INVARIANT MEASURES CONCENTRATED ON COUNTABLE STRUCTURES

NATHANAEL ACKERMAN, CAMERON FREER, AND REHANA PATEL

Abstract. Let L be a countable language. We say that a countable infinite L-structure $\mathcal{M}$ admits an invariant measure when there is a probability measure on the space of L-structures with the same underlying set as $\mathcal{M}$ that is invariant under permutations of that set, and that assigns measure one to the isomorphism class of $\mathcal{M}$. We show that $\mathcal{M}$ admits an invariant measure if and only if it has trivial definable closure, i.e., the pointwise stabilizer in $\mathrm{Aut}(\mathcal{M})$ of an arbitrary finite tuple of $\mathcal{M}$ fixes no additional points. When $\mathcal{M}$ is a Fraïssé limit in a relational language, this amounts to requiring that the age of $\mathcal{M}$ have strong amalgamation. Our results give rise to new instances of structures that admit invariant measures and structures that do not.

2010 *Mathematics Subject Classification.* Primary: 03C98; Secondary: 60G09, 37L40, 05C80, 03C75, 62E10, 05C63. *Keywords:* Invariant measure, exchangeability, infinitary logic, trivial definable closure, strong amalgamation, Scott sentence, graph limits. June 28, 2016

1. Introduction

Randomness is used to construct objects throughout mathematics, and structures resulting from symmetric random constructions often exhibit structural regularities. Here we characterize, in terms of a combinatorial criterion, those countable structures in a countable language that can be built via a random construction that is invariant under reorderings of the elements.

A probabilistic construction is *exchangeable* when its distribution satisfies the symmetry condition of being invariant under permutations of its elements. When an exchangeable construction almost surely produces a single structure (up to isomorphism), we say that the structure *admits an invariant measure*. Such structures often exhibit regularity properties such as universality or ultrahomogeneity. Two of the most important randomly constructed structures with these regularities are Rado's countable universal ultrahomogeneous graph and Urysohn's universal separable ultrahomogeneous metric space. The Rado graph may be generated as a random graph by independently choosing edges according to the Erdős–Rényi construction [ER59], and Urysohn space arises via (the completion of) an exchangeable countable metric space, by a construction of Vershik [Ver02b], [Ver04].

Because these examples have such rich internal structure, it is natural to ask which other objects admit invariant measures. One formulation of this question was posed by Cameron in [Cam90, §4.10]. Petrov and Vershik [PV10] have recently shown, using a new type of construction, that the countable universal ultrahomogeneous K_n-free graphs all admit invariant measures. In the present work, we combine methods from the model theory of infinitary logic with ideas from Petrov and Vershik's construction to give a complete characterization of those countable infinite structures in a countable language that admit invariant measures. Specifically, we show that a structure $\mathcal{M}$ admits an invariant measure if and only if the pointwise stabilizer in $\mathrm{Aut}(\mathcal{M})$ of any finite set of elements of $\mathcal{M}$ fixes no additional elements, a condition known as having *trivial definable closure*.

Many natural examples of objects admitting invariant measures are generic, in the sense of being Fraïssé limits, i.e., the countable universal ultrahomogeneous object for some class of finite structures [Hod93, §7.1]. One may ask what additional regularity properties must hold of Fraïssé limits that admit invariant measures. Fraïssé limits arise from amalgamation procedures for "gluing together" finite substructures. Our result implies that a Fraïssé limit in a countable relational language admits an invariant measure if and only if it has *strong amalgamation*, a natural restriction on the gluing procedure that produces the limit.

Our characterization gives rise to new instances of structures that admit invariant measures, and structures that do not. We apply our results to existing classifications of ultrahomogeneous graphs, directed graphs, and partial orders, as well as other combinatorial structures, thereby providing several new examples of exchangeable constructions that lead to generic structures. Among those structures for which we provide the first such constructions are the countable universal ultrahomogeneous partial order [Sch79] and certain countable universal graphs forbidding a finite homomorphism-closed set of finite connected graphs [CSS99]. Structures for which our results imply the non-existence of such constructions include Hall's countable universal group [Hod93, §7.1, Example 1], and the existentially complete countable universal bowtie-free graph [Kom99].

1.1. **Background.**

The Rado graph $\mathscr{R}$, sometimes known as the "random graph", is (up to isomorphism) the unique countable universal ultrahomogeneous graph [Rad64]. It is the Fraïssé limit of the class of finite graphs, with a first-order theory characterized by so-called "extension axioms" that have a simple syntactic form. It is also the classic example of a countable structure that has a symmetric probabilistic construction, namely, the countably infinite version of the Erdős–Rényi random graph process introduced by Gilbert [Gil59] and Erdős and Rényi [ER59]. For $0 < p < 1$, this process determines a random graph on a countably infinite set of vertices by independently flipping a coin of weight p for every pair of distinct vertices, and adding an edge between those vertices precisely when the coin flip comes up heads. Denote this random variable by $\mathbb{G}(\mathbb{N}, p)$. The random graph $\mathbb{G}(\mathbb{N}, p)$ is almost surely isomorphic to $\mathscr{R}$, for any p such that $0 < p < 1$. Moreover, each $\mathbb{G}(\mathbb{N}, p)$ is *exchangeable*, i.e., its distribution is invariant under arbitrary permutations of the vertices, and so there are continuum-many different invariant measures concentrated on $\mathscr{R}$ (up to isomorphism). It is natural to ask which other structures admit random constructions that are invariant in this way.

Consider the Henson graph $\mathscr{H}_3$, the unique (up to isomorphism) countable universal ultrahomogeneous triangle-free graph [Hen71]. Like the Rado graph, it has a first-order theory consisting of extension axioms, and can be constructed as a Fraïssé limit. Does it also admit an invariant measure? In contrast with $\mathscr{R}$, no countable random graph whose distribution of edges is independent and identically distributed (i.i.d.) can be almost surely isomorphic to $\mathscr{H}_3$. But this does not rule out the possibility of an exchangeable random graph almost surely isomorphic to $\mathscr{H}_3$. Its distribution would constitute a measure on countable graphs, invariant under arbitrary permutations of the underlying vertex set, that is concentrated on the isomorphism class of $\mathscr{H}_3$.

One might consider building an invariant measure concentrated on (graphs isomorphic to) $\mathscr{H}_3$ by "approximation from below" using uniform measures on finite triangle-free graphs, by analogy with the invariant measure concentrated on $\mathscr{R}$ obtained as the weak limit of uniform measures on finite graphs. The distribution of finite Erdős–Rényi random graphs $\mathbb{G}(n, \frac{1}{2})$ is simply the uniform measure on graphs with n labeled vertices; the sequence $\mathbb{G}(n, \frac{1}{2})$ converges in distribution to $\mathbb{G}(\mathbb{N}, \frac{1}{2})$, which is almost surely isomorphic to $\mathscr{R}$. So to obtain an invariant measure concentrated on $\mathscr{H}_3$, one might similarly consider the weak limit of the sequence of uniform measures on finite triangle-free labeled graphs of size n, i.e., of the distributions of the random graphs $\mathbb{G}(n, \frac{1}{2})$ conditioned on being triangle-free. However, by work of Erdős, Kleitman, and Rothschild [EKR76] and Kolaitis, Prömel, and Rothschild [KPR87], this sequence is asymptotically almost surely bipartite, and so its weak limit is almost surely not isomorphic to $\mathscr{H}_3$. Hence, as noted in [PV10], this particular approach does not produce an invariant measure concentrated on $\mathscr{H}_3$.

Petrov and Vershik [PV10] provided the first instance of an invariant measure concentrated on the Henson graph $\mathscr{H}_3$ (up to isomorphism); they also did likewise for Henson's other countable universal ultrahomogeneous K_n-free graphs, where $n > 3$. They produced this measure via a "top down" construction, building a continuum-sized triangle-free graph in such a way that an i.i.d. sample from its vertex set induces an exchangeable random graph that is almost surely isomorphic to $\mathscr{H}_3$.

In this paper, we address the question of invariant measures concentrated on *arbitrary* structures. Given a countable language L and a countable infinite L-structure $\mathcal{M}$, we ask whether there exists a probability measure on the space of L-structures with the same underlying set as $\mathcal{M}$, invariant under arbitrary permutations of that set, assigning measure one to the isomorphism class of $\mathcal{M}$. We provide a complete answer to this question, by characterizing such L-structures $\mathcal{M}$ as precisely those that have trivial group-theoretic definable closure, i.e., those structures $\mathcal{M}$ for which the pointwise stabilizer in $\mathrm{Aut}(\mathcal{M})$ of any finite tuple $\mathbf{a}$ from $\mathcal{M}$ fixes no elements of $\mathcal{M}$ except those in $\mathbf{a}$. We use infinitary logic to establish a setting in which, whenever $\mathcal{M}$ has trivial definable closure, we can construct continuum-sized objects that upon sampling produce invariant measures concentrated on $\mathcal{M}$. When $\mathcal{M}$ does not have trivial definable closure, we show that such invariant measures cannot exist.

Our results build on several ideas from [PV10]. In particular, Petrov and Vershik show that if a continuum-sized graph satisfies certain properties, then sampling from it produces an invariant measure concentrated on $\mathscr{H}_3$ (and similar results for $\mathscr{R}$ and the other Henson K_n-free graphs); they then proceed to construct such continuum-sized graphs. We identify a certain type of continuum-sized structure whose existence guarantees, using a similar sampling procedure, an invariant measure concentrated on a target countable structure; we then construct such a continuum-sized structure whenever the target structure has trivial definable closure.

Underlying Petrov and Vershik's construction of invariant measures, as well as ours, is the characterization of countable exchangeable (hyper)graphs as those obtained via certain sampling procedures from continuum-sized structures. These ideas were developed by Aldous [Ald81], Hoover [Hoo79], Kallenberg [Kal92] and Vershik [Ver02a] in work on the probability theory of exchangeable arrays. More recently, similar machinery has come to prominence in the combinatorial theory of limits of dense graphs via *graphons*, due to Lovász and Szegedy [LS06] and others. For an equivalence between these characterizations, see Austin [Aus08a] and Diaconis and Janson [DJ08]. The *standard recipe* of [Aus08a] provides a more general formulation of the correspondence between sampling procedures on continuum-sized objects and arbitrary countable exchangeable structures.

In the present paper, we are interested primarily in determining those countable infinite structures for which there exists at least one invariant measure concentrated on its isomorphism class. In the case of countable graphs, our construction in fact provides a new method for building graphons. In particular, the graphons we build are *random-free*, in the sense of [Jan13, §10]. Therefore our construction shows that whenever there is an invariant measure concentrated on the isomorphism class of a countable graph, there is such a measure that comes from sampling a random-free graphon.

Within mathematical logic, the study of invariant measures on countable first-order structures goes back to work of Gaifman [Gai64], Scott and Krauss [SK66], and Krauss [Kra69]. For a discussion of this earlier history and its relationship to Hoover's work on exchangeability, see Austin [Aus08a, §3.8 and §4.3]. Our countable relational setting is akin to that explored more recently in extremal combinatorics by Razborov [Raz07]; for details see [Aus08a, §4.3] and [Aus08b].

Other work in model theory has examined aspects of probabilistic constructions. Droste and Kuske [DK03] and Dolinka and Mašulović [DM12] describe probabilistic constructions of countable infinite structures, though without invariance. Usvyatsov [Usv08] has also considered a relationship between invariant measures and notions of genericity in the setting of continuous first-order logic, especially with respect to Urysohn space.

1.2. Main results.

Our main theorem characterizes countable infinite structures $\mathcal{M}$ that admit invariant measures as those for which the pointwise stabilizer, in $\mathrm{Aut}(\mathcal{M})$, of an arbitrary finite tuple of $\mathcal{M}$ fixes no additional points. For a countable language L, let Str_L be the Borel measure space of L-structures with underlying set $\mathbb{N}$. (This is a standard space on which to consider measures invariant under the action of the permutation group S_∞; we provide details in §2.3.) Then we have the following result.

Theorem 1.1. *Let L be a countable language, and let $\mathcal{M}$ be a countable infinite L-structure. The following are equivalent:*

1. *There is a probability measure on Str_L, invariant under the natural action of S_∞ on Str_L, that is concentrated on the set of elements of Str_L that are isomorphic to $\mathcal{M}$.*
2. *The structure $\mathcal{M}$ has trivial group-theoretic definable closure, i.e., for every finite tuple $\mathbf{a} \in \mathcal{M}$, we have $\mathrm{dcl}_{\mathcal{M}}(\mathbf{a}) = \mathbf{a}$, where $\mathrm{dcl}_{\mathcal{M}}(\mathbf{a})$ is the collection of elements of $\mathcal{M}$ that are fixed by all automorphisms of $\mathcal{M}$ fixing $\mathbf{a}$ pointwise.*

Note that every finite structure admits a natural probability measure that is invariant under permutations of the underlying set. But also every finite structure has nontrivial definable closure, and so the statement of this theorem does not extend to finite structures.

Our main result is the equivalence of (1) and (2); but further, an observation of Kechris and Marks shows the additional equivalence with (3) in Theorem 1.2 below.

For any structure $\mathcal{N} \in \mathrm{Str}_L$, we write $\mathrm{Aut}(\mathcal{N})$ for its automorphism group considered as a subgroup of S_∞, and take the action of $\mathrm{Aut}(\mathcal{N})$ on Str_L to be that given by the restriction of the natural action of S_∞.

Theorem 1.2. *Properties* (1) *and* (2) *from Theorem 1.1 are also equivalent to the following:*

3. *There is some $\mathcal{N} \in \mathrm{Str}_L$ that has trivial group-theoretic definable closure and is such that there is an $\mathrm{Aut}(\mathcal{N})$-invariant probability measure on Str_L concentrated on the set of elements of Str_L that are isomorphic to $\mathcal{M}$.*

Note that (3) is ostensibly weaker than (1), as in general an $\mathrm{Aut}(\mathcal{N})$-invariant measure need not be S_∞-invariant.

A structure $\mathcal{M}$ is said to be *ultrahomogeneous* when every partial isomorphism between finitely generated substructures of $\mathcal{M}$ extends to an automorphism of $\mathcal{M}$. Define the *age* of a countable L-structure $\mathcal{M}$ to be the class of all finitely generated L-structures that are isomorphic to a substructure of $\mathcal{M}$. The age of any countable infinite ultrahomogeneous L-structure has the so-called *amalgamation property*, which stipulates that any two structures in the age can be "glued together" over any common substructure, preserving this substructure but possibly identifying other elements. Countable infinite ultrahomogeneous L-structures have been characterized by Fraïssé as those obtained from their ages via a canonical "back-and-forth" construction using amalgamation; they are often called Fraïssé limits and are axiomatized by Π_2 "extension axioms". (For details, see [Hod93, Theorems 7.1.4, 7.1.7].)

A standard result [Hod93, Theorem 7.1.8] (see also [Cam90, §2.7]) states that when $\mathcal{M}$ is a countable infinite ultrahomogeneous structure in a countable relational language, $\mathcal{M}$ has trivial definable closure precisely when its age satisfies the more stringent condition known as the *strong amalgamation property*, which requires that no elements (outside the intersection) are identified during the amalgamation. Note that in [Hod93, §7.1], strong amalgamation is shown to be equivalent to a property known as "no algebraicity", which is equivalent to our notion of (group-theoretic) trivial definable closure for structures in a language with only relation symbols (but not for structures in a language with constant or function symbols). Thus we obtain the following corollary to Theorem 1.1.

Corollary 1.3. *Let L be a countable relational language, and let $\mathcal{M}$ be a countable infinite L-structure. Suppose $\mathcal{M}$ is ultrahomogeneous. The following are equivalent:*

(1) *There is a probability measure on Str_L, invariant under the natural action of S_∞ on Str_L, that is concentrated on the set of elements of Str_L that are isomorphic to $\mathcal{M}$.*

(2′) *The age of $\mathcal{M}$ satisfies the strong amalgamation property.*

At the Workshop on Homogeneous Structures, held at the University of Leeds in 2011, Anatoly Vershik asked whether an analogue of the notion of a continuum-sized *topologically universal graph* [PV10] exists for an arbitrary Fraïssé limit. We propose our notion of a (continuum-sized) *Borel L-structure strongly witnessing a theory*, defined in Section 3, as an appropriate analogue.

Our results then show that such a Borel L-structure can exist for a Fraïssé limit precisely when its age has the strong amalgamation property. If the age of a Fraïssé limit $\mathcal{M}$ in a countable relational language L has the strong amalgamation property, then the proof of our main result involves building a Borel L-structure that, just like a topologically universal graph, has a "large" set of witnesses for every (nontrivial) extension axiom. On the other hand, when the age of $\mathcal{M}$ does not have the strong amalgamation property, such a Borel L-structure cannot exist; according to the machinery of our proof, it would necessarily induce an invariant measure concentrated on $\mathcal{M}$, violating Corollary 1.3.

1.3. Outline of the paper.

We begin, in Section 2, by describing our setting and providing preliminaries. Throughout this paper we work in a countable language L. We first describe the infinitary language $\mathcal{L}_{\omega_1,\omega}(L)$. In particular, we recall the notion of a *Scott sentence*, a single infinitary sentence in $\mathcal{L}_{\omega_1,\omega}(L)$ that describes a countable structure up to isomorphism (among countable structures). We then define a certain kind of infinitary Π_2 sentence, which we call *pithy* Π_2, and which can be thought of as a "one-point" extension axiom. We go on to describe the measure space Str_L and define the natural action of S_∞ on Str_L, called the *logic action*. Using these notions, we explain what is meant by an *invariant measure* and what it means for a measure to be *concentrated* on a set of structures. We then recall the group-theoretic notion of *definable closure* and its connection to the model theory of $\mathcal{L}_{\omega_1,\omega}(L)$. Next, for any given countable L-structure $\mathcal{M}$, we describe its *canonical language* $L_{\overline{\mathcal{M}}}$ and *canonical structure* $\overline{\mathcal{M}}$; the latter is essentially equivalent to $\mathcal{M}$, but is characterized (among countable structures) by a theory $T_{\overline{\mathcal{M}}}$ consisting entirely of "one-point" extension axioms. We show that $\mathcal{M}$ admits an invariant measure if and only if $\overline{\mathcal{M}}$ does, and that $\mathcal{M}$ has trivial definable closure if and only if $\overline{\mathcal{M}}$ does. Finally, we review some basic conventions from probability theory.

In Section 3, we prove the existence of an invariant measure concentrated on an L-structure $\mathcal{M}$ that has trivial definable closure. We do so by constructing an invariant measure concentrated on its canonical structure $\overline{\mathcal{M}}$.

The invariant measures that we build in Section 3 each come from sampling a continuum-sized structure. Our method uses a similar framework to that employed by Petrov and Vershik in [PV10] for graphs. The first-order theory of the Henson graph $\mathscr{H}_3$ is generated by a set of Π_2 axioms that characterize it up to isomorphism among countable graphs. Petrov and Vershik construct an invariant measure concentrated on $\mathscr{H}_3$ by building a continuum-sized structure that realizes a "large" set of witnesses for each of these axioms.

In our generalization of their construction, we build a continuum-sized structure that satisfies $T_{\overline{\mathcal{M}}}$ in a particularly strong way, analogously to [PV10]. Specifically, given a Π_2 sentence of the form $(\forall \mathbf{x})(\exists y)\psi(\mathbf{x}, y)$ in $T_{\overline{\mathcal{M}}}$, we ensure that for every tuple $\mathbf{a}$ in the structure, the sentence $(\exists y)\psi(\mathbf{a}, y)$ has a "large" set of witnesses, whenever $\psi(\mathbf{a}, b)$ does not hold for any $b \in \mathbf{a}$. The construction proceeds inductively by defining quantifier-free types on intervals, interleaving successive refinements of existing intervals with enlargements by new intervals that provide the "large" sets of witnesses. This is possible by virtue of $T_{\overline{\mathcal{M}}}$ having a property we call *duplication of quantifier-free types*, which occurs precisely when $\overline{\mathcal{M}}$ has trivial definable closure. The continuum-sized structure built in this way is such that a random countable structure induced by sampling from it, with respect to an appropriate measure, will be a model of $T_{\overline{\mathcal{M}}}$ almost surely, thereby producing an invariant measure concentrated on $\overline{\mathcal{M}}$.

