PROPERLY ERGODIC STRUCTURES

NATHANAEL ACKERMAN, CAMERON FREER,
ALEX KRUCKMAN, AND REHANA PATEL

Abstract. We consider ergodic $\mathrm{Sym}(\mathbb{N})$-invariant probability measures on the space of L-structures with domain $\mathbb{N}$ (for L a countable relational language), and call such a measure a *properly ergodic structure* when no isomorphism class of structures is assigned measure 1. We characterize those theories in countable fragments of $\mathcal{L}_{\omega_1,\omega}$ for which there is a properly ergodic structure concentrated on the models of the theory. We show that for a countable fragment F of $\mathcal{L}_{\omega_1,\omega}$ the almost-sure F-theory of a properly ergodic structure has continuum-many models (an analogue of Vaught's Conjecture in this context), but its full almost-sure $\mathcal{L}_{\omega_1,\omega}$-theory has no models. We also show that, for an F-theory T, if there is some properly ergodic structure that concentrates on the class of models of T, then there are continuum-many such properly ergodic structures.

1. Introduction

Symmetric random constructions of mathematical structures have been extensively studied in probability theory, combinatorics, and logic. One of the

---

October 25, 2017

best-known examples is the countably infinite Erdős–Rényi random graph, whose edges are determined by an independent coin flip for each pair of vertices. With probability 1, this process produces a particular countable graph up to isomorphism, known as the Rado graph (or random graph). The paper [AFP16] provided a characterization of those countable structures that can be produced via a symmetric random construction.

In the present paper we study the *properly ergodic structures*, namely those symmetric random constructions that do not give rise to a single mathematical structure, but rather spread their probability mass across many isomorphism classes of structures. To do so, we make use of tools from infinitary model theory and from probability theory, especially the Aldous–Hoover–Kallenberg representation of exchangeable structures.

1.1. **Ergodic structures.** Let L be a countable relational language, and write Str_L for the measurable space of L-structures with domain $\mathbb{N}$. We say that a probability measure on Str_L is *invariant* when it is invariant under the natural action (called the *logic action*) of the permutation group $\mathrm{Sym}(\mathbb{N})$ on Str_L. An invariant probability measure on Str_L can be thought of as a distribution on countable structures that does not depend on the labeling of the domain. The orbits of the logic action are the isomorphism classes of L-structures in Str_L.

An invariant probability measure μ on Str_L is *ergodic* when the null and co-null sets are the only Borel sets that are almost surely invariant. In other words, μ is ergodic if, whenever $\mu(X \triangle \sigma[X]) = 0$ for all $\sigma \in \mathrm{Sym}(\mathbb{N})$, we have $\mu(X) = 0$ or $\mu(X) = 1$. The ergodic invariant probability measures are extreme points in the space of invariant probability measures on Str_L, and any invariant probability measure can be decomposed as a mixture of ergodic ones. (For details, see, e.g., [Kal05, Lemma A1.2 and Theorem A1.3].) Hence when considering invariant probability measures on Str_L, it often suffices to restrict attention to the ergodic ones.

In fact, the ergodic invariant probability measures can be thought of as random symmetric analogues of model-theoretic structures, and so we call them *ergodic structures*. An ergodic structure μ determines the "almost-sure truth value" of every sentence of the infinitary logic $\mathcal{L}_{\omega_1,\omega}$, as the set of models for a sentence of $\mathcal{L}_{\omega_1,\omega}$ is an invariant Borel set in Str_L, and hence is assigned measure 0 or 1 by μ. Therefore every ergodic structure has a complete almost-sure theory, in $\mathcal{L}_{\omega_1,\omega}$ or in any fragment of $\mathcal{L}_{\omega_1,\omega}$.

The ergodic structures have several additional nice properties. The Aldous–Hoover–Kallenberg theorem implies that every invariant probability measure on Str_L can be represented as a random process that depends on independent sources of randomness at every finite subset of $\mathbb{N}$ (see §2.4 for more details). The ergodic structures are those invariant measures with dissociated representations, i.e., in which the random process does not depend on "global" randomness (formally, randomness indexed by the empty set in the representation). Equivalently, ergodic structures are those in which the behavior on disjoint finite subsets of $\mathbb{N}$ is independent.

Further, the ergodic structures are exactly those invariant measures which arise as a limit, in the weak topology, of measures obtained by uniformly sampling from a finite structure. If we restrict to a language with a single binary relation, a rich source of ergodic structures comes from those measures obtained by sampling a graphon [Lov12]. Just as graphons arise as limits of sequences of finite graphs which are convergent in the appropriate sense, ergodic structures can be viewed as limits of convergent sequences of finite L-structures. For details, see [Kru16, §1.2].

1.2. **Properly ergodic structures.** An ergodic structure is called *properly ergodic* when it does not assign measure 1 to any $\mathrm{Sym}(\mathbb{N})$-orbit, i.e., isomorphism class of L-structures. In fact, a properly ergodic structure must assign measure 0 to every $\mathrm{Sym}(\mathbb{N})$-orbit.

The paper [AFP17] characterized those F-theories (where F is a countable fragment of $\mathcal{L}_{\omega_1,\omega}$) that are the complete almost-sure F-theory of an ergodic structure. This characterization was in terms of trivial definable closure (see §2.2), generalizing the result in [AFP16] for the case of a single $\mathrm{Sym}(\mathbb{N})$-orbit (i.e., the non-properly ergodic case). In the present paper, we are interested in understanding which F-theories are the complete almost-sure F-theory of a properly ergodic structure.

The most well-known examples of ergodic structures (such as the Erdős–Rényi random graph described above) concentrate on a single isomorphism class, and indeed, it is not immediately obvious how to construct any properly ergodic structures. At the American Institute of Mathematics workshop on *Graph and Hypergraph Limits* in 2011, Omer Angel asked whether the distribution on countable graphs induced by a graphon can have more than one isomorphism class in its support. By [LS12], the distribution on countable graphs induced by a graphon is ergodic, and so this question is asking whether there are any properly ergodic structures that concentrate on the theory of graphs. During the workshop Gábor Kun provided an example of such a properly ergodic structure; see [PSN11, §2.3] for details. In fact, a properly ergodic graph was discovered somewhat earlier by Bonato and Janssen in [BJ11]; see Example 3.1 for details.

The case of properly ergodic structures has been further considered in [AFNP16], where a class of examples was constructed that concentrate on the sets of models of certain "approximately $\aleph_0$-categorical" first-order theories with trivial definable closure. One such class of examples is described in Example 3.2.

In the present paper we characterize those F-theories that are the complete almost-sure F-theory of a properly ergodic structure. Additionally, we show that for any properly ergodic structure μ, the complete almost-sure $\mathcal{L}_{\omega_1,\omega}$-theory of μ has no models (of any cardinality), but that for any countable fragment F, the complete almost-sure F-theory of μ has continuum-many models up to isomorphism. This can be viewed as an analogue of Vaught's Conjecture in the setting of ergodic structures.

In [AFKP17] it was shown that for every countable structure M, the number of ergodic structures concentrating on its isomorphism class is zero, one, or continuum. Moreover, the case of one only occurs when M is highly homogeneous, i.e., interdefinable with one of the five reducts of the rational linear order. We extend this result to the properly ergodic case, showing that if there exists a properly ergodic structure concentrating on the class of models of an F-theory T, then there are continuum-many properly ergodic structures concentrating on the class of models of T.

1.3. **Outline of paper.** Section 2 contains some basic definitions and results, including the notion of trivial definable closure, a particular form of Π_2 sentence that we call Π_2^-, and the Aldous–Hoover–Kallenberg representation.

In Section 3, we provide a number of examples of properly ergodic structures, which illustrate some of their key features.

In Section 4 we undertake a Morley–Scott analysis of an ergodic structure μ, based on Morley's proof [Mor70] that the number of isomorphism classes of countable models of a sentence of $\mathcal{L}_{\omega_1,\omega}$ is countable, $\aleph_1$, or $2^{\aleph_0}$. This gives us a notion of Scott rank for ergodic structures and, in the properly ergodic case, allows us to find a countable fragment F of $\mathcal{L}_{\omega_1,\omega}$ in which there is a formula $\chi(\overline{x})$ which is satisfied with positive probability (under instantiations of its parameters independently sampled from μ), but which picks out continuum-many F-types, each of which has probability 0 of being realized. The analogue of Vaught's Conjecture mentioned above is a corollary of this analysis.

In Section 5, we introduce the notion of a *rooted* model of a theory. A structure M is rooted if every collection of non-isolated types (e.g., the continuum-many types of measure 0 coming from the Morley–Scott analysis) has "few" realizations in M, in a sense that we will make precise. We use the Aldous–Hoover–Kallenberg theorem to show that a structure sampled from a properly ergodic measure is almost surely rooted.

In Section 6, given a theory T having trivial definable closure, we use a single rooted model of T to guide the construction, via an inverse limit, of a rooted Borel model $\mathbb{M} \models T$ equipped with an atomless probability measure ν. By sampling from $(\mathbb{M}, \nu)$, we obtain a properly ergodic structure μ that concentrates on the class of models of T. The inverse limit construction is a refinement of the methods from [AFP16], [AFNP16], and [AFP17], which in turn generalized a construction of Petrov and Vershik [PV10]. Further, we use a technique from [AFKP17] to rescale ν, obtaining continuum-many properly ergodic structures concentrating on the class of models of T.

Putting together the results of Sections 4–6, we obtain the characterization of the complete almost-sure F-theories of properly ergodic structures.

2. Preliminaries

2.1. **The space Str_L, infinitary logic, and ergodic structures.** Throughout this paper, let L be a countable relational language. We study invariant measures on the space Str_L of L-structures with domain $\mathbb{N}$. One could formulate this work in terms of arbitrary countable languages (which allow constant and function symbols), but it turns out that one does not lose much by working in the relational case — there are no ergodic structures in languages having constant symbols, and, in an ergodic structure, the interpretation of a function symbol must take some value among its inputs almost surely (see [AFP16, §§3–4]). For more details on how to translate results about invariant measures for countable relational languages to the case of arbitrary countable languages, see [AFP17].

Definition 2.1. Str_L is the space of L-structures with domain $\mathbb{N}$. The topology is generated by the sets of the form $[\![R(\overline{a})]\!] = \{M \in \mathrm{Str}_L \mid M \models R(\overline{a})\}$ and $[\![\neg R(\overline{a})]\!] = \{M \in \mathrm{Str}_L \mid M \models \neg R(\overline{a})\}$, where R ranges over the relation symbols in L and $\overline{a}$ ranges over the $\mathrm{ar}(R)$-tuples from $\mathbb{N}$.

A structure $M \in \mathrm{Str}_L$ is uniquely determined by whether or not, for each relation symbol R in L of arity $\mathrm{ar}(R)$ and each $\mathrm{ar}(R)$-tuple $\overline{a}$ from $\mathbb{N}$,

$$M \models R(\overline{a})$$

holds. It follows that Str_L is homeomorphic to the Cantor space

$$\prod_{R\in L} 2^{(\mathbb{N}^{\mathrm{ar}(R)})}.$$

Recall that $\mathcal{L}_{\omega_1,\omega}$ is the infinitary extension of first-order logic obtained by allowing, as new formula-building operations, the conjunction or disjunction of any countable ($< \omega_1$) family of formulas with a common finite ($< \omega$) set of free variables. We ensure that all our variables come from a fixed countable supply. For a reference on $\mathcal{L}_{\omega_1,\omega}$, see [KK04].

In contrast to the infinitary logic $\mathcal{L}_{\omega_1,\omega}$, we will also be interested in the quantifier-free fragment of first-order logic, in which the only formula-building operations are negation, finite conjunction, and finite disjunction. Throughout this paper, when we speak of quantifier-free formulas and types, we mean quantifier-free first-order formulas and types.

Given a formula $\varphi(\overline{x}) \in \mathcal{L}_{\omega_1,\omega}$ and a tuple $\overline{a}$ from $\mathbb{N}$ of the same length as $\overline{x}$, we let

$$[\![\varphi(\overline{a})]\!] = \{M \in \mathrm{Str}_L \mid M \models \varphi(\overline{a})\}.$$

Every $[\![\varphi(\overline{a})]\!]$ is a Borel set in Str_L. Indeed, the formula-building operations of negation and countable conjunction and disjunction correspond to the set-building operations of complementation and countable intersection and union, and quantifiers over the countable domain also correspond to certain countable intersections and unions:

$$[\![\forall x\, \varphi(\overline{a}, x)]\!] = \bigcap_{b\in\mathbb{N}} [\![\varphi(\overline{a}, b)]\!]$$

$$[\![\exists x\, \varphi(\overline{a}, x)]\!] = \bigcup_{b\in\mathbb{N}} [\![\varphi(\overline{a}, b)]\!].$$

Restricting ourselves to finite Boolean operations, $[\![\varphi(\overline{a})]\!]$ is a clopen set when $\varphi(\overline{x})$ is a quantifier-free formula. In fact, by compactness, every clopen set in Str_L has the form $[\![\varphi(\overline{a})]\!]$ for some quantifier-free formula φ.

Let $\mathrm{Sym}(\mathbb{N})$ denote the permutation group of $\mathbb{N}$.

Definition 2.2. The **logic action** is the natural action of $\mathrm{Sym}(\mathbb{N})$ on Str_L, given by permuting the underlying set. Namely, for $\sigma \in \mathrm{Sym}(\mathbb{N})$ and $M \in \mathrm{Str}_L$, we have

$$\sigma(M) \models R(a_1, \ldots, a_n) \quad \text{if and only if} \quad M \models R\big(\sigma^{-1}(a_1), \ldots, \sigma^{-1}(a_n)\big)$$

for all $R \in L$.

Note that $\sigma(M) = N$ if and only if $\sigma\colon M \to N$ is an isomorphism of L-structures, so the orbit of a point $M \in \mathrm{Str}_L$ under the logic action is the set of all structures in Str_L which are isomorphic to M. We recall Scott's theorem, which says that this set is definable by a sentence of $\mathcal{L}_{\omega_1,\omega}$.

Theorem 2.3 (Scott, [Mar02, Theorem 2.4.15])**.** *For any countable structure M, there is a sentence φ_M of $\mathcal{L}_{\omega_1,\omega}$, the **Scott sentence of** M, such that for all countable structures N, we have $N \models \varphi_M$ if and only if $N \cong M$.*

We are now able to define the class of invariant measures on Str_L, and specifically, the ergodic and properly ergodic ones.

Definition 2.4. Let μ be a Borel probability measure on Str_L. We say that μ is **invariant** (under the logic action) if, for every Borel set X and every $\sigma \in \mathrm{Sym}(\mathbb{N})$, we have $\mu(\sigma[X]) = \mu(X)$.

Now suppose that μ is invariant. A Borel set X is **almost surely invariant** if $\mu(X \triangle \sigma[X]) = 0$ for all $\sigma \in \mathrm{Sym}(\mathbb{N})$. We say that μ is **ergodic** if, for every almost surely invariant Borel set X, either $\mu(X) = 0$ or $\mu(X) = 1$. Following terminology from [BM00, §I.2] and elsewhere, we say that μ is **properly ergodic** if $\mu(X) = 0$ for every orbit X of the logic action.

Definition 2.5. An **ergodic structure** is an ergodic invariant probability measure on Str_L.

This definition takes on a more concrete character if we restrict our attention to the measures assigned to instances of quantifier-free formulas. The following proposition is an application of the Hahn–Kolmogorov measure extension theorem [Tao11, Theorem 1.7.8, Exercise 1.7.7].

Proposition 2.6. *Let $\mathcal{B}^*$ be the Boolean algebra of clopen sets in* Str_L *(so $\mathcal{B}^*$ consists of those sets of the form $[\![\varphi(\overline{a})]\!]$, where φ is a quantifier-free formula and $\overline{a}$ is a tuple from $\mathbb{N}$). Any finitely additive measure μ^* on $\mathcal{B}^*$ extends to a unique Borel probability measure μ on Str_L. Moreover, μ is invariant if and only if μ^* is; that is, if and only if $\mu^*([\![\varphi(\overline{a})]\!]) = \mu^*([\![\varphi(\sigma(\overline{a}))]\!])$ for any $\sigma \in \mathrm{Sym}(\mathbb{N})$.*

Remark 2.7. Additionally, it follows from Theorem 2.28 below that an invariant measure μ on Str_L is ergodic if and only if the quantifier-free types of disjoint tuples from $\mathbb{N}$ are independent. That is, whenever $\varphi(\overline{x})$ and $\psi(\overline{y})$ are quantifier-free formulas and $\overline{a}$ and $\overline{b}$ are disjoint tuples from $\mathbb{N}$ whose lengths are those of $\overline{x}$ and $\overline{y}$ respectively, we have $\mu([\![\varphi(\overline{a}) \wedge \psi(\overline{b})]\!]) = \mu([\![\varphi(\overline{a})]\!])\, \mu([\![\psi(\overline{b})]\!])$.

For the remainder of this subsection, let μ be an ergodic structure.

Remark 2.8. If $\varphi(\overline{x})$ is a formula of $\mathcal{L}_{\omega_1,\omega}$ and $\overline{a}$ is a tuple of distinct elements of $\mathbb{N}$ (of the same length as $\overline{x}$), then, since μ is invariant under the logic action, the value $\mu([\![\varphi(\overline{a})]\!])$ is independent of the choice of $\overline{a}$. For convenience, we denote this quantity by $\mu(\varphi(\overline{x}))$, and refer to it as the *measure of the formula* φ. Note that under this convention, if $\varphi(\overline{x})$ implies $x_i = x_j$ for some $i \neq j$, then $\mu(\varphi(\overline{x})) = 0$.

Definition 2.9. If φ is a sentence of $\mathcal{L}_{\omega_1,\omega}$, we say μ **almost surely satisfies** φ, or μ **concentrates on** φ, if $\mu(\varphi) = 1$. We write $\mu \models \varphi$, and we set

$$\mathrm{Th}(\mu) = \{\varphi \in \mathcal{L}_{\omega_1,\omega} \mid \mu \models \varphi\}.$$

Similarly, if Σ is a set of sentences of $\mathcal{L}_{\omega_1,\omega}$, we write $\mu \models \Sigma$ if $\mu \models \varphi$ for all $\varphi \in \Sigma$, and say that μ is an **ergodic model** of Σ.

