Mind: Szemerédi is 70*, pp. 415–446, 2010.

[Mac76] Angus Macintyre. "On definable sets of p-adic numbers." *Journal of Symbolic Logic*, **41**(3):605–610, 1976.

[Mat02] Jiří Matoušek. *Lectures on Discrete Geometry.* Springer-Verlag, Berlin, Heidelberg, 2002.

[Mat04] Jiří Matoušek. "Bounded VC-dimension implies a fractional Helly theorem." *Discrete & Computational Geometry*, **31**:251–255, 2004.

[Mel10] Julien Melleray. "A note on Hjorth's oscillation theorem." *The Journal of Symbolic Logic*, **75**(4):1359–1365, 2010.

[Men02] Shahar Mendelson. "On the Size of Convex Hulls of Small Sets." *J. Mach. Learn. Res.*, **2**:1–18, mar 2002.

[Mou09] Marie-Hélène Mourgues. "Cell decomposition for P-minimal fields." *Mathematical Logic Quarterly*, **55**:487 – 492, 10 2009.

[MS14] Maryanthe Malliaris and Saharon Shelah. "Regularity lemmas for stable graphs." *Transactions of the American Mathematical Society*, **366**(3):1551–1585, 2014.

[PS98] Ya'acov Peterzil and Sergei Starchenko. "A trichotomy theorem for o-Minimal structures." *Proceedings of the London Mathematical Society*, **77**(3):481–523, 1998.

[RS80] M. Reed and B. Simon. *I: Functional Analysis.* Methods of Modern Mathematical Physics. Elsevier Science, 1980.

[Rud87] Walter Rudin. *Real and Complex Analysis, 3rd Ed.* McGraw-Hill, Inc., USA, 1987.

[SA10] Micha Sharir and Pankaj K. Agarwal. *Davenport-Schinzel Sequences and Their Geometric Applications.* Cambridge University Press, USA, 2010.

[Sau72] Norbert Sauer. "On the density of families of sets." *Journal of Combinatorial Theory, Series A*, **13**(1):145–147, 1972.

[She72] Saharon Shelah. "A combinatorial problem; stability and order for models and theories in infinitary languages." *Pacific Journal of Mathematics*, **41**(1):247–261, 1972.

[She22] Adam Sheffer. *Polynomial Methods and Incidence Theory.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2022.

[Sim13] Pierre Simon. "Distal and non-distal NIP theories." *Annals of Pure and Applied Logic*, **164**(3):294–318, 2013.

[Sim15] Pierre Simon. *A guide to NIP theories.* Cambridge University Press, 2015.

[Sim16] Pierre Simon. "A note on "Regularity lemma for distal structures"." *Proceedings of the American Mathematical Society*, **144**(8):3573–3578, 2016.

[Sim19] Pierre Simon. "Type decomposition in NIP theories." *Journal of the European Mathematical Society*, **22**(2):455–476, 2019.

[SW99] H. H. Schaefer and M. P. Wolff. *Topological Vector Spaces.* Springer-Verlag, Berlin, Heidelberg, 1999.

[Sze75] Endre Szemerédi. "On sets of integers containing no k elements in arithmetic progression." *Acta Arith*, **27**(199-245):2, 1975.

[Wai19] Martin J Wainwright. *High-dimensional statistics: A non-asymptotic viewpoint*, volume 48. Cambridge university press, 2019.

[Wal20] Miguel Walsh. "The polynomial method over varieties." *Inventiones mathematicae*, **222**, 11 2020.

[Wil96] A. J. Wilkie. "Model Completeness Results for Expansions of the Ordered Field of Real Numbers by Restricted Pfaffian Functions and the Exponential Function." *Journal of the American Mathematical Society*, **9**(4):1051–1094, 1996.