Section 4 provides the converse, for an arbitrary countable language L: If a countable infinite L-structure has nontrivial definable closure, it cannot admit an invariant measure. This is a direct argument that does not require the machinery developed in Section 3. In fact we present a generalization of the converse, due to Kechris and Marks, which states that such an L-structure cannot even admit an $\mathrm{Aut}(\mathcal{N})$-invariant measure for any $\mathcal{N} \in \mathrm{Str}_L$ having trivial definable closure.

In Section 5 we apply Theorem 1.1 and Corollary 1.3 to obtain examples of countable infinite structures that admit invariant measures, and those that do not. We describe how any structure can be "blown up" into one that admits an invariant measure and also into one that does not. This allows us to give examples of countable structures having arbitrary Scott rank that admit invariant measures, and examples of those that do not admit invariant measures. We then analyze definable closure in well-known countable structures to determine whether or not they admit invariant measures.

We conclude, in Section 6, with several connections to the theory of graph limits, and additional applications of our results.

2. Preliminaries

Throughout this paper we use uppercase letters to represent sets, lowercase letters to represent elements of a set and lowercase boldface letters to represent finite tuples (of variables, or of elements of a structure). The length $|\mathbf{x}|$ of a tuple of variables $\mathbf{x}$ is the number of entries, not the number of distinct variables, in the tuple, and likewise for tuples of elements. We use the notation $(x_1, \ldots, x_k)$ and $x_1 \cdots x_k$ interchangeably to denote a tuple of variables $\mathbf{x}$ of length k that has entries $x_1, \ldots, x_k$, in that order, and similarly for tuples of elements. When it enhances clarity, we write, e.g., $(\mathbf{x}, \mathbf{y})$ for (x_1, x_2, y_1, y_2), when $\mathbf{x} = x_1x_2$ and $\mathbf{y} = y_1y_2$. For an n-tuple $\mathbf{a} \in A^n$, we frequently abuse notation and write $\mathbf{a} \in A$.

2.1. **Infinitary logic.**

We begin by reviewing some basic definitions from logic. A *language* L, also called a *signature*, is a set $L := \mathcal{R} \cup \mathcal{C} \cup \mathcal{F}$, where $\mathcal{R}$ is a set of *relation symbols*, $\mathcal{C}$ is a set of *constant symbols*, and $\mathcal{F}$ is a set of *function symbols*, all disjoint. For each relation symbol $R \in \mathcal{R}$ and function symbol $f \in \mathcal{F}$, fix an associated positive integer, called its *arity*. We take the *equality symbol*, written $=$, to be a logical symbol, not a binary relation symbol in L. In this paper, all languages are countable. Given a language L, an L*-structure* $\mathcal{M}$ is a non-empty set M endowed with interpretations of the symbols in L. We sometimes write $x \in \mathcal{M}$ in place of $x \in M$.

We now describe the class $\mathcal{L}_{\omega_1,\omega}(L)$ of infinitary formulas in the language L. For more on infinitary logic and Scott sentences, see [Kei71], [Bar75], or [Mar02, §2.4]. For the basics of first-order languages, terms, formulas, and theories, see [Mar02, §1.1].

Definition 2.1. The class $\mathcal{L}_{\omega_1,\omega}(L)$ is the smallest collection of formulas that contains all atomic formulas of L; the formulas $(\exists x)\psi(x)$ and $\neg\chi$, where $\psi(x), \chi \in \mathcal{L}_{\omega_1,\omega}(L)$; and the formula $\bigwedge_{i\in I} \varphi_i$, where I is an arbitrary countable set, $\varphi_i \in \mathcal{L}_{\omega_1,\omega}(L)$ for each $i \in I$, and the set of free variables of $\bigwedge_{i\in I} \varphi_i$ is finite.

A formula of $\mathcal{L}_{\omega_1,\omega}(L)$ may have countably infinitely many variables, but only finitely many that are free. Note that the more familiar $\mathcal{L}_{\omega,\omega}(L)$, consisting of first-order formulas, is the restriction of $\mathcal{L}_{\omega_1,\omega}(L)$ where conjunctions are over *finite* index sets I. As is standard, we will freely use the abbreviations $\forall := \neg\exists\neg$ and $\bigvee := \neg\bigwedge\neg$, as well as binary $\wedge$ and $\vee$, in formulas of $\mathcal{L}_{\omega_1,\omega}(L)$. We will sometimes refer to formulas of $\mathcal{L}_{\omega_1,\omega}(L)$ as L*-formulas*.

A sentence is a formula with no free variables. A *(countable) theory* of $\mathcal{L}_{\omega_1,\omega}(L)$ is an arbitrary (countable) collection of sentences in $\mathcal{L}_{\omega_1,\omega}(L)$. Note that a theory need not be deductively closed.

For a formula $\varphi(x_1, \ldots, x_n)$ of $\mathcal{L}_{\omega_1,\omega}(L)$ whose free variables are among $x_1, \ldots, x_n$, all distinct, the notation $\mathcal{M} \models \varphi(a_1, \ldots, a_n)$ means that $\varphi(x_1, \ldots, x_n)$ holds in $\mathcal{M}$ when $x_1, \ldots, x_n$ are instantiated, respectively, by the elements $a_1, \ldots, a_n$ of the underlying set M. For a theory T, we write $\mathcal{M} \models T$ to mean that $\mathcal{M} \models \varphi$ for every sentence $\varphi \in T$; in this case, we say that $\mathcal{M}$ is a *model* of T. We write $T \models \varphi$ to mean that the sentence φ is true in every L-structure that is a model of T. As a special case, we write $\models \varphi$ to mean $\emptyset \models \varphi$, i.e., the sentence φ is true in every L-structure. When $\psi(\mathbf{x})$ is a formula with free variables among the entries of the finite tuple $\mathbf{x}$, we write $\models \psi(\mathbf{x})$ to mean $\models (\forall\mathbf{x})\psi(\mathbf{x})$.

A key model-theoretic property of $\mathcal{L}_{\omega_1,\omega}(L)$ is that any countable L-structure can be characterized up to isomorphism, among countable L-structures, by a single sentence of $\mathcal{L}_{\omega_1,\omega}(L)$.

Proposition 2.2 (see [Bar75, Corollary VII.6.9] or [Mar02, Theorem 2.4.15])**.** *Let L be a countable language, and let $\mathcal{M}$ be a countable L-structure. There is a sentence*

$\sigma_{\mathcal{M}} \in \mathcal{L}_{\omega_1,\omega}(L)$, *called the (canonical)* **Scott sentence** *of* $\mathcal{M}$, *such that for every countable* L*-structure* $\mathcal{N}$, *we have* $\mathcal{N} \models \varphi$ *if and only if* $\mathcal{N} \cong \mathcal{M}$.

2.2. Pithy Π_2 theories.

Countable theories consisting of "extension axioms" will play a crucial role in our main construction in Section 3. In fact, we will work with a notion of "one-point extension axioms", which allows us to realize witnesses for all possible finite configurations, one element at a time. In a sense that we make precise in §2.5, an arbitrary countable structure is essentially equivalent to one (in a different language) admitting an axiomatization consisting only of one-point extension axioms.

Definition 2.3. A sentence in $\mathcal{L}_{\omega_1,\omega}(L)$ is Π_2 when it is of the form $(\forall \mathbf{x})(\exists \mathbf{y})\psi(\mathbf{x}, \mathbf{y})$, where the (possibly empty) tuple $\mathbf{xy}$ consists of distinct variables, and $\psi(\mathbf{x}, \mathbf{y})$ is quantifier-free. A countable theory T of $\mathcal{L}_{\omega_1,\omega}(L)$ is Π_2 when every sentence $\varphi \in T$ is Π_2.

In our main construction, it will be convenient to work with a restricted kind of extension axiom, which we call *pithy*.

Definition 2.4. A Π_2 sentence $(\forall \mathbf{x})(\exists \mathbf{y})\psi(\mathbf{x}, \mathbf{y}) \in \mathcal{L}_{\omega_1,\omega}(L)$, where $\psi(\mathbf{x}, \mathbf{y})$ is quantifier-free, is said to be **pithy** when the tuple $\mathbf{y}$ consists of precisely one variable. A countable Π_2 theory T of $\mathcal{L}_{\omega_1,\omega}(L)$ is said to be pithy when every sentence in T is pithy. Note that we allow the degenerate case where $\mathbf{x}$ is the empty tuple and the Π_2 sentence is of the form $(\exists y)\psi(y)$.

Note that a pithy Π_2 sentence can be written uniquely in the form $(\forall \mathbf{x})(\exists y)\psi(\mathbf{x}, y)$, where ψ is quantifier-free, and where the free variables of ψ are among the entries of $\mathbf{x}y$.

2.3. The logic action on the measurable space Str_L.

Let L be an arbitrary countable language. Define Str_L to be the set of L-structures that have underlying set $\mathbb{N}$. For every formula $\varphi(x_1, \ldots, x_j) \in \mathcal{L}_{\omega_1,\omega}(L)$ and $n_1, \ldots, n_j \in \mathbb{N}$, where j is the number of free variables of φ, define

$$[\![\varphi(n_1, \ldots, n_j)]\!] := \{\mathcal{M} \in \mathrm{Str}_L : \mathcal{M} \models \varphi(n_1, \ldots, n_j)\}.$$

The set Str_L becomes a measurable space when it is equipped with the Borel σ-algebra generated by subbasic open sets of the form:

$$[\![R(n_1, \ldots, n_j)]\!]$$

where $R \in L$ is a j-ary relation symbol and $n_1, \ldots, n_j \in \mathbb{N}$;

$$[\![c = n]\!]$$

where $c \in L$ is a constant symbol and $n \in \mathbb{N}$; and

$$[\![f(n_1, \ldots, n_k) = n_{k+1}]\!]$$

where $f \in L$ is a k-ary function symbol and $n_1, \ldots, n_{k+1} \in \mathbb{N}$.

For any sentence φ of $\mathcal{L}_{\omega_1,\omega}(L)$, the set $[\![\varphi]\!]$ is Borel, by [Kec95, Proposition 16.7]. Given a countable L-structure $\mathcal{M}$, recall that the Scott sentence $\sigma_{\mathcal{M}} \in \mathcal{L}_{\omega_1,\omega}(L)$ determines $\mathcal{M}$ up to isomorphism among countable structures. Therefore $[\![\sigma_{\mathcal{M}}]\!] = \{\mathcal{N} \in \mathrm{Str}_L : \mathcal{N} \cong \mathcal{M}\}$, the isomorphism class of $\mathcal{M}$ in Str_L, is Borel.

Denote by S_∞ the permutation group of $\mathbb{N}$. There is a natural group action, called the *logic action*, of S_∞ on Str_L, induced by permutation of the underlying set $\mathbb{N}$; for more details, see [Kec95, §16.C]. Note that the orbit of an L-structure $\mathcal{M} \in \mathrm{Str}_L$ under this action is the isomorphism class of $\mathcal{M}$ in Str_L. We call a (Borel) measure μ on Str_L **invariant** when it is invariant under the logic action, i.e., for every Borel set $X \subseteq \mathrm{Str}_L$ and every $g \in S_\infty$, we have $\mu(X) = \mu(g \cdot X)$. Given a subgroup G of S_∞, written $G \leq S_\infty$, a (Borel) measure μ is *G-invariant* when it is invariant under the restriction of the logic action to G.

Let μ be a probability measure on Str_L. We say that μ is **concentrated** on a Borel set $X \subseteq \mathrm{Str}_L$ when $\mu(X) = 1$. We are interested in structures up to isomorphism, and for a countable infinite L-structure $\mathcal{M}$, we say that μ *is concentrated on* $\mathcal{M}$ when μ is concentrated on the isomorphism class of $\mathcal{M}$ in Str_L. We say that $\mathcal{M}$ **admits an invariant measure** when such an invariant measure μ exists. Note that when we say that μ is concentrated on some class of structures, we mean that μ is concentrated on the restriction of that class to Str_L.

2.4. Definable closure.

Our characterization of structures admitting invariant measures is in terms of the group-theoretic notion of *definable closure*.

An *automorphism* of an L-structure $\mathcal{M}$ is a bijection $g\colon M \to M$ such that

$$R^{\mathcal{M}}\big(g(a_1), \dots, g(a_j)\big) \qquad \text{if and only if} \qquad R^{\mathcal{M}}(a_1, \dots, a_j)$$

for every relation symbol $R \in L$ of arity j and all elements $a_1, \dots, a_j \in \mathcal{M}$,

$$g(c^{\mathcal{M}}) = c^{\mathcal{M}}$$

for every constant symbol $c \in L$, and

$$f^{\mathcal{M}}\big(g(b_1), \dots, g(b_k)\big) = g\big(f^{\mathcal{M}}(b_1, \dots, b_k)\big)$$

for every function symbol $f \in L$ of arity k and elements $b_1, \dots, b_k \in \mathcal{M}$.

We write $\mathrm{Aut}(\mathcal{M})$ to denote the group of automorphisms of $\mathcal{M}$.

Definition 2.5. Let $\mathcal{M}$ be an L-structure, and let $\mathbf{a} \in \mathcal{M}$. The **definable closure** of $\mathbf{a}$ in $\mathcal{M}$, denoted $\mathrm{dcl}_{\mathcal{M}}(\mathbf{a})$, is the collection of $b \in \mathcal{M}$ that are fixed by all automorphisms of $\mathcal{M}$ fixing $\mathbf{a}$ pointwise, i.e., the set of $b \in \mathcal{M}$ for which the set

$$\big\{g(b) : g \in \mathrm{Aut}(\mathcal{M}) \text{ s.t. } (\forall a \in \mathbf{a})\ g(a) = a\big\}$$

is a singleton, namely $\{b\}$.

This notion is sometimes known as the *group-theoretic definable closure*. For countable structures, it has the following equivalent formulation in terms of the formulas of $\mathcal{L}_{\omega_1,\omega}(L)$ that use parameters from the tuple $\mathbf{a}$.

Given an L-structure $\mathcal{M}$ and a tuple $\mathbf{a} \in \mathcal{M}$, let $L_{\mathbf{a}}$ denote the language L expanded by a new constant symbol for each element of $\mathbf{a}$. Then let $\mathcal{M}_{\mathbf{a}}$ denote the $L_{\mathbf{a}}$-structure given by $\mathcal{M}$ with the entries of $\mathbf{a}$ named by their respective constant symbols in $L_{\mathbf{a}}$.

Lemma 2.6 (see [Hod93, Lemma 4.1.3])**.** *Let L be a countable language, and let $\mathcal{M}$ be a countable L-structure with $\mathbf{a} \in \mathcal{M}$. Then $b \in \mathrm{dcl}_{\mathcal{M}}(\mathbf{a})$ if and only if there is a formula $\varphi \in \mathcal{L}_{\omega_1,\omega}(L_{\mathbf{a}})$ with one free variable, whose unique realization in $\mathcal{M}_{\mathbf{a}}$ is b, i.e.,*

$$\mathcal{M}_{\mathbf{a}} \models \varphi(b) \wedge [(\forall x, y)(\varphi(x) \wedge \varphi(y)) \rightarrow x = y].$$

When the first-order theory of $\mathcal{M}$ is $\aleph_0$-categorical, it suffices to consider only first-order formulas $\varphi \in \mathcal{L}_{\omega,\omega}$ in Lemma 2.6 (see [Hod93, Corollary 7.3.4]); in this case, group-theoretic definable closure coincides with the standard notion of *model-theoretic definable closure*.

Definition 2.7. We say that an L-structure $\mathcal{M}$ has **trivial definable closure** when the definable closure of every tuple $\mathbf{a} \in \mathcal{M}$ is trivial, i.e., $\mathrm{dcl}_{\mathcal{M}}(\mathbf{a}) = \mathbf{a}$ for all $\mathbf{a} \in \mathcal{M}$.

Note that if $\mathcal{M}$ has trivial definable closure, then L cannot have constant symbols, and every function of $\mathcal{M}$ is a *choice function* (or *selector*), i.e., for every function symbol $f \in L$ and every $\mathbf{a} \in \mathcal{M}$, we have $f^{\mathcal{M}}(\mathbf{a}) \in \mathbf{a}$.

It is sometimes more convenient to work with (group-theoretic) *algebraic closure*.

Definition 2.8. Let $\mathcal{M}$ be an L-structure, and let $\mathbf{a} \in \mathcal{M}$. The **algebraic closure** of $\mathbf{a}$ in $\mathcal{M}$, denoted $\mathrm{acl}_{\mathcal{M}}(\mathbf{a})$, is the collection of $b \in \mathcal{M}$ whose orbit under those automorphisms of $\mathcal{M}$ fixing $\mathbf{a}$ pointwise is finite. In other words, $\mathrm{acl}_{\mathcal{M}}(\mathbf{a})$ is the set of $b \in \mathcal{M}$ for which the set

$$\big\{g(b) : g \in \mathrm{Aut}(\mathcal{M}) \text{ s.t. } (\forall a \in \mathbf{a})\ g(a) = a\big\}$$

is finite. We say that $\mathcal{M}$ has **trivial algebraic closure** when the algebraic closure of every tuple $\mathbf{a} \in \mathcal{M}$ is trivial, i.e., $\mathrm{acl}_{\mathcal{M}}(\mathbf{a}) = \mathbf{a}$ for all $\mathbf{a} \in \mathcal{M}$.

Note that an L-structure has trivial algebraic closure if and only if it has trivial definable closure. This fact will be useful in Section 5 when we find examples of structures admitting invariant measures.

2.5. The canonical language and structure.

We now define the *canonical language* $L_{\overline{\mathcal{A}}}$ and *canonical structure* $\overline{\mathcal{A}}$ associated to each structure $\mathcal{A} \in \mathrm{Str}_L$. We will see that the canonical structure admits an invariant measure precisely when the original structure does, and has trivial definable closure precisely when the original does. In the proof of our main theorem, this will enable us to work in the setting of canonical structures. We will also establish below that canonical structures admit pithy Π_2 axiomatizations, a fact which we will use in our main construction in Section 3. For more details on canonical languages and structures, see, e.g., [BK96, §1.5].

Definition 2.9. Let $\mathcal{A} \in \mathrm{Str}_L$. For each $k \in \mathbb{N}$ let $\sim_k$ be the equivalence relation on $\mathbb{N}^k$ given by

$$\mathbf{x} \sim_k \mathbf{y} \qquad \text{if and only if} \qquad (\exists g \in \mathrm{Aut}(\mathcal{A}))\ \ g(\mathbf{x}) = \mathbf{y}.$$

Define the **canonical language for** $\mathcal{A}$ to be the (countable) relational language $L_{\overline{\mathcal{A}}}$ that consists of, for each $k \in \mathbb{N}$ and $\sim_k$-equivalence class E of $\mathcal{A}$, a k-ary relation symbol R_E. Then define the **canonical structure for** $\mathcal{A}$ to be the structure $\overline{\mathcal{A}} \in \mathrm{Str}_{L_{\overline{\mathcal{A}}}}$ in which, for each $\sim_k$-equivalence class E of $\mathcal{A}$, the interpretation of R_E is the corresponding orbit $E \subseteq \mathbb{N}^k$.

By the definition of $L_{\overline{\mathcal{A}}}$, the $\mathrm{Aut}(\mathcal{A})$-orbits of tuples in $\mathcal{A}$ are $\mathcal{L}_{\omega_1,\omega}(L_{\overline{\mathcal{A}}})$-definable in $\overline{\mathcal{A}}$. In fact, as we will see in Lemma 2.13, these orbits are already $\mathcal{L}_{\omega_1,\omega}(L)$-definable in $\mathcal{A}$. We begin by noting the folklore result that the canonical structure $\overline{\mathcal{A}}$ has elimination of quantifiers.