The following result is a connection between infinitary logic and ergodic invariant measures; see also [AFP17].

Proposition 2.10. $\mathrm{Th}(\mu)$ *is a complete and countably consistent theory of $\mathcal{L}_{\omega_1,\omega}$. That is, for every sentence φ of $\mathcal{L}_{\omega_1,\omega}$, $\varphi \in \mathrm{Th}(\mu)$ or $\neg\varphi \in \mathrm{Th}(\mu)$, and every countable subset $\Sigma \subseteq \mathrm{Th}(\mu)$ has a model.*

Proof. For any sentence φ, the set $[\![\varphi]\!]$ is an invariant Borel set. In particular, it is almost surely invariant, so by ergodicity, $\mu(\varphi) = 0$ or 1, and hence $\mu \models \varphi$ or $\mu \models \neg\varphi$. Now let Σ be a countable subset of $\mathrm{Th}(\mu)$. Since a countable intersection of measure 1 sets has measure 1, $\mu(\bigwedge_{\varphi\in\Sigma} \varphi) = 1$. In particular, $[\![\bigwedge_{\varphi\in\Sigma} \varphi]\!]$ is non-empty. □

A special case of Definition 2.9 is when the sentence φ is a Scott sentence.

Definition 2.11. If M is a countable structure, we say that μ is **almost surely isomorphic to** M, or μ **concentrates on** M, if $\mu \models \varphi_M$, where φ_M is the Scott sentence of M; equivalently, μ assigns measure 1 to the orbit of M.

Remark 2.12. If μ is properly ergodic, then $\mathrm{Th}(\mu)$ contains $\neg\varphi_M$ for every countable structure M, and thus $\mathrm{Th}(\mu)$ has no countable models. A priori,

$\mathrm{Th}(\mu)$ may have uncountable models (Löwenheim–Skolem does not apply to complete theories of $\mathcal{L}_{\omega_1,\omega}$), but we will see later (Corollary 4.9) that this is not the case: $\mathrm{Th}(\mu)$ has no models of any cardinality. Nevertheless, as noted in Proposition 2.10, every countable subset of $\mathrm{Th}(\mu)$ has countable models. This suggests that we should restrict our attention to countable fragments of $\mathcal{L}_{\omega_1,\omega}$.

Definition 2.13. A **fragment** of $\mathcal{L}_{\omega_1,\omega}$ is a set of formulas which contains all atomic formulas and is closed under subformulas, finite Boolean combinations, quantification, and substitution of free variables (from the countable supply). If F is a fragment of $\mathcal{L}_{\omega_1,\omega}$, we set

$$\mathrm{Th}_F(\mu) = \{\varphi \in F \mid \mu \models \varphi\}.$$

A countable set of formulas Φ generates a countable fragment $\langle \Phi \rangle$, the least fragment containing this set. The minimal fragment $\mathrm{FO} := \langle \emptyset \rangle$ is first-order logic.

Definition 2.14. Let F be a countable fragment of $\mathcal{L}_{\omega_1,\omega}$.

- A set of sentences T is a (complete satisfiable) F**-theory** if T has a model and, for every sentence $\varphi \in F$, either $\varphi \in T$ or $\neg\varphi \in T$. Equivalently, there is a structure M for which $T = \{\psi \in F \mid M \models \psi\}$.
- A set of formulas $p(\overline{x})$ is an F**-type** if there is a structure M and a tuple $\overline{a}$ from M such that $p(\overline{x}) = \{\psi(\overline{x}) \in F \mid M \models \psi(\overline{a})\}$. We say that $\overline{a}$ **realizes** p in M.
- An F-type p is **consistent** with an F-theory T if it is realized in some model of T, and we write $S^n_F(T)$ for the set of F-types in n variables which are consistent with T.

Remark 2.15. The Löwenheim–Skolem theorem holds for countable fragments of $\mathcal{L}_{\omega_1,\omega}$ (see [KK04, Theorem 1.5.4]). Thus, if F is countable, every F-theory has a countable model and every F-type which is consistent with T is realized in a countable model of T.

Remark 2.16. If F is countable and p is an F-type, then we denote by $\theta_p(\overline{x})$ the conjunction of all the formulas in p, i.e., $\bigwedge_{\varphi\in p} \varphi(\overline{x})$. This is a formula of $\mathcal{L}_{\omega_1,\omega}$ (although not a formula of F in general), so it is assigned a measure by our ergodic structure μ, as described in Remark 2.8. We will write $\mu(p)$ as shorthand for $\mu(\theta_p(\overline{x}))$, and refer to this as the *measure of the type* p. This is the probability, according to μ, that any given tuple of distinct elements of $\mathbb{N}$ (of the appropriate arity) satisfies p.

2.2. **Trivial definable closure.** The paper [AFP17] shows that trivial definable closure is a necessary and sufficient condition for a theory (in a countable fragment) to have an an ergodic structure which satisfies it. Here we state several definitions and basic facts, and we provide a proof of one direction of this characterization.

Definition 2.17. Let F be a fragment of $\mathcal{L}_{\omega_1,\omega}$. An F-theory T has **trivial definable closure** (abbreviated **trivial dcl**) if there is no formula $\varphi(\overline{x}, y)$ in F such that

$$T \models \exists \overline{x}\, \exists! y \left(\left(\bigwedge_{i=1}^{n} y \neq x_i \right) \wedge \varphi(\overline{x}, y) \right).$$

Here $\exists!\, y$ is the standard abbreviation for "there exists a unique y".

Remark 2.18. If T is the complete F-theory of a structure M, then T has trivial dcl if and only if M has trivial dcl for the fragment F in the usual sense: $\mathrm{dcl}_F(A) = A$ for all $A \subseteq M$, where $\mathrm{dcl}_F(A)$ is the set of all $b \in M$ such that b is the unique element of M satisfying some formula in F with parameters from A.

If $\varphi(\overline{x}, y)$ witnesses that T has nontrivial dcl, then taking φ^* to be the formula $\varphi(\overline{x}, y) \wedge \exists^{\leq 1} y\, \varphi(\overline{x}, y)$ we have the stronger condition that T proves that φ^* is a definable function on some non-empty domain. That is,

$$T \models \left(\exists \overline{x}\, \exists y \left(\bigwedge_{i=1}^{n} y \neq x_i \right) \wedge \varphi^*(\overline{x}, y)\right) \wedge \left(\forall \overline{x}\, \exists^{\leq 1} y\, \varphi^*(\overline{x}, y) \right).$$

Here $\exists^{\leq 1} y$ is the standard abbreviation for "there is at most one y".

The following argument first appeared (in a slightly different setting) in [AFP16, Theorem 4.1]; as stated, this result is from [AFP17]. We include it here for completeness.

The key observation is the standard fact that if a measure is invariant under the action of some group G, then no positive-measure set can have infinitely many almost surely disjoint images under the action of G.

Theorem 2.19. *Let μ be an ergodic structure and F a fragment of $\mathcal{L}_{\omega_1,\omega}$. Then $\mathrm{Th}_F(\mu)$ has trivial dcl.*

Proof. Suppose there is a formula $\varphi(\overline{x}, y)$ in F such that

$$\mu \left(\exists \overline{x}\, \exists! y \left(\left(\bigwedge_{i=1}^{n} y \neq x_i \right) \wedge \varphi(\overline{x}, y) \right) \right) = 1.$$

Let $\psi(\overline{x}, y)$ be the formula $\left(\bigwedge_{i=1}^{n} y \neq x_i \right) \wedge \varphi(\overline{x}, y)$.

By countable additivity of μ, there is a tuple $\overline{a}$ from $\mathbb{N}$ such that

$$\mu(\llbracket \exists! y\, \psi(\overline{a}, y) \rrbracket) > 0.$$

Let $\theta(\overline{a})$ be the formula $\forall z_1\, \forall z_2\, (\psi(\overline{a}, z_1) \wedge \psi(\overline{a}, z_2) \to (z_1 = z_2))$, so that $\exists! y\, \psi(\overline{a}, y)$ is equivalent to

$$\exists y\, \psi(\overline{a}, y) \wedge \theta(\overline{a}).$$

Since this formula has positive measure, countable additivity again implies that there is some $b \in \mathbb{N} \setminus \overline{a}$ such that

$$\beta := \mu(\llbracket \psi(\overline{a}, b) \wedge \theta(\overline{a}) \rrbracket) > 0.$$

By invariance, for any $c \in \mathbb{N} \setminus \overline{a}$, we also have

$$\mu(\llbracket \psi(\overline{a}, c) \wedge \theta(\overline{a}) \rrbracket) = \beta.$$

But θ ensures that $\psi(\overline{a}, b) \wedge \theta(\overline{a})$ and $\psi(\overline{a}, c) \wedge \theta(\overline{a})$ are inconsistent when $b \neq c$, so, computing the measure of the disjoint union,

$$\mu\left(\bigcup_{b\in\mathbb{N}\setminus\overline{a}} [\![\psi(\overline{a}, b) \wedge \theta(\overline{a})]\!]\right) = \sum_{b\in\mathbb{N}\setminus\overline{a}} \beta = \infty,$$

which is impossible. □

In the language of §2.1, the main result of [AFP16] was a characterization of those countable structures M such that there exists an ergodic structure μ which is almost surely isomorphic to M. That characterization was given in terms of trivial "group-theoretic" dcl (where the group is $\mathrm{Aut}(M)$).

Definition 2.20. A countable structure M has **trivial group-theoretic dcl** if for any finite subset $A \subseteq M$ and element $b \in M \setminus A$, there is an automorphism $\sigma \in \mathrm{Aut}(M)$ such that $\sigma(a) = a$ for all $a \in A$, but $\sigma(b) \neq b$.

Theorem 2.21 ([AFP16, Theorem 1.1])**.** *Let M be a countable structure. There exists an ergodic structure concentrating on M if and only if M has trivial group-theoretic dcl.*

Remark 2.22. The method in [AFP16] of obtaining a measure via i.i.d. sampling from a Borel structure, which we use again in Section 6, always produces an ergodic measure. This was mentioned in passing in [AFP16], though not stated as part of the main theorem; for a proof, see [AFKP17, Proposition 2.24]. See also Theorem 2.28 and Lemma 6.2 below.

It is a consequence of Scott's Theorem (Theorem 2.3) that the notion of trivial group-theoretic dcl for a countable structure M is equivalent to the usual (syntactic) trivial dcl for $\mathrm{Th}_{F_M}(M)$ in an appropriate countable fragment F_M of $\mathcal{L}_{\omega_1,\omega}$. That is, given a finite subset A of M, an element $b \in M$ is fixed by all automorphisms fixing A pointwise if and only if there is a formula from F_M with parameters from A which uniquely defines b in M.

Unlike the group-theoretic notion of trivial dcl, which is defined for a given structure, the syntactic notion of trivial dcl (Definition 2.17) is defined for theories in arbitrary countable fragments, and so is the relevant notion for this paper.

2.3. Π_2^- **sentences.** It is a well-known fact, originally due to Chang [Cha68, pp. 48–49], that if T is a theory in a countable fragment F of $\mathcal{L}_{\omega_1,\omega}$, then the models of T are exactly the reducts to L of the models of a countable first-order theory T' in a larger countable language $L' \supseteq L$ that omit a countable set of types Q.

The idea is to Morleyize: we introduce a new relation symbol R_φ for every formula $\varphi(\overline{x})$ in F and encode the intended interpretations of the R_φ in the theory T'. The role of the countable set of types Q is to achieve this for infinitary conjunctions and disjunctions, which cannot be accounted for in first-order logic.

There are two features of this construction that will be useful for us. First, it reduces F-types to quantifier-free types. Second, T' can be axiomatized by Π_1 sentences together with pithy Π_2 sentences (also called "one point extension axioms").

Definition 2.23. A first-order sentence is **pithy** Π_2 if it has the form $\forall\overline{x}\,\exists y\,\varphi(\overline{x},y)$, where $\varphi(\overline{x},y)$ is quantifier-free, $\overline{x}$ is a tuple of variables (possibly empty), and y is a single variable. We call a sentence Π_2^- if it is either pithy Π_2 or is Π_1. A Π_2^- **theory** is a set of Π_2^- sentences.

Note that, in the context of this paper, all Π_2^- theories are first-order.

Theorem 2.24. *Let F be a countable fragment of $\mathcal{L}_{\omega_1,\omega}$ and T an F-theory. Then there is a language $L'\supseteq L$, an L'-theory T' that is Π_2^-, and a countable set of partial quantifier-free L'-types Q such that the following hold.*

(a) There is a bijection between formulas $\varphi(\overline{x})$ in F and atomic L'-formulas $R_\varphi(\overline{x})$ which are not in L, such that if $M\models T'$ omits all the types in Q, then $M\models\forall\overline{x}\,\varphi(\overline{x})\leftrightarrow R_\varphi(\overline{x})$.

(b) The reduct to L is a bijection between the class of models of T' omitting all the types in Q and the class of models of T.

Proof. Let $L' = L\cup\{R_\varphi \mid \varphi(\overline{x})\in F\}$, where the arity of the relation symbol R_φ is the length of the tuple $\overline{x}$. By convention, we allow 0-ary relation symbols (i.e., propositional symbols). Thus, we include a 0-ary relation R_ψ for every sentence $\psi\in F$.

Let T_{def} be the theory consisting of the following axioms, for each formula $\varphi(\overline{x})\in F$:

(1) $\forall\overline{x}\,\big(R_\varphi(\overline{x})\leftrightarrow\varphi(\overline{x})\big)$, if $\varphi(\overline{x})$ is atomic.
(2) $\forall\overline{x}\,\big(R_\varphi(\overline{x})\leftrightarrow\neg R_\psi(\overline{x})\big)$, if φ is of the form $\neg\psi(\overline{x})$.
(3) $\forall\overline{x}\,\big(R_\varphi(\overline{x})\leftrightarrow R_\psi(\overline{x})\wedge R_\theta(\overline{x})\big)$, if φ is of the form $\psi(\overline{x})\wedge\theta(\overline{x})$.
(4) $\forall\overline{x}\,\big(R_\varphi(\overline{x})\leftrightarrow R_\psi(\overline{x})\vee R_\theta(\overline{x})\big)$, if φ is of the form $\psi(\overline{x})\vee\theta(\overline{x})$.
(5) $\forall\overline{x}\,\big(R_\varphi(\overline{x})\rightarrow R_{\psi_i}(\overline{x})\big)$ for all $i\in I$, if φ is of the form $\bigwedge_{i\in I}\psi_i(\overline{x})$.
(6) $\forall\overline{x}\,\big(R_{\psi_i}(\overline{x})\rightarrow R_\varphi(\overline{x})\big)$ for all $i\in I$, if φ is of the form $\bigvee_{i\in I}\psi_i(\overline{x})$.
(7) $\forall\overline{x}\,\big(R_\varphi(\overline{x})\leftrightarrow\forall y\,R_\psi(\overline{x},y)\big)$, if φ is of the form $\forall y\,\psi(\overline{x},y)$.
(8) $\forall\overline{x}\,\big(R_\varphi(\overline{x})\leftrightarrow\exists y\,R_\psi(\overline{x},y)\big)$, if φ is of the form $\exists y\,\psi(\overline{x},y)$.

Note that all the axioms of T_{def} are first-order and universal except for those of type (7) and (8), which are Π_2^- when put in prenex normal form.

The axioms of type (5) and (6) cannot be made into bi-implications, since arbitrary countable infinite conjunctions and disjunctions are not expressible in first-order logic. To ensure that the corresponding R_φ have their intended interpretation, we let Q consist of the partial quantifier-free types:

(i) $q_\varphi(\overline{x}) = \{R_{\psi_i}(\overline{x}) \mid i\in I\}\cup\{\neg R_\varphi(\overline{x})\}$, for all $\varphi(\overline{x})$ of the form $\bigwedge_{i\in I}\psi_i(\overline{x})$
(ii) $q_\varphi(\overline{x}) = \{\neg R_{\psi_i}(\overline{x}) \mid i\in I\}\cup\{R_\varphi(\overline{x})\}$, for all $\varphi(\overline{x})$ of the form $\bigvee_{i\in I}\psi_i(\overline{x})$.

It is now straightforward to show by induction on the complexity of formulas that if a model $M \models T_{\mathrm{def}}$ omits every type in Q, then for all $\varphi(\overline{x})$ in F and all $\overline{a}$ from M, we have $M \models \varphi(\overline{a})$ if and only if $M \models R_\varphi(\overline{a})$. This establishes *(a)*. It also implies that every L-structure N admits a unique expansion to an L'-structure N' which is a model of T_{def} and omits every type in Q. As a consequence, if we set $T' = T_{\mathrm{def}} \cup \{R_\psi \mid \psi \in T\}$, then the following hold.

- If M is a model of T' which omits every type in Q, then the reduct $M \restriction L$ is a model of T.
- If $N \models T$, then the canonical expansion N' of N is a model of T'.
- If $M \models T'$ then $(M \restriction L)' = M$.
- If $N \models T$ then $N' \restriction L = N$.

This establishes *(b)*. □

Recall that an ergodic L-structure is an ergodic invariant measure on Str_L.