Lemma 2.10. *Let $\mathcal{A} \in \mathrm{Str}_L$, and consider its canonical structure $\overline{\mathcal{A}}$. Then for all $k \in \mathbb{N}$, every $\mathrm{Aut}(\mathcal{A})$-invariant subset of $\mathbb{N}^k$ is the set of realizations of some quantifier-free formula $\psi(\mathbf{x}) \in \mathcal{L}_{\omega_1,\omega}(L_{\overline{\mathcal{A}}})$. In particular, for every formula $\varphi(\mathbf{x}) \in \mathcal{L}_{\omega_1,\omega}(L_{\overline{\mathcal{A}}})$, as the set of its realizations is $\mathrm{Aut}(\mathcal{A})$-invariant, there is a quantifier-free formula $\psi(\mathbf{x}) \in \mathcal{L}_{\omega_1,\omega}(L_{\overline{\mathcal{A}}})$ such that*

$$\overline{\mathcal{A}} \models \varphi(\mathbf{x}) \leftrightarrow \psi(\mathbf{x})$$

holds.

The following definition of *($\mathcal{L}_{\omega_1,\omega}$-)interdefinability* extends that of the standard notion of interdefinability from the setting of $\aleph_0$-categorical theories (see, e.g., [AZ86, §1]). In particular, two structures are interdefinable when they have the same underlying set (not necessarily countable) and the same $\mathcal{L}_{\omega_1,\omega}$-definable sets.

Let L_0 and L_1 be countable languages. Let $\mathcal{N}_0$ be an L_0-structure and $\mathcal{N}_1$ an L_1-structure having the same underlying set (not necessarily countable).

Definition 2.11. An **$\mathcal{L}_{\omega_1,\omega}$-interdefinition** (or simply, *interdefinition*) between $\mathcal{N}_0$ and $\mathcal{N}_1$ is a pair (Ψ_0, Ψ_1) of maps

$$\begin{aligned} \Psi_0 &: \mathcal{L}_{\omega_1,\omega}(L_0) \to \mathcal{L}_{\omega_1,\omega}(L_1) \qquad \text{and} \\ \Psi_1 &: \mathcal{L}_{\omega_1,\omega}(L_1) \to \mathcal{L}_{\omega_1,\omega}(L_0) \end{aligned}$$

satisfying, for $j \in \{0, 1\}$,

$$\begin{aligned} \mathcal{N}_{1-j} &\models \Psi_j \circ \Psi_{1-j}(\eta) \leftrightarrow \eta, \\ \mathcal{N}_{1-j} &\models \neg\Psi_j(\chi) \leftrightarrow \Psi_j(\neg\chi), \\ \mathcal{N}_{1-j} &\models \bigwedge_{i\in I} \Psi_j(\varphi_i) \leftrightarrow \Psi_j\big(\bigwedge_{i\in I} \varphi_i\big), \qquad \text{and} \\ \mathcal{N}_{1-j} &\models (\exists x)\Psi_j(\psi(x)) \leftrightarrow \Psi_j\big((\exists x)\psi(x)\big), \end{aligned}$$

where $\eta \in \mathcal{L}_{\omega_1,\omega}(L_{1-j})$ and χ, $\psi(x) \in \mathcal{L}_{\omega_1,\omega}(L_j)$, where I is an arbitrary countable set and each $\varphi_i \in \mathcal{L}_{\omega_1,\omega}(L_j)$, and such that the free variables of $\Psi_j(\xi)$ are the same as those of ξ for every $\xi \in \mathcal{L}_{\omega_1,\omega}(L_j)$.

We say that $\mathcal{N}_0$ and $\mathcal{N}_1$ are **interdefinable via** (Ψ_0, Ψ_1) when (Ψ_0, Ψ_1) is an interdefinition between $\mathcal{N}_0$ and $\mathcal{N}_1$ such that for every $k \in \mathbb{N}$ and every formula $\psi(\mathbf{x}) \in \mathcal{L}_{\omega_1,\omega}(L_0)$ with k free variables, we have

$$\{\mathbf{m} \in \mathbb{N}^k : \mathcal{N}_0 \models \psi(\mathbf{m})\} = \{\mathbf{m} \in \mathbb{N}^k : \mathcal{N}_1 \models \Psi_0(\psi)(\mathbf{m})\}.$$

We say that $\mathcal{N}_0$ and $\mathcal{N}_1$ are **interdefinable** when they are interdefinable via some interdefinition.

Note that $\mathcal{N}_0$ and $\mathcal{N}_1$ are interdefinable precisely when, for every $k \in \mathbb{N}$, a set $X \subseteq \mathbb{N}^k$ is definable in $\mathcal{N}_0$ (without parameters) by an $\mathcal{L}_{\omega_1,\omega}(L_0)$-formula if and only if it is definable in $\mathcal{N}_1$ (without parameters) by an $\mathcal{L}_{\omega_1,\omega}(L_1)$-formula.

Lemma 2.12. *Suppose (Ψ_0, Ψ_1) is an interdefinition between $\mathcal{N}_0$ and $\mathcal{N}_1$, and let $\mathcal{N}_0'$ be an L_0-structure (not necessarily countable) that satisfies the same $\mathcal{L}_{\omega_1,\omega}(L_0)$-theory as $\mathcal{N}_0$. Then there is a unique L_1-structure $\mathcal{N}_1'$ such that $\mathcal{N}_0'$ and $\mathcal{N}_1'$ are interdefinable via (Ψ_0, Ψ_1). In particular, $\mathcal{N}_1'$ satisfies the same $\mathcal{L}_{\omega_1,\omega}(L_1)$-theory as $\mathcal{N}_1$.*

Proof. Let $\mathcal{N}_1'$ be the unique structure with the same underlying set as $\mathcal{N}_0'$ such that for any atomic L_1-formula ψ, the set of realizations of ψ in $\mathcal{N}_1'$ is precisely the set of realizations of $\Psi_1(\psi)$ in $\mathcal{N}_0'$. Because $\mathcal{N}_0'$ satisfies the same sentences of $\mathcal{L}_{\omega_1,\omega}(L_0)$ as $\mathcal{N}_0$, by considering Definition 2.11 one can see that (Ψ_0, Ψ_1) is an interdefinition between $\mathcal{N}_0'$ and $\mathcal{N}_1'$; further, $\mathcal{N}_1'$ is the only such L_1-structure. It is immediate that $\mathcal{N}_1'$ satisfies the same sentences of $\mathcal{L}_{\omega_1,\omega}(L_1)$ as $\mathcal{N}_1$. □

We will use the following folklore result in the proof of our main theorem.

Lemma 2.13. *Let $\mathcal{A} \in \mathrm{Str}_L$ and let $\overline{\mathcal{A}}$ be its canonical structure. Then $\mathcal{A}$ is interdefinable with $\overline{\mathcal{A}}$.*

In fact, one can show that for $\mathcal{A} \in \mathrm{Str}_{L_0}$ and $\mathcal{B} \in \mathrm{Str}_{L_1}$, the structures $\mathcal{A}$ and $\mathcal{B}$ are interdefinable if and only if $L_{\overline{\mathcal{A}}} = L_{\overline{\mathcal{B}}}$ and $\overline{\mathcal{A}} = \overline{\mathcal{B}}$. Along with Lemma 2.13, this implies that a structure in Str_L is characterized up to interdefinability by its canonical structure.

As an immediate corollary of Lemmas 2.12 and 2.13, we see that given $\mathcal{A} \in \mathrm{Str}_L$ and an arbitrary L-structure $\mathcal{N}$ having the same $\mathcal{L}_{\omega_1,\omega}(L)$-theory as $\mathcal{A}$, there is a unique $L_{\overline{\mathcal{A}}}$-structure with which $\mathcal{N}$ is interdefinable via the interdefinition given in Lemma 2.13 between $\mathcal{A}$ and $\overline{\mathcal{A}}$. When $\mathcal{N} \in \mathrm{Str}_L$, this $L_{\overline{\mathcal{A}}}$-structure is $\overline{\mathcal{N}}$, the canonical structure of $\mathcal{N}$; in Corollary 3.22, we will call the analogous $L_{\overline{\mathcal{A}}}$-structure $\overline{\mathcal{N}}$ even when $\mathcal{N}$ is uncountable.

We now show that interdefinability preserves whether or not a countable structure admits an invariant measure and also whether or not it has trivial definable closure.

Lemma 2.14. *Suppose $\mathcal{A} \in \mathrm{Str}_{L_0}$ and $\mathcal{B} \in \mathrm{Str}_{L_1}$ are interdefinable. Then $\mathcal{A}$ admits an invariant measure if and only if $\mathcal{B}$ does.*

Proof. Let (Ψ_0, Ψ_1) be an interdefinition between $\mathcal{A}$ and $\mathcal{B}$. We first define a Borel map

$$\iota\colon \{\mathcal{C} \in \mathrm{Str}_{L_0} \;:\; \mathcal{C} \cong \mathcal{A}\} \to \{\mathcal{D} \in \mathrm{Str}_{L_1} \;:\; \mathcal{D} \cong \mathcal{B}\}$$

that commutes with the logic action. For every $\mathcal{C} \in \mathrm{Str}_{L_0}$ isomorphic to $\mathcal{A}$, let $\iota(\mathcal{C})$ be the L_1-structure, given by Lemma 2.12, that has the same $\mathcal{L}_{\omega_1,\omega}(L_1)$-theory as $\mathcal{B}$. Since $\iota(\mathcal{C})$ and $\mathcal{B}$ are countable, they are in fact isomorphic. Further, since (Ψ_0, Ψ_1) is an interdefinition between $\mathcal{C}$ and $\iota(\mathcal{C})$, it follows that ι is a bijection. Recall that the σ-algebra of Str_{L_1} is generated by sets of the form $[\![\varphi(n_1, \dots, n_j)]\!]$, for $\varphi \in \mathcal{L}_{\omega_1,\omega}(L_1)$ and $n_1, \dots, n_j \in \mathbb{N}$, where j is the number of free variables in φ. By the definition of ι, we have

$$\iota^{-1}\big([\![\varphi(n_1, \dots, n_j)]\!]\big) = \big[\!\big[\Psi_1(\varphi)(n_1, \dots, n_j)\big]\!\big],$$

which is a Borel set in Str_{L_0}. Hence the map ι is Borel.

Observe that for every $g \in S_\infty$ and $\mathcal{C} \in \mathrm{Str}_{L_0}$ isomorphic to $\mathcal{A}$, we have $\iota(g \cdot \mathcal{C}) = g \cdot \iota(\mathcal{C})$, which is interdefinable with $g \cdot \mathcal{C}$. Hence for every invariant probability measure μ concentrated on $\mathcal{A}$, its pushforward along ι is an invariant probability measure concentrated on $\mathcal{B}$. By symmetry, $\mathcal{A}$ admits an invariant measure if and only if $\mathcal{B}$ does. □

In particular, taking $\mathcal{B} = \overline{\mathcal{A}}$, we see that a countable structure admits an invariant measure if and only if its canonical structure does.

Lemma 2.15. *Suppose $\mathcal{A} \in \mathrm{Str}_{L_0}$ and $\mathcal{B} \in \mathrm{Str}_{L_1}$ are interdefinable. Then $\mathcal{A}$ has trivial definable closure if and only if $\mathcal{B}$ does.*

Proof. Suppose $\mathcal{A}$ does not have trivial definable closure. Let $\mathbf{a}, b \in \mathbb{N}$ with $b \notin \mathbf{a}$ and $b \in \operatorname{dcl}(\mathbf{a})$. By Lemma 2.6, there is a formula $\varphi \in \mathcal{L}_{\omega_1,\omega}((L_0)_{\mathbf{a}})$ whose unique realization in $\mathcal{A}_{\mathbf{a}}$ is b. Note that $\mathcal{A}_{\mathbf{a}}$ and $\mathcal{B}_{\mathbf{a}}$ are interdefinable. Hence there is a corresponding $\mathcal{L}_{\omega_1,\omega}((L_1)_{\mathbf{a}})$-formula whose unique realization in $\mathcal{B}_{\mathbf{a}}$ is b, witnessing the non-trivial definable closure of $\mathbf{a}$ in $\mathcal{B}$. Therefore $\mathcal{B}$ does not have trivial definable closure either. The result follows by symmetry. ☐

By Lemmas 2.14 and 2.15, for countable structures, the properties of having trivial definable closure, and of admitting an invariant measure, are determined up to interdefinability. Further, by Lemma 2.13, each of these properties holds of a countable structure if and only if the respective property holds of its canonical structure, and hence is determined completely by its automorphism group.

Finally, we show that for every countable structure, there is a pithy Π_2 theory in its canonical language that characterizes its canonical structure up to isomorphism among countable structures. From this, it will follow that the canonical structure is *ultrahomogeneous*.

Definition 2.16. We say that an L-structure $\mathcal{M}$ is **ultrahomogeneous** if any isomorphism between two finitely generated substructures of $\mathcal{M}$ extends to an automorphism of $\mathcal{M}$.

Proposition 2.17. *Let $\mathcal{A} \in \mathrm{Str}_L$. There is a countable $\mathcal{L}_{\omega_1,\omega}(L_{\overline{\mathcal{A}}})$-theory, every sentence of which is pithy Π_2, and all of whose countable models are isomorphic to the canonical structure $\overline{\mathcal{A}}$.*

Proof. Consider the $\mathcal{L}_{\omega_1,\omega}(L_{\overline{\mathcal{A}}})$-theory consisting of the following pithy Π_2 axioms, for each $k \in \mathbb{N}$:

- $(\forall \mathbf{x})\big(R_E(\mathbf{x}) \leftrightarrow \bigwedge\{\neg R_G(\mathbf{x}) : R_G \neq R_E \text{ is a } k\text{-ary relation symbol in } L_{\overline{\mathcal{A}}}\}\big)$,
- $(\forall \mathbf{x}) \bigvee\{R_G(\mathbf{x}) : R_G \text{ is a } k\text{-ary relation symbol in } L_{\overline{\mathcal{A}}}\}$, and
- $(\forall \mathbf{x})\big(R_E(\mathbf{x}) \to (\exists y)\, R_F(\mathbf{x}, y)\big)$,

where $|\mathbf{x}| = k$, and R_E and R_F are, respectively, k- and $(k+1)$-ary relation symbols in $L_{\overline{\mathcal{A}}}$ such that

$$\overline{\mathcal{A}} \models (\forall \mathbf{x} y)\big(R_F(\mathbf{x}, y) \to R_E(\mathbf{x})\big).$$

It is immediate that $\overline{\mathcal{A}}$ satisfies this theory. Furthermore, for any two countable models of the theory, the first two axioms require that every k-tuple in either model realizes exactly one k-ary relation. Hence given two k-tuples $\mathbf{a}, \mathbf{b}$ of $\overline{\mathcal{A}}$ satisfying the same relation, we may use the third axiom to construct an automorphism of $\overline{\mathcal{A}}$ mapping $\mathbf{a}$ to $\mathbf{b}$, by a standard back-and-forth argument. This establishes that the theory has one countable model up to isomorphism. ☐

Note that the above argument further shows the standard result that $\overline{\mathcal{A}}$ is ultrahomogeneous. The pithy Π_2 theory of Proposition 2.17 can therefore be thought of as an infinitary analogue of a Fraïssé theory. In particular, as with Fraïssé theories in first-order relational languages, the age of $\overline{\mathcal{A}}$ has strong amalgamation precisely when $\overline{\mathcal{A}}$ has trivial definable closure. (For more details on Fraïssé theories, see [Hod93, §7.1].) Therefore, even if $\mathcal{A}$ is not ultrahomogeneous itself, Corollary 1.3 could be applied to a structure that is essentially equivalent to $\mathcal{A}$, namely the canonical structure $\overline{\mathcal{A}}$. Indeed, by Lemma 2.15, $\overline{\mathcal{A}}$ has strong amalgamation precisely when $\mathcal{A}$ has trivial definable closure.

2.6. Basic probability notions.

Throughout this paper, we make use of conventions from measure-theoretic probability theory to talk about random structures having certain almost-sure properties. For a general reference on probability theory, see, e.g., [Kal02].

Let $(\Omega, \mathcal{G}, \mathbb{P})$ be a probability space, and suppose $(H, \mathcal{H})$ is a measurable space. Recall that an *H-valued random variable* Z is a $(\mathcal{G}, \mathcal{H})$-measurable function $Z \colon \Omega \to H$. Such a function Z is also sometimes called a *random element in* H. The *distribution* of Z is defined to be the probability measure $\mathbb{P} \circ Z^{-1}$.

Given a property $E \in \mathcal{H}$, we say that E holds of Z *almost surely*, abbreviated *a.s.*, when $\mathbb{P}\big(Z^{-1}(E)\big) = 1$. Sometimes, in this situation, we say instead that E holds of Z *with probability one*. For example, given a random element Z in Str_L and a Borel set $[\![\varphi]\!]$, where φ is a sentence of $\mathcal{L}_{\omega_1,\omega}(L)$, we say that $[\![\varphi]\!]$ holds of Z a.s. when $\mathbb{P}\big(Z^{-1}([\![\varphi]\!])\big) = \mathbb{P}\big(\{w \in \Omega : Z(w) \models \varphi\}\big) = 1$. In fact, we will typically not make the property explicit, and will, for instance, write that the random structure $Z \models \varphi$ a.s. when $\mathbb{P}\big(\{w \in \Omega : Z(w) \models \varphi\}\big) = 1$; this probability is abbreviated as $\mathbb{P}\{Z \models \varphi\}$.

In the proof of our main theorem, when we show that a measure μ on Str_L is concentrated on the set of models in Str_L of some sentence φ, we will do so by demonstrating that, with probability one, $Z \models \varphi$, where Z is a random structure with distribution μ.

A sequence of (H-valued) random variables is said to be *independent and identically distributed*, abbreviated *i.i.d.*, when each random variable has the same distribution and the random variables are mutually independent. When this distribution is m, we say that the sequence is *m-i.i.d.*

3. Existence of invariant measures

We now show the existence of invariant measures concentrated on a countable infinite structure having trivial definable closure. We prove this in Theorem 3.21, which constitutes one direction of Theorem 1.1, the main result of this paper.

The following is an outline of our proof; in the presentation below we will, however, develop the machinery in the reverse order. Let $\mathcal{M} \in \mathrm{Str}_L$ be a countable infinite L-structure having trivial definable closure, and $L_{\overline{\mathcal{M}}}$ its canonical language and $\overline{\mathcal{M}}$ its canonical structure, as in §2.5. Let $T_{\overline{\mathcal{M}}}$ be a countable pithy Π_2 theory of $\mathcal{L}_{\omega_1,\omega}(L_{\overline{\mathcal{M}}})$ all of whose countable models are isomorphic to $\overline{\mathcal{M}}$, as in Proposition 2.17. We show, in §3.5, that such a theory $T_{\overline{\mathcal{M}}}$ has a property that we call *duplication of quantifier-free types*. In §§3.3–3.4, we use this property to build a *Borel $L_{\overline{\mathcal{M}}}$-structure* $\mathcal{P}$ that *strongly witnesses* $T_{\overline{\mathcal{M}}}$. Roughly speaking, this means that $\mathcal{P}$ is an $L_{\overline{\mathcal{M}}}$-structure with underlying set $\mathbb{R}$, whose relations are Borel, such that for every pithy Π_2 sentence $(\forall \mathbf{x})(\exists y)\varphi(\mathbf{x}, y) \in T_{\overline{\mathcal{M}}}$ and tuple $\mathbf{a} \in \mathcal{P}$ of the appropriate length, either there is a "large" set of elements $b \in \mathcal{P}$ such that $\varphi(\mathbf{a}, b)$ holds, or else there is some $b \in \mathbf{a}$ such that $\varphi(\mathbf{a}, b)$ holds; in either case, $(\exists y)\varphi(\mathbf{a}, y)$ is "witnessed". In §§3.1–3.2 we show how to use a Borel $L_{\overline{\mathcal{M}}}$-structure that strongly witnesses $T_{\overline{\mathcal{M}}}$ to produce an invariant measure concentrated on the set of models of $T_{\overline{\mathcal{M}}}$ that are in $\mathrm{Str}_{L_{\overline{\mathcal{M}}}}$. By the initial choice of $T_{\overline{\mathcal{M}}}$, this invariant measure on $\mathrm{Str}_{L_{\overline{\mathcal{M}}}}$ is concentrated on $\overline{\mathcal{M}}$. By results of §2.5, we obtain an invariant measure on Str_L concentrated on $\mathcal{M}$.

3.1. Sampling from Borel L-structures.

We begin by introducing a certain kind of L-structure with underlying set $\mathbb{R}$, whose relations and functions are Borel (with respect to the standard topology on $\mathbb{R}$). Our definition is motivated by Petrov and Vershik's notion of a *Borel graph* [PV10, Definition 1]. The model theory of such Borel structures has earlier been studied by Harvey Friedman (published in [Ste85]). For a survey, including more recent work, see [MN13, §1].

Definition 3.1. Let $\mathcal{P}$ be an L-structure whose underlying set is $\mathbb{R}$. We say that $\mathcal{P}$ is a **Borel L-structure** if for all relation symbols $R \in L$, the set $\{\mathbf{a} \in \mathcal{P}^j : R^{\mathcal{P}}(\mathbf{a})\}$ is a Borel subset of $\mathbb{R}^j$, where j is the arity of R; and for all function symbols $f \in L$, the function $f^{\mathcal{P}} \colon \mathcal{P}^k \to \mathcal{P}$ is Borel (equivalently, the graph of $f^{\mathcal{P}}$ is Borel), where k is the arity of f.

Note that although structures with underlying set $\mathbb{R}$ will suffice for our purposes, we could have defined the notion of a *Borel L-structure* more generally, for other measure spaces.

Our first goal in §3.1 is to define a sampling procedure that, given a Borel L-structure with certain properties, yields an invariant measure on Str_L. We begin with several definitions.

Given an L-structure $\mathcal{N}$ of arbitrary cardinality, we write $\mathrm{Clo}(\mathcal{N})$ to denote the set of those sequences in $\mathcal{N}^\omega$ that contain all constants of $\mathcal{N}$ and are closed under the application of functions of $\mathcal{N}$. Such a sequence is precisely an enumeration (possibly with repetition) of the underlying set of some countable substructure of $\mathcal{N}$. Note that whenever L is relational, or when $\mathcal{N}$ has trivial definable closure, we have $\mathrm{Clo}(\mathcal{N}) = \mathcal{N}^\omega$ (but not conversely). We say that $\mathcal{N}$ is **samplable** when $\mathrm{Clo}(\mathcal{N}) = \mathcal{N}^\omega$. Observe that $\mathcal{N}$ is samplable precisely when L has no constant symbols and every function is a choice function.

Next we describe a map taking an element of $\mathrm{Clo}(\mathcal{N})$ to an L-structure with underlying set $\mathbb{N}$. In the case when $\mathcal{N}$ is samplable, we will apply this map to a random sequence of elements of $\mathcal{N}$ to induce a random L-structure with underlying set $\mathbb{N}$.

Definition 3.2. Suppose $\mathcal{N}$ is an L-structure (of arbitrary cardinality). Define the function $\mathcal{F}_{\mathcal{N}} \colon \mathrm{Clo}(\mathcal{N}) \to \mathrm{Str}_L$ as follows. For $\mathbf{A} = (a_i)_{i\in\omega} \in \mathrm{Clo}(\mathcal{N})$, let $\mathcal{F}_{\mathcal{N}}(\mathbf{A})$ be the L-structure with underlying set $\mathbb{N}$ satisfying

$$\mathcal{F}_{\mathcal{N}}(\mathbf{A}) \models R(n_1, \ldots, n_j) \qquad \text{if and only if} \qquad \mathcal{N} \models R(a_{n_1}, \ldots, a_{n_j})$$

for every relation symbol $R \in L$ and for all $n_1, \ldots, n_j \in \mathbb{N}$, where j is the arity of R; satisfying

$$\mathcal{F}_{\mathcal{N}}(\mathbf{A}) \models (c = n) \qquad \text{if and only if} \qquad \mathcal{N} \models (c = a_n)$$

for every constant symbol $c \in L$ and for all $n \in \mathbb{N}$; satisfying

$$\mathcal{F}_{\mathcal{N}}(\mathbf{A}) \models f(n_1, \ldots, n_k) = n_{k+1} \qquad \text{if and only if} \qquad \mathcal{N} \models f(a_{n_1}, \ldots, a_{n_k}) = a_{n_{k+1}}$$

for every function symbol $f \in L$ and for all $n_1, \ldots, n_{k+1} \in \mathbb{N}$, where k is the arity of f; and for which equality is inherited from $\mathbb{N}$, i.e.,

$$\mathcal{F}_{\mathcal{N}}(\mathbf{A}) \models (m \neq n)$$

just when m and n are distinct natural numbers.

When the sequence $\mathbf{A} \in \mathrm{Clo}(\mathcal{N})$ has no repeated entries, $\mathcal{F}_{\mathcal{N}}(\mathbf{A}) \in \mathrm{Str}_L$ is isomorphic to a countable infinite substructure of $\mathcal{N}$. In fact, this will hold a.s. for the random L-structures that we construct in §3.2.

Recall the definition of the Borel σ-algebra on Str_L in §2.3. Define a **subbasic formula** of $\mathcal{L}_{\omega_1,\omega}(L)$ to be a formula of the form $x_1 = x_2$, $R(x_1, \dots, x_j)$, $c = x_1$, or $f(x_1, \dots, x_k) = x_{k+1}$, where $R \in L$ is a relation symbol and j its arity, $c \in L$ is a constant symbol, $f \in L$ is a function symbol and k its arity, and the x_i are distinct variables.

Lemma 3.3. *Let $\mathcal{P}$ be a Borel L-structure. Then $\mathcal{F}_{\mathcal{P}}$ is a Borel measurable function.*

Proof. It suffices to show that the preimages of subbasic open sets of Str_L are Borel. Let ζ be a subbasic formula of $\mathcal{L}_{\omega_1,\omega}(L)$ with j free variables, and let $n_1, \dots, n_j \in \mathbb{N}$. We wish to show that $\mathcal{F}_{\mathcal{P}}^{-1}\Big([\![\zeta(n_1, \dots, n_j)]\!]\Big)$ is Borel.

Let $\pi_{n_1,\dots,n_j} : \mathcal{P}^\omega \to \mathcal{P}^j$ be the projection map defined by

$$\pi_{n_1,\dots,n_j}\big((a_i)_{i\in\omega}\big) = (a_{n_1}, \dots, a_{n_j});$$

this map is Borel. Then

$$\mathcal{F}_{\mathcal{P}}^{-1}\Big([\![\zeta(n_1, \dots, n_j)]\!]\Big) = \pi^{-1}_{n_1,\dots,n_j}\big(\{\mathbf{a} \in \mathcal{P}^j : \mathcal{P} \models \zeta(\mathbf{a})\}\big),$$

as both sides of the equation are equal to

$$\big\{(a_i)_{i\in\omega} \in \mathcal{P}^\omega : \mathcal{P} \models \zeta(a_{n_1}, \dots, a_{n_j})\big\}.$$

By Definition 3.1 we have that $\{\mathbf{a} \in \mathcal{P}^j : \mathcal{P} \models \zeta(\mathbf{a})\}$ is Borel. Hence $\mathcal{F}_{\mathcal{P}}^{-1}\Big([\![\zeta(n_1, \dots, n_j)]\!]\Big)$ is also Borel, as desired. □

We now show how to induce an invariant measure on Str_L from a samplable Borel L-structure $\mathcal{P}$. Suppose m is a probability measure on $\mathbb{R}$. Denote by m^∞ the corresponding product measure on $\mathbb{R}^\omega$, i.e., the distribution of a sequence of independent samples from m. Note that m^∞ is invariant under arbitrary reordering of the indices. We will obtain an invariant measure on Str_L by taking the distribution of the *random* structure with underlying set $\mathbb{N}$ corresponding to an m-i.i.d. sequence of elements of $\mathcal{P}$.

This technique for constructing invariant measures by sampling a continuum-sized structure was used by Petrov and Vershik [PV10], and the following notation and results parallel those in [PV10, §2.3]. A similar method of sampling is used in [LS06, §2.6] to produce the countable random graphs known as *W-random graphs* from continuum-sized *graphons*; for more details on the relationship between these notions and our construction, see §6.1.

Definition 3.4. Let $\mathcal{P}$ be a Borel L-structure, and let m be a probability measure on $\mathbb{R}$. Define the measure $\mu_{(\mathcal{P},m)}$ on Str_L to be

$$\mu_{(\mathcal{P},m)} := m^\infty \circ \mathcal{F}_{\mathcal{P}}^{-1}.$$

When $\mathcal{P}$ is samplable, $m^\infty\big(\mathcal{F}_{\mathcal{P}}^{-1}(\mathrm{Str}_L)\big) = 1$, and so $\mu_{(\mathcal{P},m)}$ is a probability measure, namely the distribution of a random element in Str_L induced via $\mathcal{F}_{\mathcal{P}}$ by an m-i.i.d. sequence on $\mathbb{R}$.

The following lemma makes precise the sense in which the invariance of m^∞ (under the action of S_∞ on $\mathbb{R}^\omega$) yields the invariance of $\mu_{(\mathcal{P},m)}$ (under the logic action).

Lemma 3.5. *Let $\mathcal{P}$ be a Borel L-structure, and let m be a probability measure on $\mathbb{R}$. Then the measure $\mu_{(\mathcal{P},m)}$ is invariant under the logic action.*

Proof. It suffices to verify that $\mu_{(\mathcal{P},m)}$ is invariant on a π-system (i.e., a family of sets closed under finite intersections) that generates the Borel σ-algebra on Str_L, by [Wil91, Lemma 1.6.b]. We first show that $\mu_{(\mathcal{P},m)}$ is invariant on subbasic open sets determined by subbasic formulas of $\mathcal{L}_{\omega_1,\omega}(L)$ along with tuples instantiating them. We then show its invariance for the π-system consisting of sets determined by finite conjunctions of such subbasic formulas.

Let ζ be a subbasic formula of $\mathcal{L}_{\omega_1,\omega}(L)$ with j free variables, and let $n_1, \dots, n_j \in \mathbb{N}$. Consider the set $[\![\zeta(n_1, \dots, n_j)]\!]$, and let $g \in S_\infty$. Note that

$$[\![\zeta\big(g(n_1), \dots, g(n_j)\big)]\!] = \big\{g \cdot \mathcal{N} : \mathcal{N} \in [\![\zeta(n_1, \dots, n_j)]\!]\big\},$$

where $\cdot$ denotes the logic action of S_∞ on Str_L. We will show that

$$\mu_{(\mathcal{P},m)}\Big([\![\zeta\big(g(n_1), \dots, g(n_j)\big)]\!]\Big) = \mu_{(\mathcal{P},m)}\Big([\![\zeta(n_1, \dots, n_j)]\!]\Big). \qquad (\star)$$

We have

$$\mathcal{F}_{\mathcal{P}}^{-1}\Big([\![\zeta\big(g(n_1), \dots, g(n_j)\big)]\!]\Big) = \pi^{-1}_{g(n_1),\dots,g(n_j)}\big(\{\mathbf{a} \in \mathcal{P}^j : \mathcal{P} \models \zeta(\mathbf{a})\}\big)$$

and

$$\mathcal{F}_{\mathcal{P}}^{-1}\Big([\![\zeta(n_1, \dots, n_j)]\!]\Big) = \pi^{-1}_{n_1,\dots,n_j}\big(\{\mathbf{a} \in \mathcal{P}^j : \mathcal{P} \models \zeta(\mathbf{a})\}\big).$$

Because m^∞ is invariant under the action of S_∞ on $\mathbb{R}^\omega$ (given by permuting coordinates of $\mathbb{R}^\omega$), the Borel subsets

$$\mathcal{F}_{\mathcal{P}}^{-1}\Big([\![\zeta\big(g(n_1), \dots, g(n_j)\big)]\!]\Big)$$

and

$$\mathcal{F}_{\mathcal{P}}^{-1}\Big([\![\zeta(n_1, \dots, n_j)]\!]\Big)$$

of $\mathbb{R}^\omega$ have equal m^∞-measure, and so ($\star$) holds.

Now consider subbasic formulas $\zeta_1, \dots, \zeta_k$ of $\mathcal{L}_{\omega_1,\omega}(L)$ with $j_1, \dots, j_k$ free variables, respectively, and let $j' := \sum_{i=1}^k j_k$. Denote the conjunction of these formulas on non-overlapping variables by

$$\varphi(x_1, \dots, x_{j'}) := \zeta_1(x_1, \dots, x_{j_1}) \wedge \cdots \wedge \zeta_k(x_{j'-j_k+1}, \dots, x_{j'}),$$

where $x_1, \dots, x_{j'}$ are distinct variables. We have

$$[\![\varphi(n_1, \dots, n_{j'})]\!] = [\![\zeta_1(n_1, \dots, n_{j_1})]\!] \cap \cdots \cap [\![\zeta_k(n_{j'-j_k+1}, \dots, n_{j'})]\!],$$

for all $n_1, \dots, n_{j'} \in \mathbb{N}$ (not necessarily distinct), and from this we see that

$$[\![\varphi\big(g(n_1), \dots, g(n_{j'})\big)]\!] = [\![\zeta_1\big(g(n_1), \dots, g(n_{j_1})\big)]\!] \cap \cdots \cap [\![\zeta_k\big(g(n_{j'-j_k+1}), \dots, g(n_{j'})\big)]\!].$$

Hence

$$\mathcal{F}_{\mathcal{P}}^{-1}\Big([\![\varphi\big(g(n_1), \dots, g(n_{j'})\big)]\!]\Big)$$

and

$$\mathcal{F}_{\mathcal{P}}^{-1}\Big([\![\varphi(n_1, \dots, n_{j'})]\!]\Big)$$

have equal measure under m^∞. Hence $\mu_{(\mathcal{P},m)}$ is invariant under the logic action. □

Recall that a measure on $\mathbb{R}$ is said to be **continuous** (or *nonatomic*) if it assigns measure zero to every singleton. When m is continuous, samples from $\mu_{(\mathcal{P},m)}$ are a.s. isomorphic to substructures of $\mathcal{P}$.

Lemma 3.6. *Let $\mathcal{P}$ be a samplable Borel L-structure and let m be a continuous probability measure on $\mathbb{R}$. Then $\mu_{(\mathcal{P},m)}$ is a probability measure on Str_L that is concentrated on the union of isomorphism classes of countable infinite substructures of $\mathcal{P}$.*

Proof. As noted before, because $\mathcal{P}$ is samplable, $m^\infty\big(\mathcal{F}_{\mathcal{P}}^{-1}(\mathrm{Str}_L)\big) = 1$, and so $\mu_{(\mathcal{P},m)}$ is a probability measure. Let $\mathbf{A} = (a_i)_{i\in\omega}$ be an m-i.i.d. sequence of $\mathbb{R}$. Note that the induced countable structure $\mathcal{F}_{\mathcal{P}}(\mathbf{A})$ is now a *random L-structure*, i.e., a Str_L-valued random variable, whose distribution is $\mu_{(\mathcal{P},m)}$. Because m is continuous, and since for any $k \neq \ell$ the random variables a_k and a_ℓ are independent, the sequence $\mathbf{A}$ has no repeated entries a.s. Hence $\mathcal{F}_{\mathcal{P}}(\mathbf{A})$ is a.s. isomorphic to a countable infinite (induced) substructure of $\mathcal{P}$. □

3.2. Strongly witnessing a pithy Π_2 theory.

We have seen how to construct invariant measures on Str_L by sampling from a samplable Borel L-structure. We now describe a property that will give us sufficient control over such measures to ensure that they are concentrated on the set of models, in Str_L, of a given countable pithy Π_2 theory T of $\mathcal{L}_{\omega_1,\omega}(L)$. For this we define when a Borel L-structure $\mathcal{P}$ and a measure m *witness* T, generalizing the key property of Petrov and Vershik's *universal measurable graphs* in [PV10, Theorem 2]. From this, we define when $\mathcal{P}$ *strongly witnesses* T, a notion that we find more convenient to apply.

We begin by defining the notions of *internal* and *external witnesses*.

Definition 3.7. Let $\mathcal{M}$ be an L-structure containing a tuple $\mathbf{a}$, and let $\psi(\mathbf{x}, y)$ be a quantifier-free formula of $\mathcal{L}_{\omega_1,\omega}(L)$, all of whose free variables are among $\mathbf{x}y$. We say that an element $b \in \mathcal{M}$ is a **witness** for $(\exists y)\psi(\mathbf{a}, y)$ when $\mathcal{M} \models \psi(\mathbf{a}, b)$. We say that such an element b is an **internal witness** when $b \in \mathbf{a}$, and an **external witness** otherwise.

Recall that a measure m on $\mathbb{R}$ is said to be **nondegenerate** when every nonempty open set has positive measure.

Definition 3.8. Let $\mathcal{P}$ be a Borel L-structure and let m be a probability measure on $\mathbb{R}$. Suppose T is a countable pithy Π_2 theory of $\mathcal{L}_{\omega_1,\omega}(L)$. We say that the pair $(\mathcal{P}, m)$ **witnesses** T if for every sentence $(\forall\mathbf{x})(\exists y)\psi(\mathbf{x}, y) \in T$, and for every tuple $\mathbf{a} \in \mathcal{P}$ such that $|\mathbf{a}| = |\mathbf{x}|$, we have either

(i) $\mathcal{P} \models \psi(\mathbf{a}, b)$ for some $b \in \mathbf{a}$, or

(ii) $m\big(\{b \in \mathcal{P} : \mathcal{P} \models \psi(\mathbf{a}, b)\}\big) > 0$.

We say that $\mathcal{P}$ **strongly witnesses** T when, for every nondegenerate probability measure m on $\mathbb{R}$, the pair $(\mathcal{P}, m)$ witnesses T.

Intuitively, the two possibilities *(i)* and *(ii)* say that witnesses for $(\exists y)\psi(\mathbf{a}, y)$ are easy to find: Either an internal witness already exists among the parameters $\mathbf{a}$, or else witnesses are plentiful elsewhere in the structure $\mathcal{P}$, according to m.

Strong witnesses simply allow us to work without keeping track of a measure m. In fact, when we build structures $\mathcal{P}$ that strongly witness a theory, in §3.4, we will be more concrete, by declaring entire intervals to be external witnesses.

These definitions generalize two of the key notions in [PV10]. Let L_G be the language of graphs. A *universal measurable graph* (X, m, E) as defined in [PV10, Definition 3] roughly corresponds to a Borel L_G-structure (X, E) with vertex set X and edge relation E for which $((X, E), m)$ witnesses the theory of the Rado graph $\mathscr{R}$. A *topologically universal graph* (X, E) as defined in [PV10, Definition 4] roughly corresponds to a Borel L_G-structure that strongly witnesses the theory of $\mathscr{R}$ by virtue of entire intervals being witnesses. Theorem 3.10 and Corollary 3.11 below are inspired directly by Petrov and Vershik's constructions.

We will use (continuum-sized) samplable Borel L-structures strongly witnessing a countable pithy Π_2 theory T to produce random *countable* structures that satisfy T almost surely. However, we first note that the property of strongly witnessing such a T is powerful enough to ensure that a samplable Borel L-structure is itself a model of T.