Corollary 2.25. *There is a bijection between the invariant measures on* Str_L *which almost surely satisfy* T *and the invariant measures on* $\mathrm{Str}_{L'}$ *which almost surely satisfy* T' *and omit all the types in* Q*. This bijection sends ergodic structures to ergodic structures and properly ergodic structures to properly ergodic structures.*

Proof. The reduct $\restriction_L$ is a continuous map $\mathrm{Str}_{L'} \to \mathrm{Str}_L$, since the preimages of clopen sets in Str_L are also clopen sets in $\mathrm{Str}_{L'}$. By Theorem 2.24, $\restriction_L$ is a bijection between the subspace X' of $\mathrm{Str}_{L'}$ consisting of models of T' which omit all the types in Q and the subspace X of Str_L consisting of models of T. Upon restricting to these subspaces, the inverse of $\restriction_L$ is a Borel map, since the image of a clopen set in X' (described by a quantifier-free formula) is a Borel set in X (described by a formula of $\mathcal{L}_{\omega_1,\omega}$). Hence $\restriction_L$ is a Borel isomorphism between these subspaces, and it induces a bijection between the set of probability measures on $\mathrm{Str}_{L'}$ concentrating on X' and the set of probability measures on Str_L concentrating on X. Moreover, $\restriction_L$ preserves the logic action, so the induced bijection on measures preserves invariance, ergodicity, and proper ergodicity. □

2.4. **The Aldous–Hoover–Kallenberg theorem and representations.** In this section, we state a version of the Aldous–Hoover–Kallenberg theorem. This theorem, which is a generalization of de Finetti's theorem to exchangeable arrays of random variables, was discovered independently by Aldous [Ald81] and Hoover [Hoo79], and further developed by Kallenberg [Kal92] and others. For proofs, we direct the reader to Kallenberg's book [Kal05, Chapter 7]. See [Ack15, §2.5] for a discussion of how to translate from the purely probabilistic statements in Kallenberg to the setting here, involving spaces of quantifier-free types. The survey by Austin [Aus08] provides details on its application to random structures.

We denote by $[n]$ the set $\{0, \ldots, n-1\}$, by $\mathbb{N}^{[n]}$ the set of n-tuples of *distinct* elements of $\mathbb{N}$ (that is, injective functions $[n] \to \mathbb{N}$), and by $\mathcal{P}_{\mathrm{fin}}(\mathbb{N})$ the set of all finite subsets of $\mathbb{N}$. Given a tuple $\overline{a} \in \mathbb{N}^{[n]}$, we denote by $\|\overline{a}\|$ the set in $\mathcal{P}_{\text{fin}}(\mathbb{N})$ enumerated by $\overline{a}$.

Let $S^n_{\text{qf}}(L)$ be the Stone space of quantifier-free n-types. Its points are the complete quantifier-free types in the variables $x_0, \dots, x_{n-1}$, and its topology is generated by the clopen sets $[\![\varphi(\overline{x})]\!] = \{p(\overline{x}) \in S^n_{\text{qf}}(L) \mid \varphi \in p\}$ for all quantifier-free formulas φ. Note that $S^n_{\text{qf}}(L)$ admits an action of the symmetric group $\text{Sym}(n)$ (the permutation group of $[n]$), by $\sigma(p(x_0, \dots, x_{n-1})) = p(x_{\sigma(0)}, \dots, x_{\sigma(n-1)})$ for $\sigma \in \text{Sym}(n)$. We write $S^{[n]}_{\text{qf}}(L)$ for the $\text{Sym}(n)$-invariant subspace of non-redundant quantifier-free types, namely those which contain $x_i \neq x_j$ for all $i \neq j$.

We let $(\xi_A)_{A \in \mathcal{P}_{\text{fin}}(\mathbb{N})}$ be a collection of independent random variables, each uniformly distributed on $[0, 1]$. We think of ξ_A as a source of randomness sitting on the subset A, which we will use to build a random L-structure with domain $\mathbb{N}$. If $\overline{a} \in \mathbb{N}^{[n]}$, the injective function $i\colon [n] \to \mathbb{N}$ enumerating $\overline{a}$ associates to each $X \in \mathcal{P}([n])$ a subset $i[X] \subseteq \|\overline{a}\|$. We denote by $\widehat{\xi_{\overline{a}}}$ the family of random variables $(\xi_{i[X]})_{X \in \mathcal{P}([n])}$.

Definition 2.26. An **AHK system** is a collection of measurable functions

$$(f_n\colon [0, 1]^{\mathcal{P}([n])} \to S^{[n]}_{\text{qf}}(L))_{n \in \mathbb{N}}$$

satisfying the coherence conditions:

- For all $\sigma \in \text{Sym}(n)$, almost surely
$$f_n((\xi_{\sigma[X]})_{X \subseteq [n]}) = \sigma(f_n((\xi_X)_{X \subseteq [n]})).$$
- For all $0 \leq m \leq n$, almost surely
$$f_m((\xi_X)_{X \subseteq [m]}) \subseteq f_n((\xi_Y)_{Y \subseteq [n]}).$$

That is, f_n takes as input a collection of values in $[0, 1]$, indexed by $\mathcal{P}([n])$, and produces a non-redundant quantifier-free n-type. Using our random variables ξ_A, we have a natural notion of **sampling** from an AHK system to obtain a non-redundant quantifier-free type $r_{\overline{a}} = f_n(\widehat{\xi_{\overline{a}}})$ for every finite tuple $\overline{a}$ from $\mathbb{N}$. Note that the order in which $\|\overline{a}\|$ is enumerated by the tuple $\overline{a}$ is significant, since f_n is, in general, not symmetric in its arguments.

The coherence conditions ensure that the quantifier-free types obtained from the function f_n cohere (almost surely), allowing us to define the random structure $\mathfrak{M}$ obtained by *sampling* from the AHK system (f_n). Namely, for every tuple $\overline{a} \in \mathbb{N}$,

$$\mathfrak{M} \models R(\overline{a}) \qquad \text{if and only if} \qquad R(\overline{x}) \in f_n(\widehat{\xi_{\overline{a}}}),$$

where n is the length of $\overline{a}$.

One may also directly describe the measure on Str_L which is the distribution of the random structure $\mathfrak{M}$: an AHK system $(f_n)_{n \in \mathbb{N}}$ gives rise to a well-defined finitely-additive probability measure μ^* on the Boolean algebra $\mathcal{B}^*$ of clopen sets in Str_L, defined by

$$\mu^*(\llbracket\varphi(\overline{a})\rrbracket) = \lambda^{\mathcal{P}([n])}(f_n^{-1}[\llbracket\varphi(\overline{x})\rrbracket]),$$

where $\lambda^{\mathcal{P}([n])}$ is the uniform product measure on $[0,1]^{\mathcal{P}([n])}$. This is the probability that $\varphi(\overline{x}) \in r_{\overline{a}}$, whenever $\overline{a}$ is a tuple of n distinct elements. The coherence conditions imply that this is well-defined: the first ensures that the order in which we list the variables in $\varphi(\overline{x})$ is irrelevant, and the second ensures that the measure is independent of the variable context $\overline{x}$.

Since the value of $\mu^*(\llbracket\varphi(\overline{a})\rrbracket)$ does not depend on the choice of tuple $\overline{a}$ of distinct elements, μ^* is manifestly invariant under the logic action. By Proposition 2.6, μ^* induces a unique invariant Borel probability measure μ on Str_L. In this case, we say that $(f_n)_{n\in\mathbb{N}}$ is an **AHK representation of** μ.

Theorem 2.27 (Aldous–Hoover–Kallenberg)**.** *Every invariant probability measure μ on Str_L has an AHK representation.*

Once a proper translation of notation is applied, Theorem 2.27 is equivalent to [Kal05, Theorem 7.22], which is usually called the Aldous–Hoover–Kallenberg theorem. For more on such a translation see [Ack15, §2.5].

The AHK representation produced by Theorem 2.27 is not unique, but it is unique up to certain appropriately measure-preserving transformations. See [Kal05, Theorem 7.28] for a precise statement.

The key fact to observe about AHK systems is that if $\overline{a}$ and $\overline{b}$ are tuples from $\mathbb{N}$ whose intersection $\|\overline{a}\| \cap \|\overline{b}\|$ is enumerated by the tuple $\overline{c}$, then the random quantifier-free types $r_{\overline{a}}$ and $r_{\overline{b}}$ are conditionally independent over $\widehat{\xi_{\overline{c}}}$. If $\overline{a}$ and $\overline{b}$ are disjoint, then $\widehat{\xi_{\overline{c}}} = \xi_\emptyset$.

The Aldous–Hoover–Kallenberg theorem also provides a characterization of the ergodic measures among the invariant measures on Str_L: they are those measures for which the random quantifier-free types $r_{\overline{a}}$ and $r_{\overline{b}}$ are independent when $\overline{a}$ and $\overline{b}$ are disjoint. Formally, for an n-tuple $\overline{a}$ from $\mathbb{N}$, let $\Sigma_{\overline{a}}$ be the σ-algebra on Str_L generated by the sets $\llbracket\varphi(\overline{a})\rrbracket$, where $\varphi(\overline{x})$ ranges over the quantifier-free formulas in the n-tuple of variables $\overline{x}$. We say that an invariant probability measure μ on Str_L is **dissociated** if whenever $\overline{a}$ and $\overline{b}$ are disjoint tuples from $\mathbb{N}$, the σ-algebras $\Sigma_{\overline{a}}$ and $\Sigma_{\overline{b}}$ are independent (see Remark 2.7 above).

Theorem 2.28 ([Kal05, Lemma 7.35])**.** *Let μ be an invariant probability measure on Str_L. The following are equivalent:*

1. *μ is ergodic.*
2. *μ is dissociated.*
3. *μ has an AHK representation in which the functions f_n do not depend on the argument indexed by $\emptyset$.*

The result [Kal05, Lemma 7.35] is stated for finite relational languages, but can be generalized to our setting by a careful modification of the proofs, as described in [Ack15, Corollary 2.18].

3. Examples

In this section, we describe some examples of properly ergodic structures and their theories, as well as theories all of whose ergodic models are not properly ergodic. In doing so, we highlight some of the key notions of the paper, including trivial definable closure and rootedness, and the relevance of infinitary logic. Certain examples are naturally described using infinite languages, but for some we also describe how they may be framed in terms of finite languages.

When we say that we pick a random element $A \in 2^{\mathbb{N}}$, we always refer to the uniform (Lebesgue) measure on $2^{\mathbb{N}}$, the infinite product of the Bernoulli($1/2$) measure on $2 = \{0, 1\}$. We identify such an $A \in 2^{\mathbb{N}}$ with both a subset of $\mathbb{N}$ and an infinite binary sequence.

Our first example of a class of properly ergodic structures arose naturally in the study of random graphs.

Example 3.1 (Random geometric graphs). Consider a metric space (X, d), a probability measure m on X, and a real number $p \in (0, 1)$. Bonato and Janssen [BJ11] define the random geometric graph given by first sampling an m-i.i.d. sequence of vertices $D \subseteq X$ and then connecting two points $x, y \in D$ such that $d(x, y) < 1$ by an edge or not based on an independent weight-p coin flip. The distribution of this random construction is an ergodic structure.

Bonato and Janssen showed that when the metric space is ℓ^n_∞ for some n, the random geometric graph is almost surely isomorphic to a single countable graph, but that on the other hand, the Euclidean plane yields a properly ergodic structure. In fact, as shown in [BBG⁺15], every normed linear space other than ℓ^n_∞ yields a properly ergodic structure.

The next class of examples, which was introduced in [AFNP16, §5.1], can be thought of as countably many overlaid instances of the Erdős–Rényi random (hyper-)graphs. These are some of the key examples of properly ergodic structures, based on which we also will build several variants.

Example 3.2 (Kaleidoscope structures). We begin by describing the case of binary relations. Let $L = \{R_n \mid n \in \mathbb{N}\}$, where each R_n is a binary relation symbol. The interpretation of each R_n will be irreflexive and symmetric; one may think of each R_n as a different "color" of edge.

Consider the random L-structure with domain $\mathbb{N}$ obtained by first picking a random $A_{\{i,j\}} \in 2^{\mathbb{N}}$ independently for each pair of distinct elements $i, j \in \mathbb{N}$, and then setting iR_nj just when $n \in A_{\{i,j\}}$. Let μ be the distribution of this random structure.

Observe that the measure μ is invariant, since the random quantifier-free type of a tuple of distinct elements does not depend on the choice of tuple. Further, μ is ergodic by Theorem 2.28, since the random quantifier-free types of disjoint tuples are independent. We call this ergodic structure μ the **kaleidoscope**

random graph. (Note that in [AFNP16], this term is used instead to refer to models of its almost-sure first-order theory.)

Note that there are continuum-many quantifier-free 2-types consistent with $\mathrm{Th}_{\mathrm{FO}}(\mu)$, each of which is realized with probability 0 in μ. Any particular countable L-structure realizes at most countably many such types, and so μ assigns measure 0 to its isomorphism class. Hence μ is properly ergodic.

In fact, for every $A \in 2^{\mathbb{N}}$, the theory $\mathrm{Th}(\mu)$ contains the sentence

$$\neg\exists x\, \exists y \Big(\bigwedge_{n\in A} xR_n y \wedge \bigwedge_{n\notin A} \neg x R_n y \Big).$$

Since all quantifier-free 2-types consistent with $\mathrm{Th}_{\mathrm{FO}}(\mu)$ are ruled out by $\mathrm{Th}(\mu)$, the theory $\mathrm{Th}(\mu)$ has no models of any cardinality. In fact, the complete $\mathcal{L}_{\omega_1,\omega}$-theory of any properly ergodic structure has no models of any cardinality, as shown in Corollary 4.9. Note, however, that any countable fragment F of $\mathcal{L}_{\omega_1,\omega}$ only contains countably many of the sentences above, so $\mathrm{Th}_F(\mu)$ only rules out countably many of the quantifier-free 2-types.

Restricting to the first-order fragment, the theory $\mathrm{Th}_{\mathrm{FO}}(\mu)$ has several nice properties. It is the model companion of the universal theory asserting that each R_n is irreflexive and symmetric. It can be axiomatized by extension axioms, analogous to those in the theory of the Rado graph: in each finite sublanguage $L^* \subseteq L$, for every finite tuple A and non-redundant quantifier-free 1-type over A in the language L^* consistent with $\mathrm{Th}_{\mathrm{FO}}(\mu)$, there is some element b satisfying that quantifier-free type. The reduct of $\mathrm{Th}_{\mathrm{FO}}(\mu)$ to any finite sublanguage is countably categorical, but $\mathrm{Th}_{\mathrm{FO}}(\mu)$ has continuum-many countable models (since there are continuum-many quantifier-free 2-types consistent with $\mathrm{Th}_{\mathrm{FO}}(\mu)$). In fact, for all countable fragments F and properly ergodic structures μ, the theory $\mathrm{Th}_F(\mu)$ has continuum-many countable models, as we also show in Corollary 4.9.

For arbitrary arity $k \geq 1$, an analogous construction produces the **kaleidoscope random k-uniform hypergraph**. We call the case $k = 1$ the **kaleidoscope random predicate**.

We now use the latter example to illustrate the distinction between group-theoretic and syntactic definable closure.

Example 3.3 (The theory of the kaleidoscope random predicate). Let T be the first-order theory of the kaleidoscope random predicate (see Example 3.2) in the language $\{P_n \mid n \in \mathbb{N}\}$. The theory T says that for every $m \in \mathbb{N}$ and every subset $A \subseteq [m]$, there is an element x such that for all $n \in [m]$, the relation $P_n(x)$ holds if and only if $n \in A$.

Now let T' be T together with the infinitary sentence

$$\forall x\, \forall y \,(\bigwedge_{n\in\mathbb{N}} (P_n(x) \leftrightarrow P_n(y)) \rightarrow x = y).$$

The kaleidoscope random predicate almost surely satisfies T'. Each of the continuum-many quantifier-free 1-types is realized with probability 0, and since the quantifier-free 1-types of distinct elements of $\mathbb{N}$ are independent, almost surely no 1-type is realized more than once.

In a model M of T', no two elements have the same quantifier-free 1-type. Hence $\mathrm{Aut}(M)$ is the trivial group, and M has non-trivial group-theoretic dcl. Let F be the countable fragment of $\mathcal{L}_{\omega_1,\omega}$ generated by T'. Then F does not contain the conjunctions of the form $\bigwedge_{n\in A} P_n(x) \wedge \bigwedge_{n\notin A} \neg P_n(x)$ for $A \subseteq \mathbb{N}$ needed to pin down elements uniquely. In fact, the complete F-theory of the kaleidoscope random predicate (which extends T') has trivial dcl, by Theorem 2.19.

We will see that the presence of a formula $\chi(\overline{x})$ of positive measure, such that every type containing χ has probability 0 of being realized, is a characteristic feature of properly ergodic structures.

In the kaleidoscope random graph (Example 3.2), $x \neq y$ is such a formula $\chi(x, y)$, since every non-redundant quantifier-free 2-type is realized with probability 0. In contrast to the kaleidoscope random graph, Example 3.4 shows that there are properly ergodic structures in which these 0-probability types have infinitely many realizations if they are realized at all.

On the other hand, in Example 3.5, we describe a transformation (known as the "blow-up"), which when applied to the Kaleidoscope random predicate, leads to each of the continuum-many 1-types being realized infinitely many times (if at all), and yet whose resulting theory has no properly ergodic models. This shows that merely having continuum-many types in a theory with trivial dcl does not imply the existence of a properly ergodic model of the theory.

These phenomena motivate the definition of rootedness in Section 5.

Example 3.4 (The max random graph). As in Example 3.2, let $L = \{R_n \mid n \in \mathbb{N}\}$, where each R_n is a binary relation symbol. We build a random L-structure with domain $\mathbb{N}$ such that the interpretation of each R_n is irreflexive and symmetric. For each $i \in \mathbb{N}$, independently choose a random element $A_i \in 2^{\mathbb{N}}$. Now for each pair $\{i, j\}$, let $A_{ij} = \max(A_i, A_j)$, where we give $2^{\mathbb{N}}$ its lexicographic order. We set iR_nj if and only if $n \in A_{ij}$.

We have continuum-many quantifier-free 2-types $\{p_A \mid A \in 2^{\mathbb{N}}\}$, where $xR_ny \in p_A$ if and only if $n \in A$, and each is realized with probability 0, since if (i, j) realizes p_A, we must have $A_i = A$ or $A_j = A$.

As long as A_i is not the constant 0 sequence (which appears with probability 0), then for any $j \neq i$, there is a positive probability, conditioned on the choice of A_i, that $A_j \leq A_i$, and hence $\mathrm{qftp}(i, j) = p_{A_i}$. Since the A_j are chosen independently, almost surely the event $A_j \leq A_i$ occurs for infinitely many j. So, almost surely, any non-redundant quantifier-free 2-type that is realized is realized infinitely many times. However, since the probability that $A_i = A_j$ when $i \neq j$ is 0, almost surely all realizations of p_{A_i} have a common intersection, namely the vertex i.

We now describe a modification of the theory of the kaleidoscope random predicate so that all of its ergodic models are not properly ergodic.

Example 3.5 (The blow-up of the theory of the kaleidoscope random predicate). Let $L = \{E\} \cup \{P_n \mid n \in \mathbb{N}\}$, and let T be the model companion of the universal theory asserting that E is an equivalence relation and the P_n are unary predicates respecting E (if xEy, then $P_n(x)$ if and only if $P_n(y)$). This is similar to the first-order theory of the kaleidoscope random predicate, but with each element replaced by an infinite E-class.

There is no properly ergodic structure that satisfies T almost surely. Indeed, suppose $\mu \models T$. Then for every quantifier-free 1-type p, there is some probability $\mu(p)$ that p is the quantifier-free type of the element $i \in \mathbb{N}$, and, by invariance, $\mu(p)$ does not depend on the choice of i. We denote by $S^1_{\mathrm{qf}}(\mu)$ the (countable) set of quantifier-free 1-types with positive measure. If $\sum_{p \in S^1_{\mathrm{qf}}(\mu)} \mu(p) = 1$, then almost surely only the types in $S^1_{\mathrm{qf}}(\mu)$ are realized, since $\mu \models \forall x \bigvee_{p \in S^1_{\mathrm{qf}}(\mu)} \bigwedge_{\varphi \in p} \varphi(x)$. Further, μ determines, for each $p \in S^1_{\mathrm{qf}}(\mu)$, the number of E-classes on which p is realized (among $\{1, 2, \ldots, \aleph_0\}$), since each of the countably many choices is expressible by a sentence of $\mathcal{L}_{\omega_1,\omega}$. The data of which quantifier-free 1-types are realized, and how many E-classes realize each, determines a unique countable L-structure up to isomorphism, so μ is not properly ergodic.

On the other hand, if $\sum_{p \in S^1_{\mathrm{qf}}(\mu)} \mu(p) < 1$, then almost surely some types that are not in $S^1_{\mathrm{qf}}(\mu)$ are realized. Any such type p is realized with probability 0, and, by ergodicity, the quantifier-free 1-types of distinct elements of $\mathbb{N}$ are independent. So, almost surely, each of the 0-probability types is realized at most once. This contradicts the fact that any realized type must be realized on an entire infinite E-class.

The next example shows why it important to use $\mathcal{L}_{\omega_1,\omega}$ when performing the Morley–Scott analysis.

Example 3.6 (A kaleidoscope-like bipartite graph). Let $L = \{P\} \cup \{R^i_j \mid i, j \in \mathbb{N}\}$, where P is a unary predicate and the R^i_j are binary relations, and let T be the model companion of the following universal theory:

(1) $\forall x \, \forall y \, (R^i_j(x, y) \to (P(x) \wedge \neg P(y)))$ for all i and j.
(2) $\forall x \forall y \, \neg(R^i_0(x, y) \wedge R^{i'}_0(x, y))$ for all $i \neq i'$.
(3) $\forall x \, \forall y \, (R^i_{j+1}(x, y) \to R^i_j(x, y))$ for all i and j.

Thus, a model of T is a bipartite graph in which each edge from x to y is labeled by some $i \in \mathbb{N}$ (in the superscript) and the set of all $j < k$ for some $k \in \mathbb{N}_+ \cup \{\infty\}$ (in the subscript), where $\mathbb{N}_+$ denotes the positive natural numbers.

Now T is a complete theory with quantifier elimination and with only countably many quantifier-free types over $\emptyset$. Hence, by countable additivity, if μ is an ergodic structure that satisfies T almost surely, then there is no positive-measure first-order formula $\chi(\overline{x})$ such that every type containing χ has measure

0. Nevertheless, we will describe a properly ergodic structure that almost surely satisfies T.

First, for each $x \in \mathbb{N}$, let $P(x)$ hold with independent probability $1/2$, and pick $A_x \in 2^{\mathbb{N}}$ independently at random. Now for each pair $x \neq y$, if $P(x)$ and $\neg P(y)$, then we choose which of the R^i_j will hold of (x, y). First independently choose $i \in \mathbb{N}$ according to a geometric distribution where $i = n$ holds with probability $2^{-(n+1)}$. Then, if $i \in A_x$, independently choose $k \in \mathbb{N}_+ \cup \{\infty\}$ according to a geometric distribution where $k = \infty$ holds with probability $1/2$ and $k = n$ holds with probability $2^{-(n+1)}$ for $n \in \mathbb{N}_+$. On the other hand, if $i \notin A_x$, then independently choose $k \in \mathbb{N}_+$ according to a geometric distribution where $k = n$ holds with probability 2^{-n}. Finally, for this choice of i and k, we let $R^i_j(x, y)$ hold for all $j < k$.

In the resulting random structure, we can almost surely recover A_x from every $x \in P$, since if $i \in A_x$, then almost surely there is some y such that $R^i_j(x, y)$ for all $j \in \mathbb{N}$ (that is, the choice $k = \infty$ was made for the pair (x, y)), whereas this outcome is impossible if $i \notin A_x$. Thus the structure encodes a countable set of elements of $2^{\mathbb{N}}$, each of which occurs with probability 0.

The information encoding A_x is part of the 1-type of x in any countable fragment of $\mathcal{L}_{\omega_1,\omega}$ containing the infinitary formulas $\{\exists y \bigwedge_{j\in\mathbb{N}} R^i_j(x, y) \mid i \in \mathbb{N}\}$, but it is not expressible in first-order logic.

With the exception of Example 3.1 (and Gábor Kun's example alluded to in §1.2), the preceding examples have all used infinite languages, as this is the easiest setting in which to split the measure over continuum-many types. We conclude with an elementary example in the language with a single binary relation, which encodes the kaleidoscope random predicate into a directed graph, in a way that we easily verify is properly ergodic.

Example 3.7 (A directed graph encoding the kaleidoscope random predicate). Let $L = \{R\}$, where R is a binary relation. In our probabilistic construction, we will enforce the following almost surely:

- Let $O = \{x \mid R(x, x)\}$, and $P = \{x \mid \neg R(x, x)\}$. Then O and P are both infinite sets.
- If $R(x, y)$, then either x and y are both in O, or x is in P and y is in O.
- R is a preorder on O. Denote by xEy the induced equivalence relation $R(x, y) \wedge R(y, x)$. Then E has infinitely many infinite classes, and R linearly orders the E-classes with order type ω.
- Given $x \in P$ and $y, z \in O$, if $R(x, y)$ and yEz, then $R(x, z)$. So R relates each element of P to some subset of the E-classes.

Thus we can interpret the kaleidoscope random predicate on P, where the n^{th} predicate P_n holds of x if and only if x is R-related to the n^{th} class in the linear order on O.

Now it is straightforward to describe the probabilistic construction: for each $i \in \mathbb{N}$, independently let $R(i,i)$ hold with probability $1/2$. This determines whether i is in O or P. If $i \in O$, we choose which E-class to put i in, under the order induced by R, selecting the n^{th} class independently with probability $2^{-(n+1)}$. These choices determine all the R-relations between elements of O. On the other hand, if $i \in P$, we pick $A_i \in 2^{\mathbb{N}}$ independently at random and relate i to each the n^{th} class in O if and only if $n \in A_i$.

This describes an ergodic structure μ, since the quantifier-free types of disjoint tuples are independent. We obtain the properties described in the bullet points above almost surely, and since ω is rigid, any isomorphism between structures satisfying these properties must preserve the order on the E-classes. For any subset of the E-classes, the probability is 0 that there is an element of P which is related to exactly those E-classes, and so μ is properly ergodic.

4. Morley–Scott analysis of ergodic structures

Throughout this section, let μ be an ergodic structure. Recall from Remark 2.16 that for a countable fragment F of $\mathcal{L}_{\omega_1,\omega}$ and an F-type p, the abbreviation $\theta_p(\overline{x})$ means $\bigwedge_{\varphi\in p} \varphi(\overline{x})$, and the notation $\mu(p)$ means $\mu(\theta_p(\overline{x}))$ and is called the *measure of* p.

Definition 4.1. We denote by $S^n_F(\mu)$ the set $\{p \mid \mu(p) > 0\}$ of **positive-measure F-types** in the variables $x_0, \ldots, x_{n-1}$. We include the case $n = 0$: $S^0_F(\mu)$ has one element, namely $\mathrm{Th}_F(\mu)$.

Lemma 4.2. *For all $n \in \mathbb{N}$, we have $|S^n_F(\mu)| \leq \aleph_0$.*

Proof. Fix a tuple $\overline{a}$ of distinct elements from ω. The sets $\{[\![\theta_p(\overline{a})]\!] \mid p \in S^n_F(\mu)\}$ are disjoint sets of positive measure in Str_L. By additivity of μ, for all $m \in \mathbb{N}$, $P_m = \{p \in S^n_F(\mu) \mid \mu(p) \geq 1/m\}$ is finite (of size at most m), so $S^n_F(\mu) = \bigcup_{m\in\omega} P_m$ is countable. □

We build a sequence $\{F_\alpha\}_{\alpha\in\omega_1}$ of countable fragments of $\mathcal{L}_{\omega_1,\omega}$ of length ω_1, depending on the ergodic structure μ:

$$F_0 = \mathrm{FO}, \text{the first-order fragment.}$$

$$F_{\alpha+1} = \text{the fragment generated by } F_\alpha \cup \Big\{\theta_p(\overline{x}) \mid p \in \bigcup_{n\in\mathbb{N}} S^n_{F_\alpha}(\mu)\Big\}.$$

$$F_\gamma = \bigcup_{\alpha<\gamma} F_\alpha, \text{if } \gamma \text{ is a limit ordinal.}$$

Definition 4.3. We say that $p \in S^n_{F_\alpha}(\mu)$ **splits** at $\beta > \alpha$ if $\mu(q) < \mu(p)$ for all types $q \in S^n_{F_\beta}(\mu)$ such that $p \subseteq q$. We say that p **splits later** if there exists β such that p splits at β. We say that μ **has stabilized** at γ if for all $n \in \mathbb{N}$, no type in $S^n_{F_\gamma}(\mu)$ splits later.

Lemma 4.4. *Let $\alpha < \beta < \gamma$.*

(1) If a type $p \in S^n_{F_\alpha}(\mu)$ splits at β, then p also splits at γ.
(2) Suppose $p \in S^n_{F_\beta}(\mu)$ splits at γ. Then $p' = p \cap F_\alpha$ is in $S^n_{F_\alpha}(\mu)$ and also splits at γ.
(3) If no type in $S^n_{F_\alpha}(\mu)$ splits later, then no type in $S^n_{F_\beta}(\mu)$ splits later.

Proof. (1) Pick $q \in S^n_{F_\gamma}(\mu)$ with $p \subseteq q$, and let $q' = q \cap F_\beta$. Then $\mu(q) \leq \mu(q') < \mu(p)$, since p splits at β.

(2) First, $0 < \mu(p) \leq \mu(p')$, so $p' \in S^n_{F_\alpha}(\mu)$. Pick $q \in S^n_{F_\gamma}(\mu)$ such that $p' \subseteq q$. If $p \subseteq q$, then $\mu(q) < \mu(p) \leq \mu(p')$, since p splits at γ. And if $p \not\subseteq q$, then $\mu(q) \leq \mu(p') - \mu(p) < \mu(p')$, since $\mu(p) > 0$. In either case, $\mu(q) < \mu(p')$, so p' splits at γ.

(3) If some type in $S^n_{F_\beta}(\mu)$ splits later, then by (2), $p' = p \cap F_\alpha$ also splits later, and $p' \in S^n_{F_\alpha}(\mu)$. □

Lemma 4.5. *There is some countable ordinal γ such that μ has stabilized at γ.*

Proof. Fix $n \in \mathbb{N}$. For each $\alpha \in \omega_1$, let

$$\mathrm{Sp}(\alpha) = \{p \in S^n_{F_\alpha}(\mu) \mid p \text{ splits later}\},$$
$$r_\alpha = \sup\{\mu(p) \mid p \in \mathrm{Sp}(\alpha)\}.$$

Note that $\mathrm{Sp}(\alpha)$ is countable, since $S^n_{F_\alpha}$ is. If $\mathrm{Sp}(\alpha)$ is non-empty, then $r_\alpha > 0$, and in fact the supremum is achieved by finitely many types, since $\sum_{p \in \mathrm{Sp}(\alpha)} \mu(p) \leq 1$.

By Lemma 4.4 (2), the measure of any type in $\mathrm{Sp}(\beta)$ is bounded above by the measure of a type in $\mathrm{Sp}(\alpha)$, namely its restriction to F_α. So we have $r_\beta \leq r_\alpha$ whenever $\alpha < \beta$.

Now assume for a contradiction that $\mathrm{Sp}(\alpha)$ is non-empty for all α. We build a strictly increasing sequence $\langle \alpha_\delta \rangle_{\delta \in \omega_1}$ in ω_1, such that $\langle r_{\alpha_\delta} \rangle_{\delta \in \omega_1}$ is a strictly decreasing sequence in $[0,1]$. Begin with $\alpha_0 = 0$.

At each successor stage, we are given $\alpha = \alpha_\delta$, and we seek $\beta = \alpha_{\delta+1}$ with $r_\beta < r_\alpha$. Since $\mathrm{Sp}(\alpha)$ is non-empty, there are finitely many types $p_1, \dots, p_n$ of maximal measure $r_\alpha > 0$. For each i, pick $\beta_i > \alpha$ such that p_i splits at β_i, and let $\beta = \max(\beta_1, \dots, \beta_n)$. By Lemma 4.4 (1), each p_i splits at β. Let q be a type in $\mathrm{Sp}(\beta)$ with $\mu(q) = r_\beta$, and let $q' = q \cap F_\alpha$. By Lemma 4.4 (2), $q' \in \mathrm{Sp}(\alpha)$. If q' is one of the p_i, then $\mu(q) < \mu(p_i) = r_\alpha$, since p_i splits at β. If not, then $\mu(q) \leq \mu(q') < r_\alpha$. In either case, $r_\beta = \mu(q) < r_\alpha$.

If λ is a countable limit ordinal, let $\alpha_\lambda = \sup_{\delta<\lambda} \alpha_\delta$. This is an element of ω_1, since ω_1 is regular. And for all $\delta < \lambda$, since $\alpha_{\delta+1} < \alpha_\lambda$, we have $r_{\alpha_\lambda} \leq r_{\alpha_{\delta+1}} < r_{\alpha_\delta}$.

Of course, there is no strictly decreasing sequence of real numbers of length ω_1, since $\mathbb{R}$ contains a countable dense set. Hence there is some $\gamma_n \in \omega_1$ such that $\mathrm{Sp}(\gamma_n)$ is empty, i.e., no type in $S^n_{F_{\gamma_n}}$ splits later. Let $\gamma = \sup_{n\in\mathbb{N}} \gamma_n \in \omega_1$. Then by Lemma 4.4 (3), μ has stabilized at γ. □

We can think of the minimal ordinal γ such that μ has stabilized at γ as an analogue of the Scott rank for the ergodic structure μ. Since no F_γ-type splits later, every positive-measure $F_{\gamma+1}$-type q is isolated by the $F_{\gamma+1}$-formula θ_p for its restriction $p = q \cap F_\gamma$, relative to $\mathrm{Th}_{F_{\gamma+1}}(\mu)$. Lemma 4.6 says that if every tuple satisfies one of these positive-measure types almost surely, then μ almost surely satisfies a Scott sentence.

Lemma 4.6. *Suppose that μ has stabilized at γ, and that for all $n \in \mathbb{N}$,*

$$\sum_{p\in S^n_{F_\gamma}(\mu)} \mu(p) = 1.$$

Then μ concentrates on a countable structure.

Proof. For each type $r(\overline{x}) \in S^n_{F_\gamma}(\mu)$ (we include the case $n = 0$), let E_r be the set of types $q(\overline{x}, y) \in S^{n+1}_{F_\gamma}(\mu)$ with $r \subseteq q$. Fix a type $p(\overline{x}) \in S^n_{F_\gamma}(\mu)$, let φ_p be the sentence

$$\forall \overline{x} \left(\theta_p(\overline{x}) \to \forall (y \notin \overline{x}) \bigvee_{q\in E_p} \theta_q(\overline{x}, y)\right),$$

and let ψ_p be the sentence

$$\forall \overline{x} \left(\theta_p(\overline{x}) \to \bigwedge_{q\in E_p} \exists (y \notin \overline{x})\, \theta_q(\overline{x}, y)\right)$$

Here $\forall(y \notin \overline{x})\rho(\overline{x}, y)$ and $\exists(y \notin \overline{x})\rho(\overline{x}, y)$ are shorthand for $\forall y((\bigwedge_{i=0}^{n-1} y \neq x_i) \to \rho(\overline{x}, y))$ and $\exists y((\bigwedge_{i=0}^{n-1} y \neq x_i) \wedge \rho(\overline{x}, y))$, respectively. We would like to show that μ satisfies φ_p and ψ_p almost surely.

By assumption, and since every $q \in S^{n+1}_{F_\gamma}(\mu)$ is in E_r for a unique $r \in S^n_{F_\gamma}(\mu)$,

$$1 = \sum_{q\in S^{n+1}_{F_\gamma}(\mu)} \mu(q) = \sum_{r\in S^n_{F_\gamma}(\mu)} \sum_{q\in E_r} \mu(q).$$

Then for all $r \in S^n_{F_\gamma}(\mu)$, we must have

$$\mu(r) = \sum_{q\in E_r} \mu(q).$$

In particular, this is true for $r = p$, so for any tuple $\overline{a}$ and any b not in $\overline{a}$, $[\![\bigvee_{q\in E_p} \theta_q(\overline{a}, b)]\!]$ has full measure in $[\![\theta_p(\overline{a})]\!]$ (this is true even when $\overline{a}$ contains repeated elements, since in that case $[\![\theta_p(\overline{a})]\!]$ has measure 0). A countable intersection (over $b \in \mathbb{N} \setminus \|\overline{a}\|$) of subsets of $[\![\theta_p(\overline{a})]\!]$ with full measure still has full measure, so

$$\mu\Big(\Big[\!\!\Big[\theta_p(\overline{a}) \to \forall (y \notin \overline{a}) \bigvee_{q \in E_p} \theta_q(\overline{a}, y)\Big]\!\!\Big]\Big) = 1.$$

Taking another countable intersection over all tuples $\overline{a}$, we have $\mu \models \varphi_p$.