Lemma 3.9. *Let $\mathcal{P}$ be a samplable Borel L-structure, and let T be a countable pithy Π_2 theory of $\mathcal{L}_{\omega_1,\omega}(L)$. If $\mathcal{P}$ strongly witnesses T, then $\mathcal{P} \models T$.*

Proof. Fix a pithy Π_2 sentence $(\forall \mathbf{x})(\exists y)\psi(\mathbf{x}, y) \in T$. Suppose $\mathbf{a} \in \mathcal{P}$, where $|\mathbf{a}| = |\mathbf{x}|$. If possibility *(i)* of Definition 3.8 holds, then there is an internal witness for $(\exists y)\psi(\mathbf{a}, y)$, i.e., there is some $b \in \mathbf{a}$ such that $\mathcal{P} \models \psi(\mathbf{a}, b)$. Otherwise, possibility *(ii)* holds, and so the set $\{b \in \mathcal{P} : \mathcal{P} \models \psi(\mathbf{a}, b)\}$ of external witnesses has positive m-measure for an arbitrary nondegenerate probability measure m on $\mathbb{R}$; in particular, this set is nonempty. Either way, for all $\mathbf{a} \in \mathcal{P}$ we have $\mathcal{P} \models (\exists y)\psi(\mathbf{a}, y)$, and therefore $\mathcal{P} \models (\forall \mathbf{x})(\exists y)\psi(\mathbf{x}, y)$. Thus $\mathcal{P} \models T$. ☐

When the measure m is continuous, samples from $\mu_{(\mathcal{P},m)}$ are a.s. models of T.

Theorem 3.10. *Let T be a countable pithy Π_2 theory of $\mathcal{L}_{\omega_1,\omega}(L)$, and let $\mathcal{P}$ be a samplable Borel L-structure. Suppose m is a continuous probability measure on $\mathbb{R}$ such that $(\mathcal{P}, m)$ witnesses T. Then $\mu_{(\mathcal{P},m)}$ is concentrated on the set of structures in Str_L that are models of T.*

Proof. Let $\mathbf{A} = (a_i)_{i\in\omega}$ be an m-i.i.d. sequence of elements of $\mathcal{P}$. Recall that by the proof of Lemma 3.6, $\mu_{(\mathcal{P},m)}$ is the distribution of the random structure $\mathcal{F}_{\mathcal{P}}(\mathbf{A})$, and so we must show that $\mathcal{F}_{\mathcal{P}}(\mathbf{A}) \models T$ a.s. Because T is countable, it suffices by countable additivity to show that for any sentence $\varphi \in T$, we have $\mathcal{F}_{\mathcal{P}}(\mathbf{A}) \models \varphi$ a.s.

Recall that T is a countable pithy Π_2 theory. Suppose $(\forall \mathbf{x})(\exists y)\psi(\mathbf{x}, y) \in T$, and let $k = |\mathbf{x}|$ (which may be 0). Our task is to show that, with probability one,

$$\mathcal{F}_{\mathcal{P}}(\mathbf{A}) \models (\forall \mathbf{x})(\exists y)\psi(\mathbf{x}, y).$$

Fix $t_1 \cdots t_k \in \mathbb{N}$. We will show that, with probability one,

$$\mathcal{F}_{\mathcal{P}}(\mathbf{A}) \models (\exists y)\psi(t_1 \cdots t_k, y). \tag{†}$$

Consider the random tuple $a_{t_1} \cdots a_{t_k}$. Because $\mathcal{P}$ strongly witnesses T, by Definition 3.8 it is *surely* the case that either

(i) for some ℓ such that $1 \le \ell \le k$, the random real a_{t_ℓ} is an internal witness for $(\exists y)\psi(a_{t_1} \cdots a_{t_k}, y)$, i.e.,

$$\mathcal{P} \models \psi(a_{t_1} \cdots a_{t_k}, a_{t_\ell}),$$

or else

(ii) the (random) set of witnesses for $(\exists y)\psi(a_{t_1} \cdots a_{t_k}, y)$ has positive measure, i.e.,

$$m\big(\{b \in \mathbb{R} : \mathcal{P} \models \psi(a_{t_1} \cdots a_{t_k}, b)\}\big) > 0.$$

In case *(i)*, we have

$$\mathcal{F}_{\mathcal{P}}(\mathbf{A}) \models \psi(t_1 \cdots t_k, t_\ell),$$

where ℓ is as above, and so (†) holds surely.

Now suppose case *(ii)* holds, and condition on $a_{t_1} \cdots a_{t_k}$. Then

$$\beta := m\big(\{b \in \mathbb{R} : \mathcal{P} \models \psi(a_{t_1} \cdots a_{t_k}, b)\}\big)$$

is a positive constant. For each $n \in \mathbb{N}$ not among $t_1, \dots, t_k$, the random element a_n is m-distributed, and so the events

$$\mathcal{P} \models \psi(a_{t_1} \cdots a_{t_k}, a_n)$$

each have probability β. These events are also mutually independent for such n, and so with probability one, there is some $s \in \mathbb{N}$ for which

$$\mathcal{F}_{\mathcal{P}}(\mathbf{A}) \models \psi(t_1 \cdots t_k, s).$$

Therefore, in this case, (†) holds almost surely. □

Finally, we show that given a Borel L-structure strongly witnessing T, we can construct an invariant measure concentrated on the set of models of T in Str_L.

Corollary 3.11. *Let T be a countable pithy Π_2 theory of $\mathcal{L}_{\omega_1,\omega}(L)$, and let $\mathcal{P}$ be a samplable Borel L-structure. Suppose that $\mathcal{P}$ strongly witnesses T. Then there is an invariant measure on Str_L that is concentrated on the set of structures in Str_L that are models of T.*

Proof. Let m be a nondegenerate probability measure on $\mathbb{R}$ that is continuous (e.g., a Gaussian or Cauchy distribution). By Lemmas 3.5 and 3.6, $\mu_{(\mathcal{P},m)}$ is an invariant measure, and by Theorem 3.10, it is concentrated on the set of models of T in Str_L. □

Our constructions of invariant measures all employ a samplable Borel L-structure $\mathcal{P}$ that strongly witnesses a countable pithy Π_2 theory T. We note that the machinery developed in §3.2 could have been used to build invariant measures via the substantially weaker condition that, for a given probability measure m, for each expression $(\exists y)\psi(\mathbf{a}, y)$ for which $(\forall \mathbf{x})(\exists y)\psi(\mathbf{x}, y) \in T$, there are witnesses in $\mathcal{P}$ only for m-almost all tuples $\mathbf{a}$. However, in this case, $\mathcal{P}$ need not be a model of T (in contrast to Lemma 3.9), nor need $(\mathcal{P}, m)$ even witness T. Also, while we defined the notion of witnessing only for L-structures on $\mathbb{R}$, we could have developed a similar notion for L-structures whose underlying set is an m-measure one subset of $\mathbb{R}$.

Next we find conditions that allow us to construct samplable Borel L-structures that strongly witness T. When all countable models of T are isomorphic to a particular countable infinite L-structure $\mathcal{M}$, we will thereby obtain an invariant measure concentrated on the isomorphism class of $\mathcal{M}$ in Str_L.

3.3. **Duplication of quantifier-free types.**

We now introduce the notion of a theory having *duplication of quantifier-free types*. We will see in §3.4 that when L is a countable relational language and T is a countable pithy Π_2 theory of $\mathcal{L}_{\omega_1,\omega}(L)$, duplication of quantifier-free types guarantees the existence of a samplable Borel L-structure strongly witnessing T. However, for the definitions and results in §3.3, we do not require L to be relational.

We first recall the notion of a *quantifier-free type*, which can be thought of as giving the full description of the subbasic formulas that could hold among the elements of some tuple. In first-order logic this is typically achieved with an infinite consistent set of formulas, but in our infinitary context, a single satisfiable formula of $\mathcal{L}_{\omega_1,\omega}(L)$ suffices.

Definition 3.12. Suppose $\mathbf{x}$ is a finite tuple of variables. Define a **(complete) quantifier-free type** $p(\mathbf{x})$ **of** $\mathcal{L}_{\omega_1,\omega}(L)$ to be a quantifier-free formula in $\mathcal{L}_{\omega_1,\omega}(L)$, whose free variables are precisely those in $\mathbf{x}$, such that the sentence $(\exists \mathbf{x})p(\mathbf{x})$ has a model, and such that for every quantifier-free formula $\varphi(\mathbf{x}) \in \mathcal{L}_{\omega_1,\omega}(L)$, either

$$\models p(\mathbf{x}) \to \varphi(\mathbf{x}) \qquad \text{or} \qquad \models p(\mathbf{x}) \to \neg\varphi(\mathbf{x})$$

holds.

Note that because we require this condition only for quantifier-free formulas φ, it suffices for $p(\mathbf{x})$ to be a quantifier-free formula such that whenever $\zeta(\mathbf{x})$ is a subbasic formula, and whenever $\mathbf{y}$ is a tuple of length equal to the number of free variables of ζ such that every variable of $\mathbf{y}$ is in the tuple $\mathbf{x}$, either

$$\models p(\mathbf{x}) \to \zeta(\mathbf{y}) \qquad \text{or} \qquad \models p(\mathbf{x}) \to \neg\zeta(\mathbf{y}).$$

By taking countable conjunctions, we see that every tuple in every L-structure satisfies some complete quantifier-free type (in the sense of Definition 3.12). This justifies our use of the word *type* for a single formula.

We will often call complete quantifier-free types of $\mathcal{L}_{\omega_1,\omega}(L)$ simply *quantifier-free types*, and sometimes refer to them as *quantifier-free $\mathcal{L}_{\omega_1,\omega}(L)$-types*. Although we have required that a quantifier-free type $p(\mathbf{x})$ have free variables precisely those in $\mathbf{x}$, when there is little possibility of confusion we will sometimes omit the tuple $\mathbf{x}$ and refer to the quantifier-free type as p.

We say that a quantifier-free type $p(\mathbf{x})$ is **consistent with** a theory T when $T \cup (\exists \mathbf{x})p(\mathbf{x})$ has a model. A tuple $\mathbf{a}$ in an L-structure $\mathcal{M}$, where $|\mathbf{a}| = |\mathbf{x}|$, is said to **realize** the quantifier-free type $p(\mathbf{x})$ when $\mathcal{M} \models p(\mathbf{a})$; in this case we say that $p(\mathbf{x})$ is *the* quantifier-free type of $\mathbf{a}$ (as it is unique up to equivalence).

Suppose that $p(\mathbf{x})$ and $q(\mathbf{y})$ are quantifier-free types, where $\mathbf{y}$ is a tuple of variables containing those in $\mathbf{x}$. We say that q **extends** p, or that p is the **restriction** of q to $\mathbf{x}$, when $\models q(\mathbf{y}) \to p(\mathbf{x})$.

Definition 3.13. A quantifier-free type $p(x_1, \ldots, x_n)$ is said to be **non-redundant** when it implies the formula $\bigwedge_{1\le i<j\le n}(x_i \neq x_j)$.

Note that every quantifier-free type is equivalent to the conjunction of a non-redundant quantifier-free type and equalities of variables, as follows. Suppose q is a quantifier-free type. Let S be the set containing those formulas $(u = v)$, for u and v free variables of q, such that q implies $(u = v)$. Then q is equivalent to

$$r \wedge \bigwedge_{\eta \in S} \eta$$

for some non-redundant quantifier-free type r.

The following notion will be used in the stages of Construction 3.18 that we call "refinement".

Definition 3.14. We say that a theory T has **duplication of quantifier-free types** when, for every non-redundant quantifier-free type $p(x, \mathbf{z})$ consistent with T, there is a non-redundant quantifier-free type $q(x, y, \mathbf{z})$ consistent with T such that

$$\models q(x, y, \mathbf{z}) \to \big(p(x, \mathbf{z}) \wedge p(y, \mathbf{z})\big),$$

where $p(y, \mathbf{z})$ denotes the quantifier-free type $p(x, \mathbf{z})$ with all instances of the variable x replaced by the variable y.

Equivalently, when T does not have duplication of quantifier-free types, there is some non-redundant quantifier-free type $p(x, \mathbf{z})$ consistent with T and some model $\mathcal{M}$ of T containing a tuple $(a, \mathbf{b})$ realizing p such that the only way for a tuple $(a', \mathbf{b})$ in $\mathcal{M}$ to also realize p is for a' to equal a.

3.4. **Construction of a Borel L-structure strongly witnessing a theory.**
Throughout §3.4, let L be a countable *relational* language and let T be a countable pithy Π_2 theory of $\mathcal{L}_{\omega_1,\omega}(L)$ that has duplication of quantifier-free types. We now construct a Borel L-structure $\mathcal{P}$ that strongly witnesses T. This construction is inspired by [PV10, Theorem 5], in which Petrov and Vershik build an analogous continuum-sized structure realizing the theory of the Henson graph $\mathscr{H}_3$. We begin with an informal description.

We construct $\mathcal{P}$ by assigning, for every increasing tuple of reals, the quantifier-free type that it realizes. This will determine the quantifier-free type of every tuple of reals, and hence determine the structure $\mathcal{P}$ on $\mathbb{R}$, as we now explain.

Definition 3.15. Given (strictly) increasing tuples of reals $\mathbf{c}$ and $\mathbf{d}$, we say that $\mathbf{c}$ **isolates d** when every left-half-open interval whose endpoints are consecutive entries of $\mathbf{c}$ contains exactly one entry of $\mathbf{d}$, i.e., for $0 \le j < \ell$,

$$d_j \in (c_j, c_{j+1}],$$

where $\mathbf{c} = c_0 \cdots c_\ell$ and $\mathbf{d} = d_0 \cdots d_{\ell-1}$.

For example, the triple $(1, 2, 5)$ isolates the pair $(2, 3)$.

Let $\mathbf{c}$ be an arbitrary tuple of reals (not necessarily in increasing order, and possibly with repetition). The quantifier-free type of $\mathbf{c}$ is determined by the quantifier-free type of the increasing tuple containing all reals of $\mathbf{c}$ along with the relative ordering of the entries of $\mathbf{c}$. For example, let $p(x, y, z)$ be the quantifier-free type of $(1, 2, 5)$. Then the quantifier-free type of $(2, 2, 1, 5)$ is the unique (up to equivalence) formula $q(y, w, x, z)$ implied by $p(x, y, z) \wedge (y = w)$. Hence in order to determine the quantifier-free type of every tuple of reals, and thereby define a structure $\mathcal{P}$ on $\mathbb{R}$, it suffices to assign the quantifier-free types of all increasing tuples.

In Construction 3.18, we will build our Borel L-structure $\mathcal{P}$ inductively, making sure that $\mathcal{P}$ strongly witnesses T. At stage $i \ge 0$ of the construction we will define the following quantities:

- $\mathbf{r}_i = (r^i_0, \ldots, r^i_{|\mathbf{r}_i|-1})$, the increasing tuple of all rationals mentioned by the end of stage i;
- p_i, the quantifier-free type of $\mathbf{r}_i$; and
- $\mathbf{v}_i = (v^i_0, \ldots, v^i_{|\mathbf{r}_i|})$, an increasing tuple of irrationals that isolates $\mathbf{r}_i$.

We call the left-half-open intervals

$$\big(-\infty, v^i_0\big],\ \big(v^i_0, v^i_1\big],\ \ldots,\ \big(v^i_{|\mathbf{r}_i|-1}, v^i_{|\mathbf{r}_i|}\big],\ \big(v^i_{|\mathbf{r}_i|}, \infty\big)$$

the *intervals determined by* $\mathbf{v}_i$.

We will define the $\mathbf{v}_i$ so that they form a nested sequence of tuples of irrationals such that every (increasing) tuple that a given $\mathbf{v}_i$ isolates, including $\mathbf{r}_i$, is assigned the same quantifier-free type p_i at stage i. The sequence of tuples $\{\mathbf{v}_j\}_{j\in\omega}$ will be such that the set of reals $\bigcup_{j\in\omega} \mathbf{v}_j$ is dense in $\mathbb{R}$. Thus for every tuple of reals $\mathbf{a}$, all of its entries occur in some increasing tuple isolated by $\mathbf{v}_i$ for some i, and so its quantifier-free type will eventually be defined. This motivates the following definitions.

Definition 3.16. For each stage i, define $\mathcal{B}_i$ to be the set of tuples $\mathbf{c} \in \mathbb{R}$ such that there is some increasing tuple $\mathbf{d} \in \mathbb{R}$ that $\mathbf{v}_i$ isolates and that contains every entry of $\mathbf{c}$.

Note that $\mathcal{B}_i \subseteq \mathcal{B}_{i'}$ for $i \leq i'$, and that $\bigcup_{j\in\omega} \mathcal{B}_j$ contains every tuple of reals.

By the end of stage i, we will have defined the quantifier-free type of every tuple that $\mathbf{v}_i$ isolates, and hence by extension, of every tuple in $\mathcal{B}_i$. For example, if $(1, 2, 5) \in \mathcal{B}_i$, then $(2, 2, 1, 5) \in \mathcal{B}_i$ also, and by the end of stage i the quantifier-free type of $(1, 2, 5)$ will be determined explicitly, and of $(2, 2, 1, 5)$ implicitly, as described above.

Next we define an equivalence relation on $\mathcal{B}_i$, which we call i-equivalence. By the end of stage i, tuples in $\mathcal{B}_i$ that are i-equivalent will have been assigned the same quantifier-free type.

Definition 3.17. Let $i \geq 0$. We say that two tuples $\mathbf{c}, \mathbf{d} \in \mathcal{B}_i$ of the same length ℓ are ***i*-equivalent**, denoted $\mathbf{c} \approx_i \mathbf{d}$, if for all $j \leq \ell$, the jth entry of $\mathbf{c}$ and jth entry of $\mathbf{d}$ both fall into the same interval determined by $\mathbf{v}_i$. Any two elements of $\mathcal{B}_i$ of different lengths are not i-equivalent.

For example, each left-half-open interval $(v^i_j, v^i_{j+1}]$ determined by $\mathbf{v}_i$, where $0 \leq j < |\mathbf{r}_i|$, is the i-equivalence class of any element in the interval.

Note that i'-equivalence refines i-equivalence for $i' > i$, in the sense that given two elements of $\mathcal{B}_i$ that are not i-equivalent, they are also not i'-equivalent. Furthermore, our construction will be such that for any two distinct i-equivalent tuples in $\mathcal{B}_i$, there is some $i' > i$ for which they are not i'-equivalent.

In the construction, we will assign quantifier-free types in such a way that $\mathcal{P}$ strongly witnesses T. Specifically, for every $\mathbf{a} \in \mathcal{B}_i$ and every pithy Π_2 sentence $(\forall \mathbf{x})(\exists y)\psi(\mathbf{x}, y) \in T$, if there is no internal witness for $(\exists y)\psi(\mathbf{a}, y)$, then at some stage $i' > i$ we will build a left-half-open interval I, disjoint from $\mathcal{B}_i$, consisting of i'-equivalent elements all of which are external witnesses for $(\exists y)\psi(\mathbf{a}, y)$. This will imply that for any $\mathbf{c} \approx_{i'} \mathbf{a}$, every $b \in I$ will witness $(\exists y)\psi(\mathbf{c}, y)$, since $\mathbf{a}b \approx_{i'} \mathbf{c}b$ and i'-equivalent tuples realize the same quantifier-free type.

We will build these external witnesses in the even-numbered stages of the construction; we call this process *enlargement*, because we extend the portion of the real line to which we assign quantifier-free types. In the odd-numbered stages, we perform *refinement* of intervals, so that distinct i-equivalent tuples in $\mathcal{P}$ are eventually not i'-equivalent for some $i' > i$; this ensures that each expression $(\exists y)\psi(\mathbf{a}, y)$ will be witnessed with respect to all possible $\mathbf{a} \in \mathcal{P}$. In fact, as we have noted earlier, by the end of stage i, we will have assigned the quantifier-free type of every tuple in $\mathcal{B}_i$ in such a way that i-equivalent tuples have the same quantifier-free type.

Construction 3.18. Fix an enumeration $\{\varphi_i\}_{i\in\omega}$ of the sentences of T such that every sentence of T appears infinitely often. Because T is a pithy Π_2 theory, for each i, the sentence φ_i is of the form

$$(\forall \mathbf{x})(\exists y)\psi_i(\mathbf{x}, y),$$

where ψ_i is a quantifier-free formula whose free variables are precisely $\mathbf{x}y$, all distinct, and where $\mathbf{x}$ is possibly empty. Consider the induced enumeration $\{\psi_i\}_{i\in\omega}$, and for each i, let k_i be one less than the number of free variables of ψ_i. Also fix an enumeration $\{q_i\}_{i\in\omega}$ of the rationals.

We now give the inductive construction. For a diagram, see Figure 1. The key inductive property is that at the end of each stage i, the quantifier-free type p_i is consistent with T, extends p_{i-1} (for $i \geq 1$), and is the (non-redundant) quantifier-free type of every tuple that $\mathbf{v}_i$ isolates, including $\mathbf{r}_i$.

Stage 0: Set $\mathbf{r}_0$ to be the tuple (0), let p_0 be any quantifier-free unary type consistent with T, and set $\mathbf{v}_0$ to be the pair $\left(-\sqrt{2}, \sqrt{2}\right)$.

Stage $\mathbf{2i+1}$ (Refinement): In stage $2i+1$, we will construct a tuple $\mathbf{r}_{2i+1}$ of rationals, a tuple $\mathbf{v}_{2i+1}$ of irrationals, and a non-redundant quantifier-free type p_{2i+1} consistent with T in such a way that these extend $\mathbf{r}_{2i}$, $\mathbf{v}_{2i}$, and p_{2i}, respectively. In doing so, we will refine the intervals determined by $\mathbf{v}_{2i}$, and assign the quantifier-free type of every increasing tuple that $\mathbf{v}_{2i+1}$ isolates. By extension, this will determine the quantifier-free type of every tuple in $\mathcal{B}_{2i+1}$, i.e., of every tuple $\mathbf{c}$ all of whose entries are contained in some tuple that $\mathbf{v}_{2i+1}$ isolates, in such a way that $(2i+1)$-equivalent tuples are assigned the same quantifier-free type.

Define $\mathbf{r}_{2i+1}$ to be the increasing tuple consisting of the entries of $\mathbf{r}_{2i}$ along with q_i. We need to define the quantifier-free type p_{2i+1} of $\mathbf{r}_{2i+1}$ so that it extends p_{2i}, the quantifier-free type of $\mathbf{r}_{2i}$. There are three cases, depending on the value of q_i.

Case 1: The "new" rational q_i is already an entry of $\mathbf{r}_{2i}$. In this case, there is nothing to be done, as $\mathbf{r}_{2i+1} = \mathbf{r}_{2i}$, and so we set $p_{2i+1} := p_{2i}$.

Case 2: We have $q_i \in (-\infty, v_0^{2i}] \cup (v_{|\mathbf{r}_{2i}|}^{2i}, \infty)$, i.e., the singleton tuple (q_i) is not in $\mathcal{B}_{2i}$. If $q_i < v_0^{2i}$, let $p_{2i+1}(x_0, \ldots, x_{|\mathbf{r}_{2i}|})$ be any non-redundant quantifier-free type consistent with T that implies $p_{2i}(x_1, \ldots, x_{|\mathbf{r}_{2i}|})$. Similarly, if $q_i > v_{|\mathbf{r}_{2i}|}^{2i}$, let $p_{2i+1}(x_0, \ldots, x_{|\mathbf{r}_{2i}|})$ be such that it implies $p_{2i}(x_0, \ldots, x_{|\mathbf{r}_{2i}|-1})$. Such quantifier-free types p_{2i+1} must exist, because p_{2i} is consistent with T and we have not yet determined the set of relations that hold of any tuple that has an entry lying outside the interval $(v_0^{2i}, v_{|\mathbf{r}_{2i}|}^{2i}]$.

Case 3: Otherwise. Namely, $q_i \in (v_j^{2i}, v_{j+1}^{2i}]$ for some j such that $0 \leq j < |\mathbf{r}_{2i}|$, and $q_i \neq r_j^{2i}$. Note that $\mathbf{v}_{2i}$ isolates each of the tuples

$$r_0^{2i} \cdots r_{|\mathbf{r}_{2i}|-1}^{2i} \qquad \text{and} \qquad r_0^{2i} \cdots r_{j-1}^{2i}\, q_i\, r_{j+1}^{2i} \cdots r_{|\mathbf{r}_{2i}|-1}^{2i}$$

and hence by our construction, the tuples both satisfy the same quantifier-free type p_{2i}. By our assumption that T has duplication of quantifier-free types, and because p_{2i} is non-redundant, there must be a non-redundant quantifier-free type $p_{2i+1}(x_0, \ldots, x_{|\mathbf{r}_{2i}|})$ consistent with T that implies

$$p_{2i}(x_0, \ldots, x_j, x_{j+2}, \ldots, x_{|\mathbf{r}_{2i}|}) \wedge p_{2i}(x_0, \ldots, x_{j-1}, x_{j+1}, \ldots, x_{|\mathbf{r}_{2i}|}).$$

Whichever case holds, now let $\mathbf{v}_{2i+1}$ be any increasing tuple of irrationals that contains every entry of $\mathbf{v}_{2i}$ and isolates $\mathbf{r}_{2i+1}$. Consider the subtuple of variables $\mathbf{z} \subseteq x_0 \cdots x_{|\mathbf{r}_{2i}|}$ that corresponds to the positions of the entries of $\mathbf{r}_{2i}$ within $\mathbf{r}_{2i+1}$. In each case above, $p_{2i+1}(x_0, \ldots, x_{|\mathbf{r}_{2i}|})$ is a non-redundant quantifier-free type consistent with T whose restriction to $\mathbf{z}$ is $p_{2i}(\mathbf{z})$. Hence we may assign the quantifier-free type of every increasing tuple that $\mathbf{v}_{2i+1}$ isolates (including $\mathbf{r}_{2i+1}$) to be p_{2i+1}. By extension, this determines the quantifier-free type of every tuple in $\mathcal{B}_{2i+1}$.

Stage $2i+2$ (Enlargement): In stage $2i+2$, we will construct a tuple $\mathbf{r}_{2i+2}$ of rationals, a tuple $\mathbf{v}_{2i+2}$ of irrationals, and a non-redundant quantifier-free type p_{2i+2} consistent with T in such a way that these extend $\mathbf{r}_{2i+1}$, $\mathbf{v}_{2i+1}$, and p_{2i+1}, respectively. As we do so, we will enlarge the portion of the real line to which we assign quantifier-free types. At the end of the stage we will have determined the quantifier-free type of every tuple in $\mathcal{B}_{2i+2}$, in such a way that $(2i+2)$-equivalent tuples are assigned the same quantifier-free type.

Our goal is to provide witnesses for $(\exists y)\psi_i(\mathbf{a}, y)$, where $\psi_i(\mathbf{x}, y)$ is from our enumeration above, for each k_i-tuple of reals $\mathbf{a} \in \mathcal{B}_{2i+1}$, i.e., each k_i-tuple whose quantifier-free type is determined by p_{2i+1}. We extend the non-redundant quantifier-free type $p_{2i+1}(x_0, \ldots, x_{|\mathbf{r}_{2i+1}|-1})$ to a non-redundant quantifier-free type $p_{2i+2}(x_0, \ldots, x_{|\mathbf{r}_{2i+1}|-1}, \mathbf{w})$ so that for every tuple $\mathbf{a}$ such that $(\exists y)\psi_i(\mathbf{a}, y)$ has no internal witness, there is an entry of $\mathbf{w}$ whose realizations provide external witnesses.

Let $\{\mathbf{z}_\ell\}_{1 \le \ell \le N_i}$ be an enumeration of those tuples of variables (possibly with repetition) of length k_i all of whose entries are among $x_0, \ldots, x_{|\mathbf{r}_{2i+1}|-1}$. We now define, by induction on ℓ, intermediate non-redundant quantifier-free types

$$s_\ell(x_0, \ldots, x_{|\mathbf{r}_{2i+1}|-1}, \mathbf{u}_\ell),$$

for $0 \le \ell < N_i$, that are consistent with T and such that each $s_{\ell+1}$ implies s_ℓ. As we step through the tuples of variables of length k_i, if we have already provided a "witness" for $(\exists y)\psi_i(\mathbf{z}_{\ell+1}, y)$ then we will do nothing; otherwise, we will extend our quantifier-free type to provide one, as we now describe.

Let $s_0 := p_{2i+1}$, and let $\mathbf{u}_0$ be the empty tuple of variables. Now consider step $\ell < N_i$ of the induction, so that $s_0, \ldots, s_\ell$ have been defined. If there is a variable t among $x_0 \cdots x_{|\mathbf{r}_{2i+1}|-1}$ or among $\mathbf{u}_\ell$ such that $s_\ell(x_0, \ldots, x_{|\mathbf{r}_{2i+1}|-1}, \mathbf{u}_\ell)$ implies $\psi_i(\mathbf{z}_{\ell+1}, t)$, then let $s_{\ell+1} := s_\ell$ and $\mathbf{u}_{\ell+1} := \mathbf{u}_\ell$. If not, then because s_ℓ is consistent with T and $(\forall \mathbf{x})(\exists y)\psi_i(\mathbf{x}, y) \in T$, there must be some non-redundant quantifier-free type $s_{\ell+1}$ consistent with T that has one more variable, $w_{\ell+1}$, than s_ℓ, such that $s_{\ell+1}$ implies both s_ℓ and $\psi_i(\mathbf{z}_{\ell+1}, w_{\ell+1})$; in this case, let $\mathbf{u}_{\ell+1} := \mathbf{u}_\ell w_{\ell+1}$. Let

$$p_{2i+2}(x_0, \ldots, x_{|\mathbf{r}_{2i+1}|-1}, \mathbf{w}) := s_{N_i}(x_0, \ldots, x_{|\mathbf{r}_{2i+1}|-1}, \mathbf{w}),$$

where $\mathbf{w} := \mathbf{u}_{N_i}$. Note that $p_{2i+2}(x_0, \ldots, x_{|\mathbf{r}_{2i+1}|-1}, \mathbf{w})$ is a non-redundant quantifier-free type that is consistent with T and extends $p_{2i+1}(x_0, \ldots, x_{|\mathbf{r}_{2i+1}|-1})$.

Next, choose $|\mathbf{w}|$-many rationals greater than all entries of $\mathbf{r}_{2i+1}$, and define $\mathbf{r}_{2i+2}$ to be the increasing tuple consisting of $\mathbf{r}_{2i+1}$ and these new rationals. Let $\mathbf{v}_{2i+2}$ be an arbitrary increasing tuple of irrationals that contains every entry of $\mathbf{v}_{2i+1}$ and isolates $\mathbf{r}_{2i+2}$. Finally, for every increasing tuple that $\mathbf{v}_{2i+2}$ isolates (including $\mathbf{r}_{2i+2}$), declare its quantifier-free type to be p_{2i+2}. As with the refinement stages, this determines by extension the quantifier-free type of every tuple in $\mathcal{B}_{2i+2}$, i.e., of every tuple $\mathbf{c}$ all of whose entries are contained in some tuple that $\mathbf{v}_{2i+2}$ isolates. In particular, for any tuple $\mathbf{a} \in \mathcal{B}_{2i+1}$ of length k_i such that $(\exists y)\psi_i(\mathbf{a}, y)$ does not have internal witnesses, we have constructed a left-half-open interval $(v_j^{2i+2}, v_{j+1}^{2i+2}]$, for some j such that $0 \leq j < |\mathbf{r}_{2i+2}|$, consisting of external witnesses for $(\exists y)\psi_i(\mathbf{a}, y)$. This ends the stage, and the construction. □

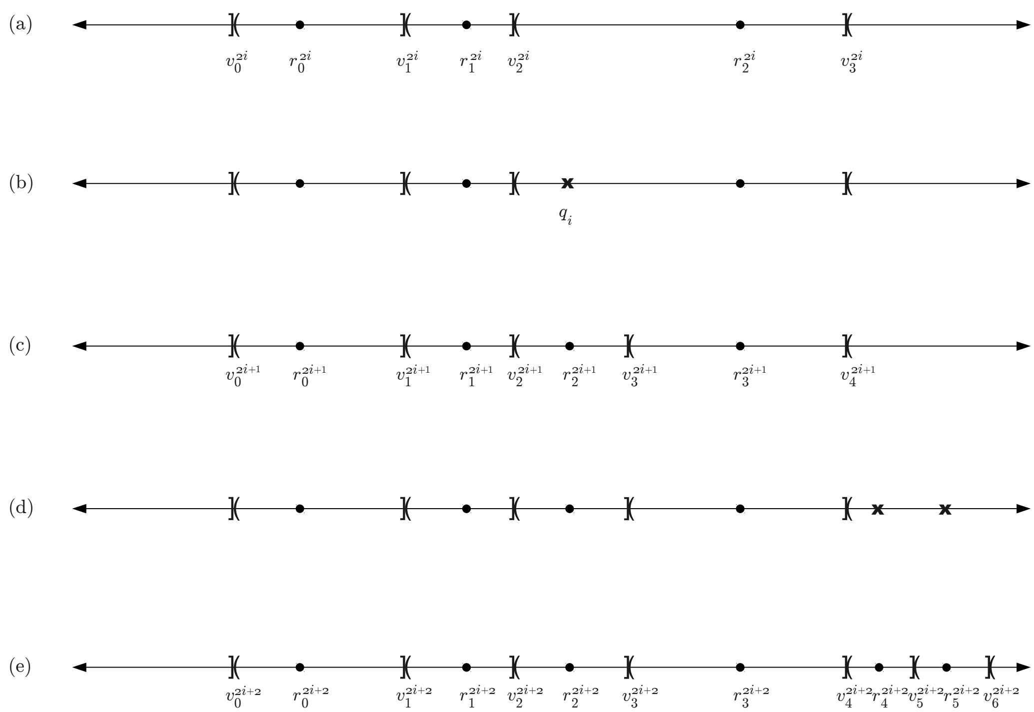

FIGURE 1. An illustration of Construction 3.18.
(a) Suppose that we start stage $2i+1$ with the tuple $\mathbf{r}_{2i} = \left(r_0^{2i}, r_1^{2i}, r_2^{2i}\right)$ of rationals and the tuple $\mathbf{v}_{2i} = \left(v_0^{2i}, v_1^{2i}, v_2^{2i}, v_3^{2i}\right)$ of irrationals.
(b) Suppose that the rational q_i falls between v_2^{2i} and r_2^{2i}.
(c) By the end of stage $2i+1$, the rational q_i has become r_2^{2i+1}, and the rationals and irrationals to its right are reindexed.
(d) Suppose that in stage $2i+2$ we need two intervals of external witnesses for $(\exists y)\psi_i(\mathbf{a}, y)$ as $\mathbf{a}$ ranges among k_i-tuples all of whose entries are entries of $\mathbf{r}_{2i+1}$. Then we select two new rational witnesses r_4^{2i+2} and r_5^{2i+2} to the right of $\mathbf{v}_{2i+1}$.
(e) At the end of stage $2i+2$, we choose irrational interval boundaries v_5^{2i+2} (between r_4^{2i+2} and r_5^{2i+2}) and v_6^{2i+2} (to the right of r_5^{2i+2}).

We now verify that this construction produces a structure with the desired properties.

Theorem 3.19. *Let L be a countable relational language, let T be a countable pithy Π_2 theory of $\mathcal{L}_{\omega_1,\omega}(L)$ that has duplication of quantifier-free types, and let $\mathcal{P}$ be the L-structure obtained via Construction 3.18. Then $\mathcal{P}$ is a samplable Borel L-structure that strongly witnesses T.*

Proof. We first show that $\mathcal{P}$ is a Borel L-structure. Fix a relation symbol $R \in L$, and let k be the arity of R. We must show that $\{\mathbf{a} \in \mathcal{P}^k : \mathcal{P} \models R(\mathbf{a})\}$ is Borel. For each $i \geq 0$ define

$$\mathcal{X}_i := \{\mathbf{a} \in \mathcal{P}^k : \mathcal{P} \models R(\mathbf{a}) \text{ and } \mathbf{a} \in \mathcal{B}_i\}.$$

Recall that by our construction, the set of reals $\bigcup_{j\in\omega} \mathbf{v}_j$ is dense in $\mathbb{R}$, and so every tuple of reals is in $\bigcup_{j\in\omega} \mathcal{B}_j$. Therefore

$$\{\mathbf{a} \in \mathcal{P}^k : \mathcal{P} \models R(\mathbf{a})\} = \bigcup_{j\in\omega} \mathcal{X}_j.$$

In particular, it suffices to show that $\mathcal{X}_j$ is Borel for each j.

Fix some $i \geq 0$, and note that for every $\mathbf{a}, \mathbf{a}' \in \mathcal{B}_i$ such that $\mathbf{a} \approx_i \mathbf{a}'$, we have

$$\mathcal{P} \models R(\mathbf{a}) \qquad \text{if and only if} \qquad \mathcal{P} \models R(\mathbf{a}'),$$

because i-equivalent tuples are assigned the same quantifier-free type. Furthermore, for every $\mathbf{a} \in \mathcal{B}_i$, the set

$$\{\mathbf{c} \in \mathcal{B}_i : \mathbf{c} \approx_i \mathbf{a}\}$$

is a k-fold product of left-half-open intervals. Hence $\mathcal{X}_i$ is Borel, and so $\mathcal{P}$ is a Borel L-structure. Moreover, $\mathcal{P}$ is samplable because L is relational.

We now show that $\mathcal{P}$ strongly witnesses T. Let m be an arbitrary nondegenerate probability measure on $\mathbb{R}$. Consider a pithy Π_2 sentence

$$(\forall \mathbf{x})(\exists y)\xi(\mathbf{x}, y) \in T,$$

and let $\mathbf{a}$ be a tuple of reals such that $|\mathbf{a}| = |\mathbf{x}|$, where $|\mathbf{x}|$ could possibly be zero.

Suppose $(\exists y)\xi(\mathbf{a}, y)$ does not have an internal witness. Let ℓ^* be the least stage such that $\mathbf{a} \in \mathcal{B}_{2\ell^*+1}$. Because each sentence of T appears infinitely often in the enumeration $\{\varphi_j\}_{j\in\omega}$, there is some $\ell \geq \ell^*$ such that $\varphi_\ell = (\forall \mathbf{x})(\exists y)\xi(\mathbf{x}, y)$, and hence such that $\xi = \psi_\ell$.

Since $\mathbf{a} \in \mathcal{B}_{2\ell+1}$, at stage $2\ell + 2$ there is some real b such that $\mathcal{P} \models \psi_\ell(\mathbf{a}, b)$. Furthermore, we have ensured that there is a left-half-open interval of reals b' such that $b' \approx_{2\ell+2} b$ and hence such that $\mathcal{P} \models \psi_\ell(\mathbf{a}, b')$. Because m is nondegenerate, this collection of external witnesses for $(\exists y)\xi(\mathbf{a}, y)$ has positive m-measure. Hence $(\mathcal{P}, m)$ witnesses T. As m was an arbitrary nondegenerate probability measure on $\mathbb{R}$, the Borel L-structure $\mathcal{P}$ strongly witnesses T, as desired. □

3.5. **Invariant measures from trivial definable closure.**

We are now ready to prove the positive direction of our main theorem, Theorem 1.1. We have seen, in §3.4, that if a countable pithy Π_2 theory T in a countable relational language L has duplication of quantifier-free types, then there exists a samplable Borel L-structure strongly witnessing T. We show below that when T has a unique countable model $\mathcal{M}$ (up to isomorphism), duplication of quantifier-free types is moreover implied by $\mathcal{M}$ having trivial definable closure; we prove the converse for relational languages in Corollary 4.4. In fact, we have the following stronger result.

Lemma 3.20. *Let T be a countable theory of $\mathcal{L}_{\omega_1,\omega}(L)$ such that every countable model of T has trivial definable closure. Then T has duplication of quantifier-free types.*

Proof. Suppose $p(x, \mathbf{z})$ is a non-redundant quantifier-free $\mathcal{L}_{\omega_1,\omega}(L)$-type consistent with T. Because every model of T has trivial definable closure,

$$T \models p(x, \mathbf{z}) \rightarrow (\exists y)\big(p(y, \mathbf{z}) \wedge (y \neq x)\big).$$

Hence there is some non-redundant quantifier-free $\mathcal{L}_{\omega_1,\omega}(L)$-type $q(x, y, \mathbf{z})$ such that

$$T \models q(x, y, \mathbf{z}) \to \big(p(x, \mathbf{z}) \wedge p(y, \mathbf{z})\big),$$

and so T has duplication of quantifier-free types. □

We now use Theorem 3.19 and Lemma 3.20 to prove the positive direction of Theorem 1.1.