We turn now to ψ_p. Since μ stabilizes at γ, there is a (necessarily unique) extension of p to a type $p^* \in S^n_{F_{\gamma+1}}(\mu)$ with $\mu(p^*) = \mu(p)$. Let $q(\overline{x}, y)$ be any type in E_p, and let $\upsilon_q(\overline{x}) \in F_{\gamma+1}$ be the formula $(\exists y \notin \overline{x})\, \theta_q(\overline{x}, y)$. Note that $\theta_q(\overline{x}, y)$ implies $\upsilon_q(\overline{x})$ and $\upsilon_q(\overline{x})$ implies $\theta_p(\overline{x})$. So $\mu(\upsilon_q(\overline{x})) \geq \mu(q) > 0$, and we must have $\upsilon_q(\overline{x}) \in p^*$, otherwise $\mu(p^*) \leq \mu(p) - \mu(\upsilon_q(\overline{x}))$. Finally, we conclude that for any tuple $\overline{a}$, the set $[\![\upsilon_q(\overline{a})]\!]$ has full measure in $[\![\theta_p(\overline{a})]\!]$, since $\mu(p) = \mu(p^*) \leq \mu(\upsilon_q(\overline{x})) \leq \mu(p)$.

As before, a countable intersection of subsets with full measure has full measure, so

$$\mu\Big(\Big[\!\!\Big[\theta_p(\overline{a}) \to \bigwedge_{q \in E_p} \exists (y \notin \overline{a})\, \theta_q(\overline{a}, y)\Big]\!\!\Big]\Big) = 1.$$

Taking another countable intersection over all tuples $\overline{a}$, we have $\mu \models \psi_p$.

Let $T = \mathrm{Th}_{F_\gamma}(\mu) \cup \{\varphi_p, \psi_p \mid p \in \bigcup_{n \in \mathbb{N}} S^n_{F_\gamma}(\mu)\}$, and note that T is countable. Since μ almost surely satisfies T, it suffices to show that any two countable models of T are isomorphic. This is a straightforward back-and-forth argument, using φ_p and ψ_p to extend a partial F_γ-elementary isomorphism defined on a realization of p by one step: φ_p tells us that each one-point extension in one model realizes one of the types in E_p, and ψ_p tells us that every type in E_p is realized in a one-point extension in the other model. To start, the empty tuples in any two models of T satisfy the same F_γ-type, namely $\mathrm{Th}_{F_\gamma}(\mu)$. □

Theorem 4.7. *Let μ be an ergodic structure. Then μ is properly ergodic if and only if for every countable fragment F of $\mathcal{L}_{\omega_1,\omega}$, there is a countable fragment $F' \supseteq F$ and a formula $\chi(\overline{x})$ in F' such that $\mu(\chi(\overline{x})) > 0$, but $\mu(p) = 0$ for every F'-type $p(\overline{x})$ containing $\chi(\overline{x})$.*

Proof. Suppose μ is properly ergodic. By Lemma 4.5, μ stabilizes at some γ, and by Lemma 4.6, there is some n such that $\sum_{p \in S^n_{F_\gamma}(\mu)} \mu(p) < 1$. Let $\chi(\overline{x})$ be the formula $\bigwedge_{p \in S^n_{F_\gamma}(\mu)} \neg\theta_p(\overline{x})$. Then $\mu(\chi(\overline{x})) > 0$.

Let F' be the countable fragment generated by $F \cup F_\gamma \cup \{\chi(\overline{x})\}$, and suppose that $p(\overline{x})$ is an F'-type containing $\chi(\overline{x})$. Let $q = p \cap F_\gamma$. Then q is an F_γ type that is consistent with $\chi(\overline{x})$, so $q \notin S^n_{F_\gamma}(\mu)$, and $\mu(p) \leq \mu(q) = 0$.

Conversely, suppose we have such a fragment F' and such a formula $\chi(\overline{x})$. Since $\mu(\chi(\overline{x})) > 0$, by ergodicity, $\mu \models \exists \overline{x}\, \chi(\overline{x})$. Let M be a countable structure. If M contains no tuple satisfying χ, then μ assigns measure 0 to the isomorphism class of M. On the other hand, if M contains a tuple $\overline{a}$ satisfying $\chi(\overline{x})$, then since μ assigns measure 0 to the set of structures realizing $\mathrm{tp}_{F'}(\overline{a})$, it also assigns measure 0 to the isomorphism class of M. So μ is properly ergodic. $\square$

By countable additivity, if a sentence φ of $\mathcal{L}_{\omega_1,\omega}$ has only countably many countable models up to isomorphism, then any ergodic structure μ that almost surely satisfies φ is almost surely isomorphic to one of its models. That is, no ergodic model of φ is properly ergodic. We show now that the same is true if φ is a counterexample to Vaught's conjecture, i.e., a sentence with uncountably many, but fewer than continuum-many, countable models.

Corollary 4.8 ("Vaught's Conjecture for ergodic structures")**.** *Let φ be a sentence of $\mathcal{L}_{\omega_1,\omega}$. If there is a properly ergodic structure μ such that $\mu \models \varphi$, then φ has continuum-many countable models up to isomorphism.*

Proof. This is a consequence of Theorem 4.7 and an observation due to Morley [Mor70]: for any countable fragment F of $\mathcal{L}_{\omega_1,\omega}$ containing φ and any $n \in \mathbb{N}$, the set $S^n_F(\varphi)$ of F-types consistent with φ is an analytic subset of 2^F. Since analytic sets have the Perfect Set Property, if $|S^n_F(\varphi)| > \aleph_0$, then $|S^n_F(\varphi)| = 2^{\aleph_0}$. And since a countable structure realizes only countably many n-types, if $|S^n_F(\varphi)| = 2^{\aleph_0}$, then φ must have continuum-many countable models up to isomorphism.

Now let μ be the given properly ergodic structure, let F be a countable fragment containing φ, let F' and $\chi(\overline{x})$ be as in Theorem 4.7, let n be the length of the tuple $\overline{x}$, and suppose for a contradiction that $|S^n_{F'}(\varphi)| \leq \aleph_0$. Let $U_\chi = \{p \in S^n_{F'}(\varphi) \mid \chi(\overline{x}) \in p\}$. Then U_χ is countable, and, by our choice of $\chi(\overline{x})$, we have $\mu(p) = 0$ for all $p \in U_\chi$. Since $\mu(\llbracket\varphi\rrbracket) = 1$, for any tuple $\overline{a}$ of distinct elements of $\mathbb{N}$, we have

$$0 < \mu(\llbracket\chi(\overline{a})\rrbracket) = \mu(\llbracket(\varphi \wedge \chi)(\overline{a})\rrbracket) = \mu\Big(\bigcup_{p \in U_\chi} \llbracket\theta_p(\overline{a})\rrbracket\Big) = \sum_{p \in U_\chi} \mu(p),$$

which is a contradiction, by countable additivity of μ. $\square$

Kechris has observed (in private communication) that Corollary 4.8 also follows from a result in descriptive set theory [Kec95, Exercise 17.14]: an analogue for measure of a result of Kuratowski about category [Kur76]. However, our proof above provides additional model-theoretic information about properly ergodic structures.

Recall that $\mathrm{Th}(\mu)$ is the complete $\mathcal{L}_{\omega_1,\omega}$-theory of μ. As noted in Remark 2.12, μ is properly ergodic if and only if $\mathrm{Th}(\mu)$ has no countable models. In fact, if μ is properly ergodic, then $\mathrm{Th}(\mu)$ has no models at all. This is stronger, since the Löwenheim–Skolem theorem fails for complete theories of $\mathcal{L}_{\omega_1,\omega}$.

Corollary 4.9. *If μ is properly ergodic, then $\mathrm{Th}(\mu)$ has no models (of any cardinality). However, for any countable fragment F of $\mathcal{L}_{\omega_1,\omega}$, the theory $\mathrm{Th}_F(\mu)$ has continuum-many countable models up to isomorphism.*

Proof. Starting with any countable fragment F (e.g., $F = \mathrm{FO}$), let F' and $\chi(\overline{x})$ be as in Theorem 4.7. Then $\mu(\chi(\overline{x})) > 0$, so by ergodicity, $\exists \overline{x}\, \chi(\overline{x}) \in \mathrm{Th}(\mu)$. Now if $\mathrm{Th}(\mu)$ has a model M, then there is some tuple $\overline{a}$ from M satisfying $\chi(\overline{x})$. Let p be the F'-type of $\overline{a}$. Since p contains $\chi(\overline{x})$, we have $\mu(p) = 0$, and so $\neg\exists \overline{x}\, \theta_p(\overline{x}) \in \mathrm{Th}(\mu)$, a contradiction.

The last assertion follows from Corollary 4.8, taking $\varphi = \bigwedge_{\psi \in \mathrm{Th}_F(\mu)} \psi$. □

Corollary 4.9 describes a general version of two phenomena observed for the Kaleidoscope random graph in Example 3.2.

5. Rooted models

The Morley–Scott analysis in Section 4 showed that proper ergodicity of μ can always be explained by a positive-measure formula $\chi(\overline{x})$ such that any type containing $\chi(\overline{x})$ has measure 0. In a countable structure sampled from μ, each of these types of measure 0 will be realized "rarely". Sometimes "rarely" means "at most once", as in Examples 3.2 and 3.3. But in Example 3.4, the max graph, we saw that a type p of measure 0 can be realized by infinitely many tuples, all of which share a common element $i \in \mathbb{N}$. In that example, if some element $A \in 2^{\mathbb{N}}$ is randomly selected at a vertex i, then for any other vertex j, there is a positive probability that $\mathrm{qftp}(i, j)$ is the type p_A encoding A. In other words, the fact that the type p_A is realized infinitely many times is explained by the fact that p_A has positive measure, after the random choice of A "living at" the vertex i. In this section, we will use the Aldous–Hoover–Kallenberg theorem from §2.4 to show that this behavior is typical.

Throughout this section, let F be a countable fragment of $\mathcal{L}_{\omega_1,\omega}$, and let T be an F-theory. We write $S_F^{[n]}(T)$ for the subspace of $S_F^n(T)$ consisting of non-redundant F-types on $x_0, \ldots, x_{n-1}$, i.e., those which contain $x_i \neq x_j$ for all $i \neq j$.

Definition 5.1. Let $p \in S_F^{[n]}(T)$ be a type realized in $M \models T$. An element $a \in M$ is called a **root** of p in M if a is an element of every tuple realizing p in M. We use the same terminology for quantifier-free types in $S_{\mathrm{qf}}^{[n]}(T)$.

Remark 5.2. If a type p has a unique realization in M, then p has a root in M (take any element of the unique tuple realizing p). When $n = 1$, the converse is true: a realized type $p(x) \in S_F^{[1]}(T)$ (or $S_{\mathrm{qf}}^{[1]}(T)$) has a root in M if and only if it has a unique realization in M.

Definition 5.3. Let $\chi(\overline{x})$ is a formula in F such that $\chi(\overline{x}) \to (\bigwedge_{i \neq j} x_i \neq x_j) \in T$. Then a model $M \models T$ is χ**-rooted** if every type $p(\overline{x}) \in S_F^{[n]}(T)$ which contains χ and is realized in M has a root in M. Again, we use the same terminology for quantifier-free formulas and types.

Remark 5.4. We note that the property of χ-rootedness is expressible by a sentence of $\mathcal{L}_{\omega_1,\omega}$, although not necessarily a sentence of F, which asserts that for every tuple $\overline{a}$ of distinct elements satisfying $\chi(\overline{x})$, there is some element a_i of the tuple such that every other tuple $\overline{b}$ with the same F-type as $\overline{a}$ contains a_i. Hence the set of χ-rooted models of T is a Borel set in Str_L.

Our goal is to prove the following theorem.

Theorem 5.5. *Let μ be a properly ergodic structure, F a countable fragment of $\mathcal{L}_{\omega_1,\omega}$, and $\chi(\overline{x})$ a formula in F such that $\mu(\chi(\overline{x})) > 0$ and $\chi(\overline{x}) \to (\bigwedge_{i\neq j} x_i \neq x_j) \in \mathrm{Th}_F(\mu)$. Suppose that $\mu(p) = 0$ for every F-type p containing $\chi(\overline{x})$. Then μ assigns measure 1 to the set of χ-rooted models of* $\mathrm{Th}_F(\mu)$.

The idea of the proof is as follows: We take an AHK representation of μ, sampling from which involves a family of i.i.d. random variables $(\xi_A)_{A\in\mathcal{P}_{\mathrm{fin}}(\mathbb{N})}$. For a set B with $0 \leq |B| \leq n$, we say that an n-type p is *likely given* $\widehat{\xi}_B$ if after conditioning on the random variables $(\xi_A)_{A\in\mathcal{P}(B)}$, the type p has a positive probability of being realized on a tuple containing all the elements of B (see Definition 5.7 below).

Now for $C \subseteq B$, it happens with probability 0 that a particular type p jumps from being not likely given $\widehat{\xi}_C$ to being likely given $\widehat{\xi}_B$. As a consequence, it is almost surely the case that for every type p, the family of sets $N(p) = \{C \mid p \text{ is likely given } \widehat{\xi}_C\}$ is closed under intersection: given sets A and B, the probability that the same type jumps from being not likely given $A\cap B$ to being likely given both A and B is 0, since $\widehat{\xi}_A$ and $\widehat{\xi}_B$ are conditionally independent over $\widehat{\xi}_{A\cap B}$.

Now for any type p containing χ, the family $N(p)$ contains all the sets on which p is realized, and it does not contain $\emptyset$ (since p has measure 0, and, by ergodicity, the random variable $\xi_\emptyset$ is irrelevant) so the intersection of all sets on which p is realized is almost surely nonempty, i.e., if p is realized, then it almost surely has a root.

Unfortunately, the need to handle all continuum-many types uniformly introduces some technical complications in formalizing this intuitive argument. We will now tackle those technicalities.

Suppose $(f_n)_{n\in\mathbb{N}}$ is an AHK representation of an invariant measure μ. It is a consequence of Lusin's theorem on measurable functions [Kec95, Theorem 17.12] that every measurable function differs from a Borel function on a set of measure 0. If g_n agrees with f_n almost everywhere, we may replace f_n by g_n in the AHK system and obtain another AHK representation of μ. Hence, we may assume that each f_n is Borel measurable.

We adopt the notation of §2.4 for the random variables $(\xi_A)_{A\in\mathcal{P}_{\mathrm{fin}}(\mathbb{N})}$: for a tuple $\overline{b}$, we will write $\widehat{\xi}_{\overline{b}}$ to denote the family of random variables $(\xi_A)_{A\subseteq\|\overline{b}\|}$. Similarly, we will write $\widehat{x}_B$ as a shorthand for a family of values $(x_A)_{A\in\mathcal{P}(B)} \in$

$[0,1]^{\mathcal{P}(B)}$. If $C \subseteq B$, we separate the family $\widehat{x}_B$ into $\widehat{x}_C = (x_A)_{A \in \mathcal{P}(C)}$ and $\widehat{x}_{B/C} = (x_A)_{A \in \mathcal{P}(B) \setminus \mathcal{P}(C)}$. If B is enumerated as a tuple $\overline{b}$, we will write $\widehat{x}_{\overline{b}}$.

Definition 5.6. Let $p \in S^{[n]}_{\mathrm{qf}}(L)$, and let $B \in \mathcal{P}_{\mathrm{fin}}(\mathbb{N})$, with $|B| = n$. Fix values $\widehat{x}_B \in [0,1]^{\mathcal{P}(B)}$. We say that p **is realized given** $\widehat{x}_B$ if there is some enumeration of B as a tuple $\overline{b}$ such that $p = f_n(\widehat{x}_{\overline{b}})$.

While the particular type $f_n(\widehat{x}_{\overline{b}})$ depends on the order in which B is enumerated as a tuple, the set of types which are realized given $\widehat{x}_B$ does not depend on the order. Since there are $n!$ ways of enumerating B as a tuple, at most $n!$ quantifier-free types are realized given $\widehat{x}_B$. Recall that part of the definition of an AHK system is that these types almost surely form an orbit under the action of $\mathrm{Sym}(n)$ on $S^{[n]}_{\mathrm{qf}}(L)$ by permuting variables.

The set $R(B) = \{(p, \widehat{x}_B) \mid p \text{ is realized given } \widehat{x}_B\} \subseteq S^{[n]}_{\mathrm{qf}}(L) \times [0,1]^{\mathcal{P}(B)}$ is Borel. Indeed,

$$R(B) = \bigcup_{\overline{b} \text{ enumerating } B} \{(f_n(\widehat{x}_{\overline{b}}), \widehat{x}_B) \mid \widehat{x}_B \in [0,1]^{\mathcal{P}(B)}\},$$

and the graph of a Borel function is a Borel set.