Theorem 3.21. *Let L be a countable language and let $\mathcal{M}$ be a countable infinite L-structure. If $\mathcal{M}$ has trivial definable closure, then there is an invariant probability measure on Str_L that is concentrated on the isomorphism class of $\mathcal{M}$ in Str_L.*

Proof. Without loss of generality, we may assume that $\mathcal{M} \in \mathrm{Str}_L$. Let $\overline{\mathcal{M}}$ be the canonical structure of $\mathcal{M}$ and $L_{\overline{\mathcal{M}}}$ its canonical language. By Proposition 2.17, there is a pithy Π_2 $\mathcal{L}_{\omega_1,\omega}(L_{\overline{\mathcal{M}}})$-theory $T_{\overline{\mathcal{M}}}$ all of whose countable models are isomorphic to $\overline{\mathcal{M}}$. By Lemmas 2.13 and 2.15 and the fact that $\mathcal{M}$ has trivial definable closure, the unique (up to isomorphism) countable model $\overline{\mathcal{M}}$ of $T_{\overline{\mathcal{M}}}$ has trivial definable closure. Hence by Lemma 3.20, the theory $T_{\overline{\mathcal{M}}}$ has duplication of quantifier-free types.

Since $L_{\overline{\mathcal{M}}}$ is relational, by Theorem 3.19 there is a samplable Borel $L_{\overline{\mathcal{M}}}$-structure $\mathcal{Q}$ strongly witnessing $T_{\overline{\mathcal{M}}}$. Therefore, by Corollary 3.11 there is an invariant probability measure on $\mathrm{Str}_{L_{\overline{\mathcal{M}}}}$ that is concentrated on the set of countable $L_{\overline{\mathcal{M}}}$-structures that are isomorphic to $\overline{\mathcal{M}}$, i.e., those that are models of $T_{\overline{\mathcal{M}}}$. Finally, by Lemmas 2.13 and 2.14, there is an invariant probability measure on Str_L that is concentrated on $\mathcal{M}$. □

Although the proof of Theorem 3.21 produces a samplable Borel $L_{\overline{\mathcal{M}}}$-structure, where $L_{\overline{\mathcal{M}}}$ may be quite different from L (in particular, $L_{\overline{\mathcal{M}}}$ is always infinite and relational), we can obtain essentially the same invariant measure via a samplable Borel L-structure. We will use this fact in §6.1.

Corollary 3.22. *Let L be a countable language and let $\mathcal{M}$ be a countable infinite L-structure. If $\mathcal{M}$ has trivial definable closure, then there is a samplable Borel L-structure $\mathcal{P}$ such that for any continuous nondegenerate probability measure m on $\mathbb{R}$, the invariant measure $\mu_{(\mathcal{P},m)}$ is concentrated on the isomorphism class of $\mathcal{M}$ in Str_L.*

Proof. Let m be an arbitrary continuous nondegenerate probability measure on $\mathbb{R}$, and let $\overline{\mathcal{M}}$, $L_{\overline{\mathcal{M}}}$, $T_{\overline{\mathcal{M}}}$, and $\mathcal{Q}$ be as in the proof of Theorem 3.21. In particular, $\mathcal{Q}$ strongly witnesses $T_{\overline{\mathcal{M}}}$, and so by Theorem 3.10, the invariant measure $\mu_{(\mathcal{Q},m)}$ is concentrated on $\overline{\mathcal{M}}$.

Note that, by Lemma 3.9, $\mathcal{Q}$ and $\overline{\mathcal{M}}$ have the same $\mathcal{L}_{\omega_1,\omega}(L_{\overline{\mathcal{M}}})$-theory. By Lemma 2.13, $\mathcal{M}$ and $\overline{\mathcal{M}}$ are interdefinable; let (Ψ_0, Ψ_1) be an interdefinition between $\mathcal{M}$ and $\overline{\mathcal{M}}$. As in Lemma 2.12, let $\mathcal{P}$ be the L-structure that is interdefinable with $\mathcal{Q}$ via the interdefinition (Ψ_0, Ψ_1), so that $\overline{\mathcal{P}} = \mathcal{Q}$, using the notation described after Lemma 2.13. In particular, $\mathcal{P}$ and $\mathcal{M}$ have the same $\mathcal{L}_{\omega_1,\omega}(L)$-theory.

Let $\varphi(\mathbf{x})$ be an arbitrary $\mathcal{L}_{\omega_1,\omega}(L_{\overline{\mathcal{M}}})$-formula. By Lemma 2.10, there is some quantifier-free $\mathcal{L}_{\omega_1,\omega}(L_{\overline{\mathcal{M}}})$-formula $\psi_\varphi(\mathbf{x})$ such that

$$\overline{\mathcal{M}} \models \varphi(\mathbf{x}) \leftrightarrow \psi_\varphi(\mathbf{x}).$$

Because $\overline{\mathcal{P}}$ and $\overline{\mathcal{M}}$ have the same $\mathcal{L}_{\omega_1,\omega}(L_{\overline{\mathcal{M}}})$-theory, we also have

$$\overline{\mathcal{P}} \models \varphi(\mathbf{x}) \leftrightarrow \psi_\varphi(\mathbf{x}).$$

Hence any countable substructure of $\overline{\mathcal{P}}$ isomorphic to $\overline{\mathcal{M}}$ must in fact be an $\mathcal{L}_{\omega_1,\omega}(L_{\overline{\mathcal{M}}})$-elementary substructure of $\overline{\mathcal{P}}$.

As every $\mathcal{L}_{\omega_1,\omega}(L_{\overline{\mathcal{M}}})$-definable set in $\overline{\mathcal{P}}$ is equivalent to a quantifier-free definable set, every definable set in $\overline{\mathcal{P}}$ is Borel. However every $\mathcal{L}_{\omega_1,\omega}(L)$-definable set in $\mathcal{P}$ is an $\mathcal{L}_{\omega_1,\omega}(L_{\overline{\mathcal{M}}})$-definable set in $\overline{\mathcal{P}}$ and so every $\mathcal{L}_{\omega_1,\omega}(L)$-definable set in $\mathcal{P}$ is Borel. Hence for any subbasic formula ζ of $\mathcal{L}_{\omega_1,\omega}(L)$, the set in $\mathcal{P}$ defined by ζ is Borel. Thus $\mathcal{P}$ is a Borel L-structure.

Because $\mathcal{M}$ has trivial definable closure, L has no constant symbols and every function of $\mathcal{M}$ is a choice function. Since $\mathcal{P}$ satisfies the same $\mathcal{L}_{\omega_1,\omega}(L)$-theory as $\mathcal{M}$, every function of $\mathcal{P}$ is also a choice function. Hence $\mathcal{P}$ is samplable.

By Lemmas 3.5 and 3.6, we have that $\mu_{(\mathcal{P},m)}$ is an invariant probability measure on Str_L that is concentrated on the union of isomorphism classes of countable infinite substructures of $\mathcal{P}$. It remains to show that $\mu_{(\mathcal{P},m)}$ is concentrated on the isomorphism class of $\mathcal{M}$, using the fact that $\mu_{(\overline{\mathcal{P}},m)}$ is concentrated on the isomorphism class of $\overline{\mathcal{M}}$.

Suppose a countable infinite set $N \subseteq \mathbb{R}$ is such that the substructure $\mathcal{N}^*$ of $\overline{\mathcal{P}}$ having underlying set N is isomorphic to $\overline{\mathcal{M}}$. We will show that the substructure $\mathcal{N}^\times$ of $\mathcal{P}$ having underlying set N is isomorphic to $\mathcal{M}$. Again as in Lemma 2.12, let $\mathcal{N}$ be the L-structure that is interdefinable with $\mathcal{N}^*$ via the interdefinition (Ψ_0, Ψ_1), so that $\overline{\mathcal{N}} = \mathcal{N}^*$. As noted above, $\overline{\mathcal{N}}$ is an $\mathcal{L}_{\omega_1,\omega}(L_{\overline{\mathcal{M}}})$-elementary substructure of $\overline{\mathcal{P}}$. Therefore for any $\mathcal{L}_{\omega_1,\omega}(L_{\overline{\mathcal{M}}})$-formula $\varphi(\mathbf{x})$,

$$\big\{\mathbf{a} \in \overline{\mathcal{N}} : \overline{\mathcal{P}} \models \varphi(\mathbf{a})\big\} = \big\{\mathbf{a} \in \overline{\mathcal{N}} : \overline{\mathcal{N}} \models \varphi(\mathbf{a})\big\}. \tag{‡}$$

By (‡) and the fact that (Ψ_0, Ψ_1) is an interdefinition between $\mathcal{P}$ and $\overline{\mathcal{P}}$, and between $\mathcal{N}$ and $\overline{\mathcal{N}}$, we have

$$\begin{aligned}\big\{\mathbf{a} \in \mathcal{N}^\times : \mathcal{P} \models \psi(\mathbf{a})\big\} &= \big\{\mathbf{a} \in \overline{\mathcal{N}} : \overline{\mathcal{P}} \models \Psi_0(\psi)(\mathbf{a})\big\} \\ &= \big\{\mathbf{a} \in \overline{\mathcal{N}} : \overline{\mathcal{N}} \models \Psi_0(\psi)(\mathbf{a})\big\} \\ &= \big\{\mathbf{a} \in \mathcal{N} : \mathcal{N} \models \Psi_1(\Psi_0(\psi))(\mathbf{a})\big\} \\ &= \big\{\mathbf{a} \in \mathcal{N} : \mathcal{N} \models \psi(\mathbf{a})\big\}\end{aligned}$$

for every $\mathcal{L}_{\omega_1,\omega}(L)$-formula $\psi(\mathbf{x})$. Further, as $\mathcal{N}^\times$ is a substructure of $\mathcal{P}$, for every *quantifier-free* $\mathcal{L}_{\omega_1,\omega}(L)$-formula $\psi(\mathbf{x})$, we have

$$\big\{\mathbf{a} \in \mathcal{N}^\times : \mathcal{N}^\times \models \psi(\mathbf{a})\big\} = \big\{\mathbf{a} \in \mathcal{N}^\times : \mathcal{P} \models \psi(\mathbf{a})\big\},$$

and so

$$\big\{\mathbf{a} \in \mathcal{N}^\times : \mathcal{N}^\times \models \psi(\mathbf{a})\big\} = \big\{\mathbf{a} \in \mathcal{N} : \mathcal{N} \models \psi(\mathbf{a})\big\}.$$

Hence $\mathcal{N} = \mathcal{N}^\times$, and so $\mathcal{N}^\times$ is isomorphic to $\mathcal{M}$, as $\overline{\mathcal{N}}$ is isomorphic to $\overline{\mathcal{M}}$.

By the fact that $\mu_{(\overline{\mathcal{P}},m)}$ is concentrated on the isomorphism class of $\overline{\mathcal{M}}$ in $\mathrm{Str}_{L_{\overline{\mathcal{M}}}}$, we have

$$m^\infty\Big(\big\{A \in \mathbb{R}^\omega : \mathcal{F}_{\overline{\mathcal{P}}}(A) \cong \overline{\mathcal{M}}\big\}\Big) = 1,$$

as $\mu_{(\overline{\mathcal{P}},m)} = m^\infty \circ \mathcal{F}_{\overline{\mathcal{P}}}^{-1}$. Now suppose $A \in \mathbb{R}^\omega$ is such that $\mathcal{F}_{\overline{\mathcal{P}}}(A) \cong \overline{\mathcal{M}}$, and consider the the set $N \subseteq \mathbb{R}$ of entries of A. By the above, the substructure $\mathcal{N}^\times$ of $\mathcal{P}$ having underlying set N is isomorphic to $\mathcal{M}$, and so $\mathcal{F}_{\mathcal{P}}(A) \cong \mathcal{M}$. Hence

$$\big\{A \in \mathbb{R}^\omega : \mathcal{F}_{\overline{\mathcal{P}}}(A) \cong \overline{\mathcal{M}}\big\} \subseteq \big\{A \in \mathbb{R}^\omega : \mathcal{F}_{\mathcal{P}}(A) \cong \mathcal{M}\big\},$$

and so

$$m^\infty\Big(\big\{A \in \mathbb{R}^\omega : \mathcal{F}_\mathcal{P}(A) \cong \mathcal{M}\big\}\Big) = 1.$$

Therefore $\mu_{(\mathcal{P},m)}$ is concentrated on the isomorphism class of $\mathcal{M}$ in Str_L. □

4. Non-existence of invariant measures

In this section we complete the proofs of Theorems 1.1 and 1.2. We begin by considering the converse of Theorem 3.21, namely, that for any countable language L, a countable infinite L-structure having nontrivial definable closure cannot admit an invariant measure.

Suppose a countable L-structure $\mathcal{M}$ admits an invariant measure. If there exists an element $b \in \mathrm{dcl}_\mathcal{M}(\emptyset)$, then for every $n \in \mathbb{N}$ the measure assigns the same positive probability to the event that n satisfies the quantifier-free type of b, which is not possible. More generally, $\mathcal{M}$ having non-trivial definable closure leads to a contradiction, as we show below; a special case of this has been observed in [Cam90, (4.29)].

In fact, an even more general result holds. Upon taking the special case of $G = S_\infty$, Theorem 4.1 below completes the proof of our main result by establishing that property (1) implies property (2) in Theorem 1.1. Indeed, initially we proved only this special case. However, Alexander Kechris and Andrew Marks noticed that with minor modifications to our original proof, the stronger result Theorem 4.1 holds. We have included this with their permission.

Theorem 4.1. *Let L be a countable language, and let $\mathcal{M}$ be a countable infinite L-structure. Suppose that $G \leq S_\infty$ is the automorphism group of a structure in* Str_L *that has trivial definable closure. If there is a G-invariant probability measure μ on* Str_L *that is concentrated on the isomorphism class of $\mathcal{M}$, then $\mathcal{M}$ must have trivial definable closure.*

Proof. Without loss of generality, we may assume that $\mathcal{M} \in \mathrm{Str}_L$. Suppose, for a contradiction, that there is a tuple $\mathbf{a} \in \mathcal{M}$ and an element $b \in \mathcal{M}$ such that $b \in \mathrm{dcl}_\mathcal{M}(\mathbf{a}) - \mathbf{a}$. By considering the Scott sentence of the structure obtained by expanding $\mathcal{M}$ by constants for the tuple $\mathbf{a}$ (see, e.g., [KK04, Theorem 3.3.5]), we can find a formula $p(\mathbf{x}, y) \in \mathcal{L}_{\omega_1,\omega}(L)$ such that $\mathcal{M} \models p(\mathbf{a}, b)$ and whenever $\mathcal{M} \models p(\mathbf{c}, d)$ there is an automorphism of $\mathcal{M}$ taking $\mathbf{a}b$ to $\mathbf{c}d$ pointwise.

In particular,

$$\mathcal{M} \models (\exists \mathbf{x} y) p(\mathbf{x}, y).$$

Since the measure μ is concentrated on $\mathcal{M}$, hence on L-structures that satisfy $(\exists \mathbf{x} y) p(\mathbf{x}, y)$, we have

$$\mu\big([\![(\exists \mathbf{x} y) p(\mathbf{x}, y)]\!]\big) = 1.$$

By the countable additivity of μ, there is some tuple $\mathbf{m}n \in \mathbb{N}$ such that $\mu\big([\![p(\mathbf{m}, n)]\!]\big) > 0$. Fix such an $\mathbf{m}n$ and let $\alpha := \mu\big([\![p(\mathbf{m}, n)]\!]\big)$.

Note that if $\mathcal{M} \models p(\mathbf{a}, c)$ for some $c \in \mathbb{N}$, then there is an automorphism of $\mathcal{M}$ that fixes $\mathbf{a}$ pointwise and sends b to c; hence, as $b \in \mathrm{dcl}(\mathbf{a})$, such an element c is equal to b. Therefore we have

$$\mathcal{M} \models (\forall \mathbf{x} y_1 y_2)\big((p(\mathbf{x}, y_1) \wedge p(\mathbf{x}, y_2)) \rightarrow (y_1 = y_2)\big).$$

Hence for $j, k \in \mathbb{N}$,

$$[\![p(\mathbf{m}, j)]\!] \cap [\![p(\mathbf{m}, k)]\!] = \emptyset$$

whenever $j \neq k$, and so

$$\mu\big([\![(\exists y)p(\mathbf{m}, y)]\!]\big) = \sum_{j\in\mathbb{N}} \mu\big([\![p(\mathbf{m}, j)]\!]\big).$$

Now let

$$X := \big\{j \in \mathbb{N} : (\exists g \in G)\,\big((g(\mathbf{m}) = \mathbf{m}) \wedge (g(n) = j)\big)\big\}.$$

By the G-invariance of μ, for all $j \in X$ we have

$$\mu\big([\![p(\mathbf{m}, j)]\!]\big) = \alpha.$$

However, because G is the automorphism group of a structure having trivial definable closure, X is infinite. But

$$1 \geq \mu\big([\![(\exists y)p(\mathbf{m}, y)]\!]\big) \geq \sum_{j\in X} \mu\big([\![p(\mathbf{m}, j)]\!]\big) = \sum_{j\in X} \alpha,$$

which is a contradiction as $\alpha > 0$ and X is infinite. □

Consider the properties (1), (2), and (3) of Theorems 1.1 and 1.2. Theorem 4.1 establishes that (3) implies (2), and Theorem 3.21 shows that (2) implies (1). Finally, (1) trivially implies (3), by considering the unique countable infinite structure in the empty language, which has automorphism group S_∞. Hence the proof of Theorem 1.2 is complete.

Kechris and Marks also observed that the equivalence of properties (2) and (3) from Theorem 1.2 can be established more easily than the full theorem, in particular while entirely avoiding the machinery of Section 3. We again include the argument with their permission.

Corollary 4.2 (Kechris–Marks)**.** *Let L be a countable language, and let $\mathcal{M}$ be a countable infinite L-structure. The following are equivalent:*

- (2) *The structure $\mathcal{M}$ has trivial group-theoretic definable closure, i.e., for every finite tuple $\mathbf{a} \in \mathcal{M}$, we have $\mathrm{dcl}_{\mathcal{M}}(\mathbf{a}) = \mathbf{a}$, where $\mathrm{dcl}_{\mathcal{M}}(\mathbf{a})$ is the collection of elements $b \in \mathcal{M}$ that are fixed by all automorphisms of $\mathcal{M}$ fixing $\mathbf{a}$ pointwise.*
- (3) *There is some $\mathcal{N} \in \mathrm{Str}_L$ that has trivial group-theoretic definable closure and is such that there is an $\mathrm{Aut}(\mathcal{N})$-invariant probability measure on Str_L concentrated on the set of elements of Str_L that are isomorphic to $\mathcal{M}$.*

Proof. We already have that (3) implies (2) by Theorem 4.1. But that (2) implies (3) follows by taking $G = \mathrm{Aut}(\mathcal{M})$ and μ to be a Dirac delta measure on the structure $\mathcal{M}$ itself (where we again take $\mathcal{M}$ to be in Str_L without loss of generality). □

As an immediate corollary of Theorem 4.1, we see that any countable infinite structure that admits an invariant measure cannot have constants, and all of its functions must be choice functions. This observation has been used in [Ack16] to classify those commutative n-semigroups as well as those ultrahomogeneous n-semigroups that admit an invariant measure.

Corollary 4.3. *Let L be a countable language, and let $\mathcal{M}$ be a countable infinite L-structure. Suppose that either L has constant symbols or that there is a function symbol $f \in L$ and tuple $\mathbf{a} \in \mathcal{M}$ for which $f^{\mathcal{M}}(\mathbf{a}) \notin \mathbf{a}$. Then there is no invariant probability measure on Str_L that is concentrated on the isomorphism class of $\mathcal{M}$.*

Note that this implies that Corollary 1.3, which characterizes those Fraïssé limits in relational languages that admit invariant measures, does not extend to structures with constants or functions. This is demonstrated, e.g., by Hall's countable universal locally finite group, whose age has the strong amalgamation property [Hod93, §7.1, Example 1], but which does not have group-theoretic trivial definable closure.