Let $C \in \mathcal{P}_{\mathrm{fin}}(\mathbb{N})$, with $0 \leq |C| \leq n$, and pick some $C \subseteq B$ with $|B| = n$. Let $1_{R(B)}$ be the indicator function of the event $R(B)$. The Fubini–Tonelli theorem for Borel measurable functions [Tao11, Theorem 1.7.15] tells us that for all $p \in S^{[n]}_{\mathrm{qf}}(L)$ and $\widehat{x}_C \in [0,1]^{\mathcal{P}(C)}$, the integral

$$\int_{\widehat{x}_{B/C} \in [0,1]^{\mathcal{P}(B) \setminus \mathcal{P}(C)}} 1_{R(B)}(p, \widehat{x}_C, \widehat{x_{B/C}})\, d\lambda_0^{\mathcal{P}(B) \setminus \mathcal{P}(C)}$$

is defined (here λ_0 is the Lebesgue measure on $[0,1]$, restricted to the Borel σ-algebra), and that the function

$$P_C \colon (p, \widehat{x}_C) \mapsto \int_{\widehat{x}_{B/C} \in [0,1]^{\mathcal{P}(B) \setminus \mathcal{P}(C)}} 1_{R(B)}(p, \widehat{x}_C, \widehat{x}_{B/C})\, d\lambda_0^{\mathcal{P}(B) \setminus \mathcal{P}(C)}$$

is Borel measurable. Abusing terminology somewhat, we call $P_C(p, \widehat{x}_C)$ the **probability of** p **given** $\widehat{x}_C$.

Observe that the definition of P_C is independent of the choice of B, since if $C \subseteq B'$ and $f \colon B \to B'$ is a bijection fixing C, the induced map $S^{[n]}_{\mathrm{qf}}(L) \times [0,1]^{\mathcal{P}(B)} \to S^{[n]}_{\mathrm{qf}}(L) \times [0,1]^{\mathcal{P}(B')}$ carries $R(p, B)$ to $R(p, B')$.

Definition 5.7. With notation as above, we say that p is **likely given** $\widehat{x}_C$ if the probability of p given $\widehat{x}_C$ is positive. For fixed values $\widehat{x}_C$, we denote by $S(\widehat{x}_C)$ the set of non-redundant quantifier-free n-types which are likely given $\widehat{x}_C$.

We'd like to show that $S(\widehat{x}_C)$ is always countable. We'll need the following basic measure theory lemma.

Lemma 5.8. *Let $(\Omega, \mathcal{F}, \nu)$ be a probability space, and let $(E_i)_{i\in I}$ be an uncountable family of events, each of positive measure. Then there is some $x \in \Omega$ such that x is in infinitely many of the E_i.*

Proof. Since there are uncountably many events in the family, there is some $\varepsilon > 0$ such that infinitely many have measure at least ε. Let $\{E_{i_n} \mid n \in \mathbb{N}\}$ be a countable sequence with $\nu(E_{i_n}) \geq \varepsilon$ for all n. Then define $E'_N = \bigcup_{n\geq N} E_{i_n}$, for $n \in N$. We have $\nu(E'_N) \geq \varepsilon$. By continuity, $\nu(\bigcap_{N\in\mathbb{N}} E'_N) \geq \varepsilon$, so there is some $x \in \bigcap_{N\in\mathbb{N}} E'_N$. This x is in infinitely many of the E_{i_n}. □

Lemma 5.9. *For any set C with $0 \leq |C| \leq n$, and any $\widehat{x}_C \in [0,1]^{\mathcal{P}(C)}$, the set $S(\widehat{x}_C)$ is countable.*

Proof. Pick some $C \subseteq B$ with $|B| = n$. A type p is likely given $\widehat{x}_C$ if and only if the event $E_p = \{\widehat{x}_{B/C} \mid (p, \widehat{x}_C, \widehat{x}_{B/C}) \in R(B)\} \subseteq [0,1]^{\mathcal{P}(B)\setminus\mathcal{P}(C)}$ has positive measure. Since any point $\widehat{x}_{B/C}$ is in at most $n!$ of the events E_p (corresponding to the types $f_n(\widehat{x}_{\bar{b}})$ for the $n!$ enumerations of B as a tuple), by Lemma 5.8, the set $S(\widehat{x}_C) = \{p \mid E_p \text{ has positive measure}\}$ is countable. □

Lemma 5.10. *For sets $D \subseteq C$ with $0 \leq |D| \leq |C| \leq n$, and any $\widehat{x}_D \in [0,1]^{\mathcal{P}(D)}$, if p is not likely given $\widehat{x}_D$, then the set $\{\widehat{x}_{C/D} \mid p \text{ is likely given } \widehat{x}_C\}$ has measure 0 in $[0,1]^{\mathcal{P}(C)\setminus\mathcal{P}(D)}$.*

Proof. This is just the Fubini–Tonelli theorem. Pick some $C \subseteq B$ with $|B| = n$ (recall that p is an n-type). Since p is not likely given $\widehat{x}_D$, we have

$$\begin{aligned} 0 &= \int_{\widehat{x}_{B/D}} 1_{R(B)}(p, \widehat{x}_D, \widehat{x}_{B/D})\, d\lambda_0^{\mathcal{P}(B)\setminus\mathcal{P}(D)} \\ &= \int_{\widehat{x}_{C/D}} \Big(\int_{\widehat{x}_{B/C}} 1_{R(B)}(p, \widehat{x}_D, \widehat{x}_{C/D}, \widehat{x}_{B/C})\, d\lambda_0^{\mathcal{P}(C)\setminus\mathcal{P}(D)} \Big) d\lambda_0^{\mathcal{P}(B)\setminus\mathcal{P}(C)}. \end{aligned}$$

And the interior integral, which is 0 for almost all values of $\widehat{x}_{C/D}$, is the probability of p given $\widehat{x}_C$. □

Lemma 5.11. *Fix sets $A, B \in \mathcal{P}_{\mathit{fin}}(\mathbb{N})$, with $0 \leq |A| \leq n$ and $0 \leq |B| \leq n$, and let $C = A \cap B$. Let $\widehat{\xi}_C$, $\widehat{\xi}_A = (\widehat{\xi}_C, \widehat{\xi}_{A/C})$ and $\widehat{\xi}_B = (\widehat{\xi}_C, \widehat{\xi}_{B/C})$ be our random variables on these sets. Almost surely, every quantifier-free type which is likely given $\widehat{\xi}_A$ and likely given $\widehat{\xi}_B$ is likely given $\widehat{\xi}_C$. That is, $S(\widehat{\xi}_C) \subseteq S(\widehat{\xi}_A) \cap S(\widehat{\xi}_B)$.*

Proof. Observe that for any set D, the set $\{(p, \widehat{x}_D) \mid p \text{ is likely given } \widehat{x}_D\} \subseteq S_{\mathrm{qf}}^{[n]}(L) \times [0,1]^{\mathcal{P}(D)}$ is Borel. Indeed, it is the preimage of $(0,1]$ under the Borel function P_D. It follows that the set

$$\{(p, \widehat{x}_C, \widehat{x}_{A/C}, \widehat{x}_{B/C}) \mid p \text{ is likely given } \widehat{x}_A \text{ and } \widehat{x}_B \text{ but not } \widehat{x}_C\}$$

is a Borel subset of $S_{\mathrm{qf}}^{[n]}(L) \times [0,1]^{\mathcal{P}(C)} \times [0,1]^{\mathcal{P}(A)\setminus\mathcal{P}(C)} \times [0,1]^{\mathcal{P}(B)\setminus\mathcal{P}(C)}$. Projecting out the first coordinate, we see that the set

$$X^{A,B} = \{(\widehat{x}_C, \widehat{x}_{A/C}, \widehat{x}_{B/C}) \mid \text{some } p \text{ is likely given } \widehat{x}_A \text{ and } \widehat{x}_B \text{ but not } \widehat{x}_C\}$$

is analytic, and hence measurable (see [Kec95, Theorem 21.10]). Since our random variables are i.i.d. Lebesgue on $[0,1]$, we would like to show that $X^{A,B}$ has measure 0 with respect to the Lebesgue measure on $[0,1]^{\mathcal{P}(C)} \times [0,1]^{\mathcal{P}(A)\setminus\mathcal{P}(C)} \times [0,1]^{\mathcal{P}(B)\setminus\mathcal{P}(C)}$. The knowledge that $X^{A,B}$ is measurable enables us to analyze its measure fiber-wise, using the Fubini–Tonelli theorem (the version for complete measures this time, [Tao11, Theorem 1.7.18]).

So consider the fiber $X^{A,B}_{\widehat{x}_A}$ over $\widehat{x}_A = (\widehat{x}_C, \widehat{x}_{A/C}) \in [0,1]^{\mathcal{P}(C)} \times [0,1]^{\mathcal{P}(A)\setminus\mathcal{P}(C)}$.

$$\begin{aligned} X^{A,B}_{\widehat{x}_A} &= \{\widehat{x}_{B/C} \mid \text{some } p \text{ is likely given } \widehat{x}_A \text{ and } \widehat{x}_B \text{ but not } \widehat{x}_C\} \\ &= \bigcup_{p \in S(\widehat{x}_A)\setminus S(\widehat{x}_C)} \{\widehat{x}_{B/C} \mid p \text{ is likely given } \widehat{x}_B\} \end{aligned}$$

By Lemma 5.9, this is a countable union, and by Lemma 5.10, each set in the union has measure 0, so $X^{A,B}_{\widehat{x}_A}$ has measure 0, and, by Fubini–Tonelli, $X^{A,B}$ has measure 0. □

Proof of Theorem 5.5. We have a properly ergodic structure μ, a countable fragment F of $\mathcal{L}_{\omega_1,\omega}$, and a distinguished formula $\chi(\overline{x})$ in F. Let n be the length of the tuple $\overline{x}$. Let L', T', and Q be the language, Π_2^- theory, and countable set of partial quantifier-free types, respectively obtained from Theorem 2.24 for the fragment F and empty theory T. By Corollary 2.25, μ corresponds to an ergodic L'-structure μ', concentrated on those models of T' that omit all the types in Q.

For such models, each formula $\varphi(\overline{y})$ in F is equivalent to the atomic L'-formula $R_\varphi(\overline{y})$, so we have $\mu'(R_\chi(\overline{x})) > 0$, and for every quantifier-free type q containing $R_\chi(\overline{x})$, we have $\mu'(q) = 0$. It suffices to show that an L'-structure $\mathfrak{M}$ sampled from μ' is almost surely $R_\chi(\overline{x})$-rooted with respect to quantifier-free types.

By Theorem 2.27, μ' has an AHK representation $(f_m)_{m\in\mathbb{N}}$. And since μ' is ergodic, by Theorem 2.28, we can pick the functions f_m so they do not depend on the argument indexed by $\emptyset$. As a consequence, $S(\widehat{\xi_\emptyset}) = \{p \mid \mu'(p) > 0\}$. Indeed, the probability of a type p given $\widehat{\xi_\emptyset}$ is obtained by integrating out all of the variables except $\xi_\emptyset$, which is irrelevant to f_n, so it is simply the probability that p is realized on an arbitrary set of size n.

On the other hand, if $|B| = n$, then $S(\widehat{\xi_B})$ is almost surely equal to the set of quantifier-free types realized on the $n!$ tuples $\overline{b}$ enumerating B. Indeed, the probability of a type p given $\widehat{\xi_B}$ is simply the indicator function $1_{R(B)}$ (no variables are integrated out).

Now for any tuple $\overline{b}$, letting $B = \|\overline{b}\|$, if $\overline{b}$ satisfies a type p containing R_χ, then almost surely the family of sets C such that $p \in S(\widehat{\xi_C})$ contains B, does not contain $\emptyset$, and is closed under intersection (by Lemma 5.11). In particular, the intersection of this family is non-empty. Therefore, almost surely, $\mathfrak{M}$ is χ-rooted. □

We conclude this section with a discussion of the tension between rootedness and trivial definable closure.

Let M be a χ-rooted model of an F-theory T. Suppose that $p(\overline{x}) \in S_F^{[n]}(T)$ contains $\chi(\overline{x})$ and is realized in M, and let a be a root of p in M. If $M \models p(a, \overline{b})$, then a is the unique element of M satisfying $p(x, \overline{b})$, since if $c \neq a$, then a is not in $c\overline{b}$, so $c\overline{b}$ does not realize p. This implies that M has non-trivial group-theoretic definable closure, since every automorphism of M fixing $\overline{b}$ also fixes a. Note that T may still have trivial definable closure, since p is an F-type and, in general, is not equivalent to a formula in F.

We can conclude, however, that if an F-theory T with trivial definable closure has a χ-rooted model, then no non-redundant type that contains χ is isolated. Thus isolated types are not dense in $S_F^n(T)$. By standard facts about model theory in countable fragments of $\mathcal{L}_{\omega_1,\omega}$ (see [KK04]), this implies that T does not have a prime model with respect to F-elementary embeddings, and that there are continuum-many types in $S_F^n(T)$ containing $\chi(\overline{x})$.

Theorem 5.5 implies that the theory of every properly ergodic structure μ exhibits this behavior: using Theorem 4.7 to obtain a countable fragment F and an F-formula $\chi(\overline{x})$ of positive measure such that every type containing χ has measure 0, the theory $\mathrm{Th}_F(\mu)$ has many χ-rooted models and (by Theorem 2.19) trivial dcl.

Of course, given any particular F-type p containing $\chi(\overline{x})$, we can try to bring χ-rootedness into direct conflict with trivial dcl by moving to a larger countable fragment F' which contains the formula $\theta_p(\overline{x}) := \bigwedge_{\varphi\in p} \varphi(\overline{x})$ isolating p. But since p has measure 0, the theory $\mathrm{Th}_{F'}(\mu)$ contains the sentence $\forall \overline{x}\, \neg\theta_p(\overline{x})$, ruling out troublesome realizations of p.

Of course, a countable fragment F' can only isolate and rule out countably many of the continuum-many types of measure 0 containing $\chi(\overline{x})$. For example, given the kaleidoscope random graph (Example 3.2), we could extend from the first-order fragment to a countable fragment F of $\mathcal{L}_{\omega_1,\omega}$ containing some of the conjunctions $\bigwedge_{n\in A} xR_ny \wedge \bigwedge_{n\notin A} \neg xR_ny$, for $A \in 2^{\mathbb{N}}$. Then the theory $\mathrm{Th}_{F'}(\mu)$ is essentially the same as $\mathrm{Th}_{\mathrm{FO}}(\mu)$, but with countably many of the continuum-many quantifier-free 2-types forbidden.

6. Constructing properly ergodic structures

In this section, given an F-theory T having trivial dcl, we will use a single χ-rooted model M of T to construct a properly ergodic model of T. The strategy is to build a Borel structure $\mathbb{M}$ equipped with a probability measure ν, via an inverse limit of finite probability spaces. We use M as a guide in the construction to ensure that $\mathbb{M}$ is also χ-rooted. Then our ergodic structure μ will be obtained by i.i.d. sampling of countably many points from $\mathbb{M}$ according to ν and taking the induced substructure.

Having built the Borel structure $\mathbb{M}$, we proceed to rescale ν, using a technique from [AFKP17], to obtain not just one but continuum-many properly ergodic structures concentrating on T.

Definition 6.1. A **Borel structure** $\mathbb{M}$ is an L-structure whose domain is a standard Borel space such that for every relation symbol R of arity $\mathrm{ar}(R)$ in L, the subset $R \subseteq M^{\mathrm{ar}(R)}$ is Borel. A **measured structure** is a Borel structure $\mathbb{M}$ equipped with an atomless probability measure ν.

Given a measured structure $(\mathbb{M}, \nu)$, there is a canonical measure $\mu_{\mathbb{M},\nu}$ on Str_L, obtained by sampling a countable ν-i.i.d. sequence (of almost surely distinct points) from $\mathbb{M}$ and taking the induced substructure. Somewhat more formally, $\mu_{\mathbb{M},\nu}$ is the distribution of a random structure in Str_L whose atomic diagram on $\mathbb{N}$ is given by that of the random substructure of $\mathbb{M}$ with underlying set $\{a_i \mid i \in \mathbb{N}\}$, where $(a_i)_{i\in\mathbb{N}}$ is a ν-i.i.d. sequence of (almost surely unique) elements in $\mathbb{M}$.

We now describe an AHK representation of the measures $\mu_{\mathbb{M},\nu}$ in the sense of §2.4. Choose a measure-preserving Borel isomorphism h from $[0, 1]$ equipped with the uniform measure to the domain of $\mathbb{M}$ equipped with ν, and for each n let $\star_n$ be an arbitrary element of $S^{[n]}_{\mathrm{qf}}(L)$. Then define functions $f_n \colon [0, 1]^{\mathcal{P}_{\mathrm{fin}}([n])} \to S^{[n]}_{\mathrm{qf}}(L)$ by

$$f_n\big((\xi_A)_{A\subseteq[n]}\big) = \begin{cases} \mathrm{qftp}(h(\xi_{\{0\}}), \dots, h(\xi_{\{n-1\}})) & \text{if } \xi_{\{i\}} \neq \xi_{\{j\}} \text{ for } i < j \in [n]; \\ \star_n & \text{otherwise.} \end{cases}$$

Informally, these functions ignore the random variables ξ_A when $|A| \neq 1$ and view the $(\xi_{\{a\}})_{a\in\mathbb{N}}$ as independent random variables with distribution ν taking their values in $\mathbb{M}$.

Now $(f_n)_{n\in\mathbb{N}}$ is an AHK system, so it induces an invariant measure on Str_L. This measure is clearly the same as $\mu_{\mathbb{M},\nu}$ described above via sampling of a random substructure. Since the f_n do not depend on the argument indexed by $\emptyset$, the measure $\mu_{\mathbb{M},\nu}$ is ergodic (Theorem 2.28), which establishes the following lemma.

Lemma 6.2. *Given a measured structure* $(\mathbb{M}, \nu)$*, the measure* $\mu_{\mathbb{M},\nu}$ *on* Str_L *is an ergodic structure.*

In fact, this AHK system is "random-free". This terminology comes from the world of graphons: a graphon is said to be **random-free** [Jan13, §10] when it is $\{0, 1\}$-valued almost everywhere. This can be thought of as "having randomness" only at the level of vertices (and not at higher levels — namely edges, in the case of graphs). See also 0–1 valued graphons in [LS10] and the simple arrays of [Kal99]. A graphon is random-free if and only if the corresponding AHK system is random-free in the following sense.

Definition 6.3. An AHK system $(f_n)_{n\in\mathbb{N}}$ is **random-free** if each function f_n depends only on the singleton variables $\xi_{\{a\}}$ for $a \in \mathbb{N}$. An ergodic structure μ is **random-free** if it has a random-free AHK representation.

We would like to transfer properties of $\mathbb{M}$ to almost-sure properties of $\mu_{\mathbb{M},\nu}$. It is not true in general that $\mu_{\mathbb{M},\nu} \models \mathrm{Th}_F(\mathbb{M})$. But the following property will allow us to transfer satisfaction in $\mathbb{M}$ to satisfaction in $\mu_{\mathbb{M},\nu}$ for Π_2^- sentences.

Definition 6.4. Let $(\mathbb{M}, \nu)$ be a measured structure, and let φ be a Π_2^- sentence. We say that $(\mathbb{M}, \nu)$ **satisfies** φ **with strong witnesses** (or **has strong witnesses for** φ) if the following hold.

- If φ is universal, then $\mathbb{M} \models \varphi$.
- If φ is pithy Π_2, i.e., of the form $\forall \overline{x}\, \exists y\, \rho(\overline{x}, y)$, then for every tuple $\overline{a}$ from $\mathbb{M}$, the set $\rho(\overline{a}, \mathbb{M}) = \{b \in \mathbb{M} \mid \rho(\overline{a}, b)\}$ either contains an element of the tuple $\overline{a}$ or has positive ν-measure.

For a Π_2^- theory T, we say that $(\mathbb{M}, \nu)$ **satisfies** T **with strong witnesses** when it satisfies φ with strong witnesses for all $\varphi \in T$.

Note that if $(\mathbb{M}, \nu)$ satisfies T with strong witnesses, then $\mathbb{M} \models T$.

Lemma 6.5. *Let $(\mathbb{M}, \nu)$ be a measured structure, and let $\mu = \mu_{\mathbb{M},\nu}$.*

(i) Let Q be a countable set of partial quantifier-free types. If $\mathbb{M}$ omits all the types in Q, then μ almost surely omits all the types in Q.
(ii) Let T be a Π_2^- theory. If $(\mathbb{M}, \nu)$ satisfies T with strong witnesses, then μ almost surely satisfies T.
(iii) Further, if there is a quantifier-free formula $\chi(\overline{x})$ such that $\mathbb{M}$ is χ-rooted with respect to quantifier-free types, then μ is properly ergodic.

Proof. (i) If no tuple from $\mathbb{M}$ realizes a quantifier-free type $q \in Q$, then no tuple from any countable substructure sampled from $\mathbb{M}$ realizes q.

(ii) Every universal sentence $\forall \overline{x}\, \psi(\overline{x})$ in T is almost surely satisfied by μ, since every tuple $\overline{v}$ from $\mathbb{M}$ satisfies the quantifier-free formula $\psi(\overline{x})$.

Next, consider sentences of the form $\forall \overline{x}\, \exists y\, \rho(\overline{x}, y)$. Fix an n-tuple $\overline{a}$ from $\mathbb{N}$, where n is the length of $\overline{x}$. Corresponding to this tuple, we have a random tuple $\overline{v} := (v_{a_1}, \ldots, v_{a_n})$ sampled from $\mathbb{M}$. By the Fubini–Tonelli theorem, it suffices to show that for a measure one collection of values of this random tuple (e.g., those for which coordinates indexed by distinct natural numbers take distinct values), there is almost surely some $b \in \mathbb{N}$ such that $\mathbb{M} \models \rho\big(\overline{v}, v_b\big)$.

By strong witnesses, $\rho(\overline{v}, \mathbb{M})$ either contains an element v_{a_i} of the tuple $\overline{v}$ or has positive measure. In the first case, v_{a_i} serves as our witness. In the second case, since there are infinitely many other independent random elements $(v_b)_{b\in\mathbb{N}\setminus\|\overline{a}\|}$, almost surely infinitely many of them land in the set $\rho(\overline{c}, \mathbb{M})$.

(iii) By *(ii)*, $\mu \models T$, and since $\chi(\overline{x}) \wedge (\bigwedge_{i\neq j} x_i \neq x_j)$ is consistent with T, $\mu(\chi(\overline{x})) > 0$. Let p be any type containing $\chi(\overline{x})$, and let q be its restriction to

the quantifier-free formulas. To show that $\mu(p) = 0$, it suffices to show that $\mu(q) = 0$.

Now since $\mathbb{M}$ is χ-rooted with respect to quantifier-free types, q has a root v in $\mathbb{M}$. The probability that a tuple sampled from $\mathbb{M}$ satisfies q is bounded above by the probability that the tuple contains the root v. This probability is 0, since the measure ν is atomless. Hence, by Theorem 4.7, μ is properly ergodic. $\square$

Thus, after applying the Π_2^- transformation from §2.3 to an F-theory T, we have reduced the problem of constructing a properly ergodic structure almost surely satisfying T to that of constructing a measured structure with the properties in Lemma 6.5.

Theorem 6.6. *Let F be a countable fragment of $\mathcal{L}_{\omega_1,\omega}$, let T be a complete F-theory with trivial dcl, let $\chi(\overline{x})$ be a formula in F, and let M be a χ-rooted model of T. Then there are continuum-many properly ergodic structures μ such that $\mu \models T$.*

Proof. We begin by applying Theorem 2.24 to obtain a language $L' \supseteq L$, a Π_2^- theory T', and a countable set of partial quantifier-free types Q. Let M' be the natural expansion of M to an L'-structure. Then M' is R_χ-rooted, where $R_\chi(\overline{x})$ is the atomic L'-formula corresponding to the L-formula $\chi(\overline{x})$. By Corollary 2.25, it suffices to construct a properly ergodic L'-structure which almost surely satisfies T' and omits the types in Q.

Part 1: The inverse system

We construct a sequence $(A_k)_{k\in\mathbb{N}}$ of finite L'-structures, each of which is identified with a substructure of M'. Given a structure A, we define the structure A^* to have underlying set $A \cup \{*\}$, where no new relations hold involving $*$. For each k, we equip the underlying set of each A_k^* with a discrete probability measure ν_k that assigns positive measure to every element, and we fix a finite sublanguage L_k of L'. Finally, we define connecting maps $g_k\colon A_{k+1}^* \to A_k^*$ such that $g_k(*) = *$ for all k, which preserve the measures and certain quantifier-free types, as follows:

(1) $\nu_{k+1}(g_k^{-1}[X]) = \nu_k(X)$ for all $X \subseteq A_k^*$.
(2) If $\overline{a}$ is a tuple of distinct elements from A_{k+1} such that $g_k(\overline{a})$ is a tuple of distinct elements of A_k, then $\mathrm{qftp}_{L_k}(\overline{a}) = \mathrm{qftp}_{L_k}(g_k(\overline{a}))$. Note that we make no requirement if g_k is not injective on $\|\overline{a}\|$ or if any element of $\overline{a}$ is mapped to $*$.

We enumerate all pithy Π_2 sentences in T' as $\langle\varphi_k\rangle_{k\in\mathbb{N}}$ and the types in Q as $\langle q_i\rangle_{k\in\mathbb{N}}$ with redundancies, so that each sentence and each type appears infinitely often in its list. We also enumerate the symbols in the language L' as $\langle R_k\rangle_{k\in\mathbb{N}}$.

At stage 0, we start with $A_0 = \emptyset$, the empty substructure of M'. Then $A_0^* = \{*\}$, and we set $\nu_0(\{*\}) = 1$ and $L_0 = \emptyset$.

At stage $k+1$, we are given A_k, ν_k, and L_k. We define A_{k+1}, ν_{k+1}, L_{k+1}, and the connecting map g_k in four steps.

Step 1: Splitting the elements of A_k.

Enumerate the elements of A_k as $\langle a_1, \dots, a_m \rangle$. We build intermediate substructures $B_i = \{a_1, \dots, a_m, a'_1, \dots, a'_i\}$ of M', where each new element a'_j is a "copy" of a_j to be defined. We start with $B_0 = A_k$.

Given B_i, let $\varphi_{B_i}(x_1, \dots, x_m, x'_1, \dots, x'_i)$ be the conjunction of all atomic and negated atomic L_k formulas holding on B_i, so that φ_{B_i} encodes the quantifier-free L_k-type of B_i. Now there is an L-formula ψ_{B_i} in F such that ψ_{B_i} has the same realizations as φ_{B_i} in M'. Since $T = \mathrm{Th}_F(M)$ has trivial dcl, we can find another realization $a'_{i+1} \neq a_{i+1}$ of $\psi_{B_i}(a_1, \dots, x_{i+1}, \dots, a_m, a'_1, \dots, a'_i)$ in $M' \setminus B_i$. Set $B_{i+1} = B_i \cup \{a'_{i+1}\}$. We have

$$(\dagger)\quad \mathrm{qftp}_{L_k}(a_1, \dots, a_{i+1}, \dots, a_m, a'_1, \dots, a'_i) = \mathrm{qftp}_{L_k}(a_1, \dots, a'_{i+1}, \dots, a_m, a'_1, \dots, a'_i).$$

At the end of Step 1, we have a structure $B_m = \{a_1, \dots, a_m, a'_1, \dots, a'_m\}$.

Step 2: Splitting $*$.

The pithy Π_2 sentence φ_k has the form $\forall \overline{x}\, \exists y\, \rho(\overline{x}, y)$, where $\overline{x}$ is a tuple of length j and $\rho(\overline{x}, y)$ is quantifier-free. Suppose there is a tuple $\overline{a}$ from B_m such that $B_m \models \neg\exists y\, \rho(\overline{a}, y)$. Then, since $M' \models \exists y\, \rho(\overline{a}, y)$, we can choose some witness $c_{\overline{a}}$ to the existential quantifier in $M' \setminus B_m$. Let $W = \{c_{\overline{a}} \mid \overline{a} \in B_m^j \text{ and } B_m \models \neg\exists y\, \rho(\overline{a}, y)\}$ be the (finite) set of chosen witnesses. Note that if $\overline{x}$ is the empty tuple of variables, then W is either empty or consists of a single witness, depending on whether $B_m \models \exists y\, \rho(y)$.

Let $A_{k+1} = B_m \cup W$ if W is non-empty, and otherwise let $A_{k+1} = B_m \cup \{c\}$, where c is any new element in $M' \setminus B_m$.

Step 3: Defining g_k and ν_{k+1}.

Recall that g_k is to be a map from A^*_{k+1} to A^*_k. We set $g_k(a_i) = g_k(a'_i) = a_i$ and $g_k(c) = g_k(*) = *$ for $c \in A_{k+1} \setminus B_m$.

We define ν_{k+1} by splitting the measure of an element of A^*_k evenly among its preimages under g_k. So $\nu_{k+1}(a_i) = \nu_{k+1}(a'_i) = \frac{1}{2}\nu_k(a_i)$, and $\nu_{k+1}(c) = \nu_{k+1}(*) = \frac{1}{N}\nu_k(*)$, where $N = |A^*_{k+1} \setminus B_m| \geq 2$. Note that every element of A^*_{k+1} has positive measure, by induction.

Step 4: Defining L_{k+1}.

We expand the current language L_k to L_{k+1} by adding finitely many new symbols from L'.

(a) Add R_k to L_{k+1} if it is not already included.
(b) Since A_{k+1} is a substructure of M', no tuple from A_{k+1} realizes q_k. That is, for every tuple $\overline{a}$ from A_{k+1}, there is some quantifier-free formula $\varphi_{\overline{a}}(\overline{x}) \in q_k$ such that $M' \models \neg\varphi_{\overline{a}}(\overline{a})$. Add the finitely many relation symbols appearing in $\varphi_{\overline{a}}$ to L_{k+1}.

(c) Let n be the number of free variables in $\chi(\overline{x})$. For every pair of n-tuples $\overline{a}$ and $\overline{b}$ from A_{k+1} that realize distinct quantifier-free L'-types in M', there is some relation symbol $R_{\overline{a},\overline{b}}$ that separates their types. Add $R_{\overline{a},\overline{b}}$ to L_{k+1}.

This completes stage $k+1$ of the construction. Let us check that conditions (1) and (2) above are satisfied by the connecting map g_k.

(1): Since ν_k and ν_{k+1} are discrete measures on finite spaces, it suffices to check that $\nu_k(a) = \sum_{b \in g_k^{-1}[\{a\}]} \nu_{k+1}(b)$ for every singleton $a \in A_k^*$. This follows immediately from our definitions of g_k and ν_{k+1}.

(2): Let $\overline{b}$ be a tuple from A_{k+1}. The assumption that $g_k(\overline{b})$ is a tuple of distinct elements of A_k means that every element of $\overline{b}$ is in B_m (since the other elements are mapped to $*$) and that a_i and a_i' are not both in $\overline{b}$ for any i. For any function $\gamma\colon [m] \to [2]$, let $\overline{a}^\gamma$ be the m-tuple which contains a_i if $\gamma(i) = 0$ and a_i' if $\gamma(i) = 1$. Then, expanding $\overline{b}$ to an m-tuple of the form $\overline{a}^\gamma$, it suffices to show that $\mathrm{qftp}_{L_k}(\overline{a}^\gamma) = \mathrm{qftp}_{L_k}(g_k(\overline{a}^\gamma)) = \mathrm{qftp}_{L_k}(\overline{a})$. This follows by several applications of instances of the equality (†) above.

Part 2: The measured structure

Let $\mathbb{X}$ be the inverse limit of the system of sets A_k^* and surjective connecting maps g_k. For each k, let π_k be the projection map $\mathbb{X} \to A_k \cup \{*\}$. Then $\mathbb{X}$ is a profinite set, so it has a natural topological structure as a Stone space, in which the basic clopen sets are exactly the preimages under the maps π_k of subsets of the sets A_k^*. Note that $\mathbb{X}$ is separable, so it is a standard Borel space.

Let ν^* be the finitely additive measure on the Boolean algebra $\mathcal{B}^*$ of clopen subsets of $\mathbb{X}$ defined by $\nu^*(\pi_k^{-1}[X]) = \nu_k(X)$. This is well defined by condition (1). By the Hahn–Kolmogorov Measure Extension Theorem [Tao11, Theorem 1.7.8], ν^* extends to a Borel probability measure ν on $\mathbb{X}$.

Now each element a of A_k^* has at least 2 preimages in A_{k+1}^*, each of which have measure at most $\frac{1}{2}\nu_k(a)$. Hence, by induction, the measure of each element of A_k^* is at most 2^{-k}. So for all $x \in \mathbb{X}$, the point x is contained in a basic clopen set $X_k = \pi_k^{-1}[\{\pi_k(x)\}]$ with $\nu(X_k) \leq 2^{-k}$ for all k. This implies that $\nu(\{x\}) = 0$ and ν is non-atomic.

Note that there is a unique element $*$ of $\mathbb{X}$ with the property that $\pi_k(*) = *$ for all k. We define a Borel L'-structure $\mathbb{M}$ with domain $\mathbb{X} \setminus \{*\}$ (which is also a standard Borel space). Since we have only removed a measure 0 set from $\mathbb{X}$, the probability measure ν on $\mathbb{X}$ restricts to a probability measure on $\mathbb{M}$, which we also call ν.

We define the structure on $\mathbb{M}$ by specifying the quantifier-free type of every tuple of distinct elements from $\mathbb{M}$. By Step 4 (a), $\bigcup_{k=0}^\infty L = L'$. Given a tuple $\overline{a}$ of distinct elements from $\mathbb{M}$ and a quantifier-free formula $\varphi(\overline{x})$, we choose k large enough so that L_k contains all of the relation symbols appearing in $\varphi(\overline{x})$

and so that $\pi_k(\overline{a})$ is a tuple of distinct elements from A_k. We set $\mathbb{M} \models \varphi(\overline{a})$ if and only if $A_k \models \varphi(\pi_k(\overline{a}))$. This is well-defined by condition (2).

By Step 4 (a), $\bigcup_{k=0}^{\infty} L_k = L'$. Given a tuple $\overline{a}$ from $\mathbb{M}$ and a symbol R in L', we choose k large enough so that $R \in L_k$ and distinct elements of $\overline{a}$ are mapped by π_k to distinct elements of A_k. We set $\mathbb{M} \models R(\overline{a})$ if and only if $A_k \models R(\pi_k(\overline{a}))$. This is well-defined by condition (2).

According to this definition, to determine whether a quantifier-free formula $\varphi(\overline{x})$ holds of a tuple $\overline{a}$ with repeated elements, we can remove the redundancies from $\overline{a}$ and replace the corresponding variables in $\overline{x}$. For example, if $a_i = a_j$, we can remove a_j and replace instances of x_j in $\varphi(\overline{x})$ with x_i. This is equivalent to choosing k large enough so that distinct elements of $\overline{a}$ are mapped by π_k to distinct elements of A_k and checking whether $A_k \models \varphi(\pi_k(\overline{a}))$.

The interpretation of a relation symbol R is then a Borel subset of $\mathbb{M}^{\mathrm{ar}(R)}$. Indeed, fixing k, the set of tuples $\overline{a}$ such that distinct elements of $\overline{a}$ are mapped by π_k to distinct elements of A_k and $\pi_k(\overline{a})$ satisfies R is closed (the finite union of certain boxes intersected with certain diagonals), and the interpretation of R is the countable union (over k) of these sets. Hence $\mathbb{M}$ is a Borel structure.

We now verify the conditions of Lemma 6.5 for the measured structure $(\mathbb{M}, \nu)$, the Π_2^- theory T', the quantifier-free types Q, and the quantifier-free formula $R_\chi(\overline{x})$.

(i) $\mathbb{M}$ omits all the types in Q.

Let $q(\overline{x})$ be a type in Q, and let $\overline{a}$ be a tuple from $\mathbb{M}$. Let k be large enough so that $\pi_k(\overline{a})$ is a tuple of distinct elements of A_k. Since q appears infinitely many times in our enumeration of Q, there is some $l > k$ such that $q = q_l$. Then $\overline{b} := \pi_{l+1}(\overline{a})$ is also a tuple of distinct elements of A_{l+1}. In Step 4 (b) of stage $l+1$ of the construction, we ensured that L_{l+1} includes the relation symbols appearing in a quantifier-free formula $\varphi_{\overline{b}}(\overline{x}) \in q_k$ such that $A_{l+1} \models \neg\varphi_{\overline{b}}(\overline{b})$. Then also $\mathbb{M} \models \neg\varphi_{\overline{b}}(\overline{a})$, and hence $\overline{a}$ does not realize q.

(ii) $(\mathbb{M}, \nu)$ satisfies T' with strong witnesses.

Let φ be a Π_2^- sentence in T'. Then φ has the form $\forall \overline{x}\, \psi(\overline{x})$, where $\psi(\overline{x})$ is either quantifier-free or has a single existential quantifier. Let $\overline{a}$ be a tuple from $\mathbb{M}$. Let k be large enough so that all the symbols in φ are in L_k and $\pi_k(\overline{a})$ is a tuple of disjoint elements of A_k.

If $\psi(\overline{x})$ is quantifier-free, then $\mathbb{M} \models \psi(\overline{a})$ if and only if $A_k \models \psi(\pi_k(\overline{a}))$. The latter holds, since A_k is a substructure of $\mathbb{M}$, and $\mathbb{M} \models \varphi$.

Otherwise, $\psi(\overline{x})$ has the form $\exists y\, \rho(\overline{x}, y)$, and since φ appears infinitely many times in our enumeration of the pithy Π_2 sentences in T', there is some $l > k$ such that $\varphi = \varphi_l$. Then $\pi_l(\overline{a})$ is a tuple of distinct elements of A_l, and $\overline{b} := \pi_{l+1}(\overline{a})$ is a tuple of distinct elements of A_{l+1}. In Step 2 of stage $l+1$ of the construction, we ensured that there was some witness $c_{\overline{b}}$ such that $A_{l+1} \models \rho(\overline{b}, c_{\overline{b}})$. If $c_{\overline{b}}$ is not an element of the tuple $\overline{b}$, then for any $c \in \mathbb{M}$ such that $\pi_{l+1}(c) = c_{\overline{b}}$, we have $\mathbb{M} \models \rho(\overline{a}, c)$. Since $\nu(\pi_{l+1}^{-1}[\{c\}]) = \nu_{l+1}(c) > 0$, the set $\rho(\overline{a}, \mathbb{M})$ has positive ν-measure. On the other hand, if $c_{\overline{b}}$ is an element of the tuple $\overline{b}$, say b_i, then $\mathbb{M} \models \rho(\overline{a}, a_i)$.

(iii) $\mathbb{M}$ is R_χ-rooted with respect to quantifier-free types.

We would like to show that every non-redundant quantifier-free n-type containing $R_\chi(\overline{x})$ that is realized in $\mathbb{M}$ has a root in $\mathbb{M}$. Suppose not. Then there is a quantifier-free type $p(\overline{x})$ and a family of n-tuples $(\overline{a}^i)_{i\in I}$ from $\mathbb{M}$ such that each $\overline{a}^i$ realizes p, but there is no element a which is in every $\overline{a}^i$. Note that if such a family exists, then we can find one containing only finitely many tuples: picking some $\overline{a}$ in the family, for each element a_j in $\overline{a}$ there is another tuple in the family which does not contain a_j, so $n + 1$ tuples suffice.

Let $(\overline{a}^1, \dots, \overline{a}^m)$ be our finite family of tuples. Let k be large enough so that $R_\chi \in L_k$ and π_k is injective on $\bigcup_{i=1}^m \|\overline{a}^i\|$. For all i, let $\overline{b}^i = \pi_k(\overline{a}^i)$. Then all of the tuples $\overline{b}^i$ realize the same quantifier-free L_k-type $p' = p \restriction L_k$ in A_k, and p' contains $R_\chi(\overline{x})$. By Step 4 (c) of stage k of our construction, the tuples $\overline{b}^i$ must actually realize the same quantifier-free L' type $q \supseteq p'$ in M' (which may be distinct from p). But there is no element which appears in all of these tuples, contradicting the fact that M' is R_χ-rooted.

Let $\mu = \mu_{\mathbb{M},\nu}$. By Lemma 6.5, μ is a properly ergodic structure that almost surely satisfies T' and omits the types in Q.

Part 3: Rescaling to obtain continuum-many properly ergodic structures

Again, let n be the number of free variables in $\chi(\overline{x})$. For any quantifier-free formula $\varphi(x_1, \dots, x_n)$, we define the quantifier-free formula $\varphi^*(x_1, \dots, x_n)$:

$$\bigvee_{\sigma\in\mathrm{Sym}(n)} \varphi(x_{\sigma(1)}, \dots, x_{\sigma(n)}).$$

Note that the set $\varphi^*(\mathbb{M}) = \{\overline{a} \in \mathbb{M}^n \mid \mathbb{M} \models \varphi^*(\overline{a})\}$ is invariant under the natural action of the symmetric group $\mathrm{Sym}(n)$ on $\mathbb{M}^n$ by permuting coordinates.

We will use the following claim to apply the rescaling technique from [AFKP17].

Claim: There is some quantifier-free formula $\varphi(\overline{x})$ such that $0 < \nu^n(\varphi^*(\mathbb{M})) < 1$.
Proof of Claim. Suppose not. Then for every quantifier-free formula $\varphi(\overline{x})$, $\nu^n(\varphi^*(\mathbb{M}))$ is equal to 0 or 1. Note that $\nu^n(\varphi^*(\mathbb{M})) > 0$ if and only if $\nu^n(\varphi(\mathbb{M})) > 0$. In particular, since $\nu^n(R_\chi(\mathbb{M})) > 0$, we have $\nu^n(R^*_\chi(\mathbb{M})) = 1$.

Let $A \subseteq \mathbb{M}^n$ be the set of tuples satisfying the partial quantifier-free type $\{\varphi^*(\overline{x}) \mid \nu^n(\varphi^*(\mathbb{M})) = 1\} \cup \{\neg\varphi^*(\overline{x}) \mid \nu^n(\varphi^*(\mathbb{M})) = 0\}$. Since A is a countable intersection of measure 1 sets, it has measure 1. Pick a tuple $\overline{a} \in A$. Some permutation of $\overline{a}$ satisfies R_χ, and A is $\mathrm{Sym}(n)$-invariant, so we may assume that $\mathbb{M} \models R_\chi(\overline{a})$. Let $p(\overline{x}) = \mathrm{qftp}(\overline{a})$. Since $R_\chi(\overline{x}) \in p(\overline{x})$, some coordinate a_i of $\overline{a}$ is a root for $p(\overline{x})$.

Now ν is atomless, so the set of tuples in $\mathbb{M}^n$ containing a_i has measure 0. Thus we can pick another tuple $\overline{b} \in A$ which does not contain a_i. By rootedness, no permutation of $\overline{b}$ satisfies $p(\overline{x})$.

In particular, there is some quantifier-free formula $\psi(\overline{x}) \in p(\overline{x})$ such that no permutation of $\overline{b}$ satisfies $\psi(\overline{x})$ (explicitly, take the conjunction of $n!$ formulas in $p(\overline{x})$, one separating $p(\overline{x})$ from $\mathrm{qftp}(\sigma(\overline{b}))$ for each $\sigma \in \mathrm{Sym}(n)$). We therefore have $\mathbb{M} \models \neg\psi^*(\overline{b})$. But we also have that $\psi^*(\overline{x}) \in p(\overline{x})$ and so $\mathbb{M} \models \psi^*(\overline{a})$. Hence every tuple in A must satisfy $\psi^*(\overline{x})$, contradicting the fact that $\overline{b} \in A$. □

In [AFKP17], a method is described for rescaling a probability measure μ according to a weight $\mathcal{W}$ (essentially an assignment of weights to the pieces of a finite partition of the domain) to obtain a new probability measure $\mu^{\mathcal{W}}$. In that paper, all probability measures are continuous measures on $\mathbb{R}$, but the results apply equally well to measures on $\mathbb{M}$, since this is a standard Borel space.

The main observation about this construction is that μ and $\mu^{\mathcal{W}}$ are equivalent measures, in the sense that they are absolutely continuous with respect to each other. It follows that for our measure ν on $\mathbb{M}$, any measure of the form $\nu^{\mathcal{W}}$ is an atomless probability measure on $\mathbb{M}$, with the property that $(\mathbb{M}, \nu^{\mathcal{W}})$ satisfies T' with strong witnesses, and hence the measure $\mu_{\mathcal{W}} = \mu_{\mathbb{M},\nu^{\mathcal{W}}}$ on $\mathrm{Str}_{L'}$ is a properly ergodic structure which almost surely satisfies T' and omits the types in Q.

Now by the key proposition [AFKP17, Proposition 3.8], since the set $\varphi^*(\mathbb{M})$ from the Claim above is an $\mathrm{Sym}(n)$-invariant Borel set with ν^n-measure strictly between 0 and 1, the expression $\nu^{\mathcal{W}}(\varphi^*(\mathbb{M}))$ takes on continuum-many values as $\mathcal{W}$ varies through the possible weights. And since $\mu_{\mathcal{W}}([\![\varphi^*(\overline{a})]\!]) = \nu^{\mathcal{W}}(\varphi^*(\mathbb{M}))$ for any tuple $\overline{a}$ of distinct elements of $\mathbb{N}$, this construction produces continuum-many properly ergodic structures of the form $\mu_{\mathcal{W}}$. □

Theorem 6.6, along with the results of the previous sections, gives a "measure-free" characterization of those theories which admit properly ergodic models.

Theorem 6.7. *Suppose Σ is a set of sentences in some countable fragment F of $\mathcal{L}_{\omega_1,\omega}$. The following are equivalent:*

(1) There is a properly ergodic structure μ such that $\mu \models \Sigma$.
(2) There are continuum-many properly ergodic structures μ such that $\mu \models \Sigma$.
(3) There is a countable fragment $F' \supseteq F$ of $\mathcal{L}_{\omega_1,\omega}$, a complete F'-theory $T \supseteq \Sigma$ with trivial dcl, a formula $\chi(\overline{x})$ in F', and a model $M \models T$ which is χ-rooted.

Proof. (3) → (2): By Theorem 6.6, there are continuum-many properly ergodic structures μ such that $\mu \models T$, and $\Sigma \subseteq T$.

(2) → (1): Clear.

(1) → (3): Theorem 4.7 gives us a countable fragment $F' \supseteq F$, and a formula $\chi(\overline{x})$ in F' such that $\mu(\chi(\overline{x})) > 0$, but for every F'-type p containing $\chi(\overline{x})$, $\mu(p) = 0$. Let $T = \mathrm{Th}_F(\mu)$. Then $\Sigma \subseteq T$, and T has trivial dcl by Theorem 2.19.

Now by Theorem 5.5, the set of χ-rooted models of T has measure 1. In particular, it is non-empty. □

Remark 6.8. The conditions in Theorem 6.7 (3) can sometimes be satisfied with $F' = F$. In fact, for many of the examples in Section 3, we could take F' to be first-order logic. However, Example 3.6 shows that, in general, the move to a larger fragment of $\mathcal{L}_{\omega_1,\omega}$ is necessary.

The following two corollaries, which may be of interest independently of Theorem 6.7, follow immediately from its proof in the case that μ is properly ergodic and from the analogous construction in [AFP16] in the case that μ is almost surely isomorphic to a countable structure.

Corollary 6.9. *If μ is an ergodic structure, then for any countable fragment F of $\mathcal{L}_{\omega_1,\omega}$, the theory $\mathrm{Th}_F(\mu)$ has a Borel model (of cardinality $2^{\aleph_0}$).*

Corollary 6.10. *For every countable fragment F of $\mathcal{L}_{\omega_1,\omega}$, every ergodic structure μ is F-elementarily equivalent to a random-free ergodic structure μ'. That is, there exists a random-free ergodic structure μ' such that $\mathrm{Th}_F(\mu) = \mathrm{Th}_F(\mu')$.*

Moreover, except in the case that μ concentrates on the isomorphism type of a highly homogeneous structure M, there exist continuum-many such μ'.

For a definition and discussion of high homogeneity, and the previously-mentioned characterization [Cam76] of highly homogeneous structures as those interdefinable with one of the five reducts of the rational linear order, see [AFKP17, §2.3].

Acknowledgments

Portions of this work appear in A.K.'s PhD thesis [Kru16, Chapters 1 and 2], completed under the supervision of Thomas Scanlon, whom we thank for his support and insightful comments. We also thank Will Boney, Alexander Kechris, and Slawomir Solecki for helpful discussions, and Andrew Marks for bringing Example 3.1 to our attention.

This research was facilitated by the Mathematical Sciences Research Institute program on Model Theory, Arithmetic Geometry, and Number Theory (January–May 2014), the workshop at U.C. Berkeley on Vaught's Conjecture (June 2015), and the Lorentz Center workshop on Logic and Random Graphs (August–September 2015).

Work on this publication by C.F. was made possible through the support of ARO grant W911NF-13-1-0212 and a grant from Google.

References

[Ack15] N. Ackerman, *Representations of Aut(M)-invariant measures: Part 1*, arXiv e-print 1509.06170 (2015).

[AFKP17] N. Ackerman, C. Freer, A. Kwiatkowska, and R. Patel, *A classification of orbits admitting a unique invariant measure*, Annals of Pure and Applied Logic **168** (2017), no. 1, 19–36.

[AFNP16] N. Ackerman, C. Freer, J. Nešetřil, and R. Patel, *Invariant measures via inverse limits of finite structures*, Eur. J. Combin. **52** (2016), 248–289.

[AFP16] N. Ackerman, C. Freer, and R. Patel, *Invariant measures concentrated on countable structures*, Forum Math. Sigma **4** (2016), no. e17, 59 pp.

[AFP17] ———, *Countable infinitary theories admitting an invariant measure*, arXiv e-print 1710.06128 (2017).

[Ald81] D. J. Aldous, *Representations for partially exchangeable arrays of random variables*, J. Multivariate Anal. **11** (1981), no. 4, 581–598.

[Aus08] T. Austin, *On exchangeable random variables and the statistics of large graphs and hypergraphs*, Probab. Surv. **5** (2008), 80–145.

[BBG+15] P. Balister, B. Bollobás, K. Gunderson, I. Leader, and M. Walters, *Random geometric graphs and isometries of normed spaces*, arXiv e-print 1504.05324 (2015).

[BJ11] A. Bonato and J. Janssen, *Infinite random geometric graphs*, Ann. Comb. **15** (2011), no. 4, 597–617.

[BM00] M. B. Bekka and M. Mayer, *Ergodic theory and topological dynamics of group actions on homogeneous spaces*, London Math. Soc. Lecture Note Ser., vol. 269, Cambridge Univ. Press, Cambridge, 2000.

[Cam76] P. J. Cameron, *Transitivity of permutation groups on unordered sets*, Math. Z. **148** (1976), no. 2, 127–139.

[Cha68] C. C. Chang, *Some remarks on the model theory of infinitary languages*, The syntax and semantics of infinitary languages (Jon Barwise, ed.), Lecture Notes in Mathematics, vol. 72, Springer-Verlag, Berlin, 1968, pp. 36–63.

[Hoo79] D. N. Hoover, *Relations on probability spaces and arrays of random variables*, Preprint, Institute for Advanced Study, Princeton, NJ (1979).

[Jan13] S. Janson, *Graphons, cut norm and distance, couplings and rearrangements*, New York J. Math. Monographs, vol. 4, State Univ. of New York, Albany, NY, 2013.

[Kal92] O. Kallenberg, *Symmetries on random arrays and set-indexed processes*, J. Theoret. Probab. **5** (1992), no. 4, 727–765.

[Kal99] ———, *Multivariate sampling and the estimation problem for exchangeable arrays*, J. Theoret. Probab. **12** (1999), no. 3, 859–883.

[Kal05] ———, *Probabilistic symmetries and invariance principles*, Probability and its Applications, Springer-Verlag, New York, 2005.

[Kec95] A. S. Kechris, *Classical descriptive set theory*, Grad. Texts Math., vol. 156, Springer-Verlag, New York, 1995.

[KK04] H. J. Keisler and J. F. Knight, *Barwise: infinitary logic and admissible sets*, Bull. Symbolic Logic **10** (2004), no. 1, 4–36.

[Kru16] A. Kruckman, *Infinitary limits of finite structures*, Ph.D. thesis, University of California, Berkeley, 2016.

[Kur76] K. Kuratowski, *A certain theorem on ideals, and its applications to the Baire property of Polish spaces*, Uspehi Mat. Nauk **31** (1976), no. 5 (191), 108–111.

[Lov12] L. Lovász, *Large networks and graph limits*, Amer. Math. Soci. Colloq. Publ., vol. 60, Amer. Math. Soc., Providence, RI, 2012.

[LS10] L. Lovász and B. Szegedy, *Regularity partitions and the topology of graphons*, An irregular mind: Szemerédi is 70, Bolyai Soc. Math. Stud., vol. 21, János Bolyai Math. Soc., Budapest, 2010, pp. 415–446.

[LS12] L. Lovász and B. Szegedy, *Random graphons and a weak Positivstellensatz for graphs*, J. Graph Theory **70** (2012), no. 2, 214–225.

[Mar02] D. Marker, *Model theory: an introduction*, Grad. Texts Math., vol. 217, Springer-Verlag, New York, 2002.

[Mor70] M. Morley, *The number of countable models*, J. Symbolic Logic **35** (1970), 14–18.

[PSN11] O. Pikhurko, B. Szegedy, and J. Nešetřil (editors), *Report on the AIM workshop on Graph and hypergraph limits*, American Institute of Mathematics, August 2011, https://aimath.org/pastworkshops/graphlimitsrep.pdf.

[PV10] F. Petrov and A. Vershik, *Uncountable graphs and invariant measures on the set of universal countable graphs*, Random Struct. Algor. **37** (2010), no. 3, 389–406.

[Tao11] T. Tao, *An introduction to measure theory*, Grad. Stud. Math., vol. 126, Amer. Math. Soc., Providence, RI, 2011.

Department of Mathematics, Harvard University, Cambridge, MA 02138
E-mail address: nate@math.harvard.edu

Remine, Fairfax, VA 22031
E-mail address: cameron@remine.com

Department of Mathematics, Indiana University, Bloomington, IN 47405
E-mail address: akruckma@indiana.edu

Wheaton College, Norton, MA, 02766
E-mail address: patel_rehana@wheatoncollege.edu