Finally, Theorem 4.1 has the following consequence, which (for relational languages) extends Lemma 3.20 on the relationship between trivial definable closure and duplication of quantifier-free types.

Corollary 4.4. *Let L be a countable relational language, and let T be a countable pithy Π_2 theory of $\mathcal{L}_{\omega_1,\omega}(L)$ that has a unique countable model $\mathcal{M}$ (up to isomorphism). Then $\mathcal{M}$ has trivial definable closure if and only if T has duplication of quantifier-free types.*

Proof. From Lemma 3.20 it is immediate that if $\mathcal{M}$ has trivial definable closure then T must have duplication of quantifier-free types.

For the other direction, suppose that T has duplication of quantifier-free types. By Theorem 3.19 there is a samplable Borel L-structure strongly witnessing T, which by Corollary 3.11 induces an invariant probability measure concentrated on the set of models of T in Str_L. Hence there is an invariant measure concentrated on $\mathcal{M}$. But Theorem 4.1 then implies that $\mathcal{M}$ has trivial definable closure. ☐

5. Structures admitting invariant measures

We now consider several important classes of structures, and examine whether or not the structures in these classes admit invariant measures. In §5.1, we show how any countable infinite structure is a quotient of one with trivial definable closure, and of one without. We use this fact to construct countable structures of arbitrary Scott rank that have trivial definable closure and hence admit invariant measures, and to construct ones that do not. In §5.2 we apply our main results, Theorem 1.1 and Corollary 1.3, to examine certain well-known countable infinite structures, and ask whether or not they admit invariant measures. We make use of existing classifications to provide complete lists of countable infinite ultrahomogeneous partial orders, permutations, directed graphs, and graphs, for which such invariant measures exist.

5.1. **Structures with an equivalence relation.**

Suppose we are given a countable infinite structure in a countable relational language with a binary relation symbol. Further, suppose that this symbol is interpreted as an equivalence relation such that every equivalence class has at least two elements. Consider the quotient map on the underlying set that is induced by this equivalence relation. In the case where this quotient map *respects* the remaining relations (in a sense that we will make precise), we can characterize when the original structure does or does not have trivial definable closure. On the other hand, starting with an arbitrary countable infinite structure in a countable relational language, we can "blow up" each element into an equivalence class, and characterize when the resulting structure has trivial definable closure.

We will thereby see, in Corollary 5.4, that every countable structure in a countable relational language is the quotient of one with trivial definable closure, and of one without. We then apply this result to further yield Corollary 5.5, and obtain structures of arbitrary Scott rank that admit invariant measures, as well as ones that do not.

We begin by describing what it means for an equivalence relation to respect the remaining relations in the language.

Definition 5.1. For a relational language L, define $L^+ := L \cup \{\equiv\}$, where $\equiv$ is a new binary relation symbol. Let $\mathcal{N}$ be an L^+-structure. We say that $\equiv$ **respects** L **in** $\mathcal{N}$ if for each k-ary (non-equality) relation symbol $R \in L$,

$$\mathcal{N} \models (\forall x_1, \ldots, x_k, y_1, \ldots, y_k) \bigwedge_{1 \le i \le k} (x_i \equiv y_i) \to \big(R(x_1, \ldots, x_k) \leftrightarrow R(y_1, \ldots, y_k)\big).$$

Such a relation $\equiv$ is often referred to as an *equality*.

When $\equiv$ respects L in $\mathcal{N}$, the structure $\mathcal{N}$ cannot "L-distinguish" between $\equiv$-equivalent elements. In particular, there is a quotient structure induced by the $\equiv$ relation on the underlying set.

For a countable infinite L^+-structure $\mathcal{N}$ in which $\equiv$ respects L, and where every $\equiv$-equivalence class has at least two elements, the size of the $\equiv$-equivalence classes completely determines whether or not $\mathcal{N}$ admits an invariant measure.

Lemma 5.2. *Suppose $\mathcal{N}$ is a countable infinite L^+-structure such that $\equiv$ respects L in $\mathcal{N}$ and such that no $\equiv$-equivalence class has only one element. The following are equivalent:*

(1) *Every $\equiv$-equivalence class of $\mathcal{N}$ has infinitely many elements.*
(2) *There is an invariant measure on Str_{L^+} that is concentrated on the isomorphism class of $\mathcal{N}$.*

Proof. Assume that (1) holds. Whenever $c, c' \in \mathcal{N}$ are such that $\mathcal{N} \models (c \equiv c')$, define $g_{c,c'} : \mathcal{N} \to \mathcal{N}$ to be the map that interchanges c and c' but is constant on all other elements of $\mathcal{N}$. Since $\equiv$ respects L, the map $g_{c,c'}$ is an automorphism of $\mathcal{N}$.

Suppose, for a contradiction, that there are $\mathbf{a}, b \in \mathcal{N}$ such that $b \in \mathrm{dcl}_{\mathcal{N}}(\mathbf{a}) - \mathbf{a}$. Each $\equiv$-equivalence class has infinitely many elements, and so there must be some $b' \in \mathcal{N}$ satisfying $b' \notin \mathbf{a}b$ and $\mathcal{N} \models (b \equiv b')$. Now, $g_{b,b'}$ fixes $\mathbf{a}$ pointwise by construction. Because $b \in \mathrm{dcl}_{\mathcal{N}}(\mathbf{a})$, the map $g_{b,b'}$ also fixes b. Hence $b = g_{b,b'}(b) = b'$, a contradiction. Therefore $\mathcal{N}$ has trivial definable closure, and so by Theorem 1.1, $\mathcal{N}$ admits an invariant measure.

For the converse, assume that (1) fails. Let $A \subseteq \mathcal{N}$ be a finite $\equiv$-equivalence class. By hypothesis, A has at least two elements. Hence for each $a \in A$, the set $A - \{a\}$ is nonempty, and so $a \in \mathrm{dcl}_{\mathcal{N}}(A - \{a\})$. Therefore $\mathcal{N}$ has nontrivial definable closure, and so (2) fails by Theorem 1.1. □

From Lemma 5.2 we can see that, in a sense, every countable infinite L-structure is "close to" one that admits an invariant measure, and also to infinitely many that do not. Specifically, if we take a countable infinite L-structure and "blow up" every element into n-many elements, where n is a cardinal satisfying $1 < n \le \aleph_0$, then the resulting structure admits an invariant measure if and only if $n = \aleph_0$.

Definition 5.3. Let L be a relational language, and let $\mathcal{M}$ be a countable infinite L-structure with underlying set M. Suppose n is a cardinal satisfying $1 \le n \le \aleph_0$. Define $\mathcal{M}_n^+$ to be the L^+-structure with underlying set $M \times n$ such that

$$\mathcal{M}_n^+ \models (a, j) \equiv (a', j') \qquad \text{if and only if} \qquad a = a',$$

for every $(a, j), (a', j') \in M \times n$, and

$$\mathcal{M}_n^+ \models R\big((a_1, j_1), \ldots, (a_k, j_k)\big) \qquad \text{if and only if} \qquad \mathcal{M} \models R(a_1, \ldots a_k),$$

for every relation $R \in L$ and every $(a_1, j_1), \ldots, (a_k, j_k) \in M \times n$, where k is the arity of R.

In the case when $\mathcal{M}$ is a graph, this construction is known as the *lexicographic product* of $\mathcal{M}$ with the empty graph on n vertices.

Note that $\equiv$ is an equivalence relation on $\mathcal{M}_n^+$ that respects L in $\mathcal{M}_n^+$. Hence we may take the quotient of $\mathcal{M}_n^+$ by $\equiv$ to obtain a structure isomorphic to $\mathcal{M}$. Moreover, every $\equiv$-equivalence class of $\mathcal{M}_n^+$ has n-many elements.

As an immediate corollary of Lemma 5.2 we have the following.

Corollary 5.4. *Let L be a relational language, let $\mathcal{M}$ be a countable infinite L-structure, and let n be a cardinal such that $1 < n \leq \aleph_0$. Then $\mathcal{M}_{\aleph_0}^+$ admits an invariant measure, while for $1 < n < \aleph_0$, the structure $\mathcal{M}_n^+$ does not admit an invariant measure.*

Note that this shows that every countable structure in a countable language is interpretable (see, e.g., [Mar02, Definition 1.3.9]) in a structure that admits an invariant measure, as well as in a structure that does not.

The *Scott rank* of a structure provides a measure of the complexity of the Scott sentence of the structure. (For details, see [Gao07].) Corollary 5.4 provides a method by which to build countable structures of arbitrary Scott rank that admit invariant measures, as well as ones that do not.

Corollary 5.5. *Let α be an arbitrary countably infinite ordinal. Define $\mathcal{T}^\alpha$ to be a countable linear order isomorphic to the well-order $(\alpha, \in)$ of height α. Then the structure $(\mathcal{T}^\alpha)_{\aleph_0}^+$ has Scott rank α and admits an invariant measure, whereas for $1 \leq n < \aleph_0$, the structure $(\mathcal{T}^\alpha)_n^+$ has Scott rank α and does not admit an invariant measure.*

Proof. For $1 \leq n \leq \aleph_0$, the structure $(\mathcal{T}^\alpha)_n^+$ has Scott rank α, as can be seen by a simple back-and-forth argument with $(\mathcal{T}^\beta)_n^+$ for $\beta < \alpha$. For $1 < n \leq \aleph_0$, the result follows by Corollary 5.4. When $n = 1$, the result follows from the fact that the least element of $(\mathcal{T}^\alpha)_1^+$ is in the definable closure of the empty set. ☐

5.2. **Classifications and other examples.**

Here we examine certain well-known countable infinite structures, and note whether or not they admit invariant measures. In some cases, such as the countable universal ultrahomogeneous partial order, our results provide the first demonstration that the structure admits an invariant measure. In several instances, the existence of invariant measures was known previously, though our results provide a simple way to check this. For example, it has been known nearly since its initial construction that the Rado graph $\mathscr{R}$ admits an invariant measure, and Petrov and Vershik [PV10] have more recently constructed invariant measures concentrated on the Henson graph $\mathscr{H}_3$ and on the other countable universal ultrahomogeneous K_n-free graphs.

Our results may be used to determine whether a particular structure admits an invariant measure either by checking directly whether it has trivial definable closure and applying Theorem 1.1, or, in the case of an ultrahomogeneous structure in a relational language, by determining whether its age has strong amalgamation and applying Corollary 1.3. It will be convenient sometimes to use the fact that a structure has trivial definable closure if and only if it has trivial algebraic closure, as mentioned in §2.4.

In the examples below, all graphs, directed graphs, and partial orders are considered to be structures in a language with a single binary relation symbol.

5.2.1. *Countable infinite ultrahomogeneous partial orders.*

These have been classified by Schmerl [Sch79] as follows.

- *(a)* The rationals, $(\mathbb{Q}, <)$.
- *(b)* The countable universal ultrahomogeneous partial order.
- *(c)* The countable infinite antichain.
- *(d)* The antichain of n copies of $\mathbb{Q}$ $(1 < n \leq \omega)$.
- *(e)* The $\mathbb{Q}$-chain of antichains, each of size n $(1 \leq n < \omega)$.
- *(f)* The $\mathbb{Q}$-chain of antichains, each of size ω.

All but *(e)* admit invariant measures: Their amalgamation problems can be solved by taking the transitive closure and, when needed, linearizing, and so their ages exhibit strong amalgamation. Example *(e)* clearly has nontrivial algebraic closure, and so does not admit an invariant measure.

5.2.2. *Countable infinite ultrahomogeneous permutations.*

Finite permutations have a standard interpretation as structures in a language with two binary relation symbols [Cam03] (see [Che11, §4.1] for a discussion). A permutation σ on $\{1, \dots, n\}$ can be viewed as two linear orders, $<$ and $\lhd$, on $\{1, \dots, n\}$, where $<$ is the usual order, and $\lhd$ is the permuted order, i.e., $\sigma(a) \lhd \sigma(b)$ if and only if $a < b$. One may extend this perspective on permutations to the infinite case, and consider structures that consist of a single infinite set endowed with two linear orders. Such structures describe relative finite rearrangements without completely determining a permutation on the infinite set. The countable infinite ultrahomogeneous permutations, so defined, have been classified by Cameron [Cam03] as follows.

- *(a)* The rationals, i.e., where each linear order has order type $\mathbb{Q}$ and they are equal to each other.
- *(b)* The reversed rationals, i.e., where each linear order has order type $\mathbb{Q}$ and the second is the reverse of the first.
- *(c)* Rational blocks of reversed rationals, i.e., where each linear order is the lexicographic product of $\mathbb{Q}$ with itself, and the second order is the reverse of the first *within* each block.
- *(d)* Reversed rational blocks of rationals, i.e., where each linear order is the lexicographic product of $\mathbb{Q}$ with itself, and the second order is the reverse of the first *between* the blocks.
- *(e)* The countable universal ultrahomogeneous permutation.

All five have trivial definable closure and hence admit invariant measures.

5.2.3. *Countable infinite ultrahomogeneous tournaments.*

A *tournament* is a structure consisting of a single irreflexive, binary relation, $\to$, such that for each pair a, b of distinct vertices, either $a \to b$ or $b \to a$, but not both. For example, any linear order is a tournament. The countable infinite ultrahomogeneous tournaments have been classified by Lachlan [Lac84] as follows.

(a) The rationals, $(\mathbb{Q}, <)$.
(b) The countable universal ultrahomogeneous tournament, T^∞.
(c) The *circular tournament* $S(2)$, also known as the *local order*, which consists of a countable dense subset of a circle where no two points are antipodal, with $x \to y$ if and only if the angle of xOy is less than π, where O is the center of the circle.

The ages of all three exhibit strong amalgamation (see [Che98, §2.1]).

5.2.4. *Countable infinite ultrahomogeneous directed graphs.*

A *directed graph* is a structure consisting of a single irreflexive, binary relation, $\to$, that is asymmetric, i.e., such that for each pair a, b of distinct vertices, $a \to b$ and $b \to a$ do not both hold. The countable infinite ultrahomogeneous directed graphs have been classified by Cherlin [Che98] (see also [Che87] for the imprimitive case). Macpherson [Mac11] describes the classification as follows (with some overlap between classes).

(a) The countable infinite ultrahomogeneous partial orders.
(b) The countable infinite ultrahomogeneous tournaments.
(c) Henson's countable infinite ultrahomogeneous directed graphs with forbidden sets of tournaments.
(d) The countable infinite ultrahomogeneous directed graph omitting I_n, the edgeless directed graph on n vertices $(1 < n < \omega)$.
(e) Four classes of directed graphs that are imprimitive, i.e., for which there is a nontrivial equivalence relation definable without parameters.
(f) Two exceptional directed graphs: a shuffled 3-tournament $S(3)$, defined analogously to the local order (defined above in 5.2.3(c)) with angle $2\pi/3$, and the *dense local partial order* $\mathcal{P}(3)$, a modification of the countable universal ultrahomogeneous partial order.

The structures in *(a)* and *(b)* are discussed above, in §5.2.1 and §5.2.3, respectively.

Henson [Hen72] described the class *(c)* of $2^{\aleph_0}$-many nonisomorphic countable infinite ultrahomogeneous directed graphs with forbidden sets of tournaments. The age of each has *free* amalgamation, i.e., its amalgamation problem can be solved by taking the disjoint union over the common substructure and adding no new relations. Free amalgamation implies strong amalgamation; hence on Henson's ultrahomogeneous directed graphs there are invariant measures.

The ages of the structures in *(d)* have strong amalgamation.

The first imprimitive class in *(e)* consists of the wreath products $T[I_n]$ and $I_n[T]$ where T is a countable infinite ultrahomogeneous tournament (as discussed above in §5.2.3) and $1 < n < \omega$. Each $T[I_n]$ has nontrivial definable closure because there is a definable equivalence relation, each class of which has n elements. Each $I_n[T]$ has trivial definable closure because it is the disjoint union of copies of an infinite tournament that has strong amalgamation.

The second imprimitive class in *(e)* consists of $\widehat{\mathbb{Q}}$ and $\widehat{T^\infty}$, modifications of the rationals and the countable universal ultrahomogeneous tournament, respectively, in which the algebraic closure of each point has size 2, namely itself and the unique other point to which it is not related. Hence neither directed graph has trivial definable closure.

The third imprimitive class in *(e)* consists of directed graphs $n * I_\infty$, for $1 < n \leq \omega$, which are universal subject to the constraint that non-relatedness is an equivalence relation with n classes. All such directed graphs have trivial definable closure.

The fourth imprimitive class in *(e)* consists of a *semigeneric* variant of $\omega * I_\infty$ with a parity constraint, which also has trivial definable closure.

The ages of $S(3)$ and $\mathcal{P}(3)$ exhibit strong amalgamation.

5.2.5. *Countable infinite ultrahomogeneous graphs.*

These have been classified by Lachlan and Woodrow [LW80] as follows.

- *(a)* The Rado graph $\mathscr{R}$.
- *(b)* The Henson graph $\mathscr{H}_3$ and the other countable universal ultrahomogeneous K_n-free graphs $(n > 3)$, and their complements.
- *(c)* Finite or countably infinite union of K_ω, and their complements.
- *(d)* Countably infinite union of K_n (for $1 < n < \omega$), and their complements.

The ages of the structures in *(a)* through *(c)* all have strong amalgamation; in fact, for the Rado graph, Henson's $\mathscr{H}_3$ and other K_n-free graphs, and the complement of K_ω, the amalgamation is free. Hence the structures in *(a)* through *(c)* all admit invariant measures. The structures in *(d)* clearly have nontrivial algebraic closure, and so do not admit invariant measures.

5.2.6. *Countable universal C-free graphs.*

Let C be a finite set of finite connected graphs. A graph $\mathcal{G}$ is said to be *C-free*, or to *forbid* C, when no member of C is isomorphic to a (graph-theoretic) subgraph of $\mathcal{G}$, i.e., when no member of C embeds as a weak substructure of $\mathcal{G}$. A countable infinite C-free graph $\mathcal{G}$ is said to be *universal* when every countable C-free graph is isomorphic to an induced subgraph of $\mathcal{G}$, i.e., embeds as a substructure of $\mathcal{G}$. When there is a universal such graph, there is one (up to isomorphism) that is distinguished by being *existentially complete*.

Only a limited number of examples are known of finite sets C of finite connected graphs for which a countable universal C-free graph exists (see the introduction to [CSS99] for a discussion). The best-known are when $C = \{K_n\}$, for $n \geq 3$; Henson's countable universal ultrahomogeneous K_n-free graph is universal for countable graphs that forbid $\{K_n\}$. We consider two other families here.

(a) The set C is homomorphism-closed, i.e., closed under maps that preserve edges but not necessarily non-edges. For example, take C to be the set of cycles of all odd lengths up to a fixed $2n + 1$. Cherlin, Shelah, and Shi [CSS99, Theorem 4] have shown that for a homomorphism-closed set C, an existentially complete countable universal C-free graph exists and has trivial algebraic closure. Hence these graphs admit invariant measures. Such graphs have also been considered in [HN14].

(b) The singleton set $C = \{K_m \dotplus K_n\}$ for some $m, n > 2$, where $K_m \dotplus K_n$ is the graph on $m + n - 1$ vertices consisting of complete graphs K_m and K_n joined at a single vertex. For example, $K_3 \dotplus K_3$ is the so-called *bowtie*. An existentially complete countable universal $(K_m \dotplus K_n)$-free graph exists if and only if $\min(m, n) = 3$ or 4, or $\min(m, n) = 5$ but $m \neq n$ ([Kom99], [CSS99], and [CT07]). Any such graph has nontrivial algebraic closure because, by existential completeness, it must contain a copy $\mathcal{K}$ of K_{m+n-2}, but for any vertex $v \in \mathcal{K}$, the algebraic closure of $\{v\}$ in the graph is all of $\mathcal{K}$.

5.2.7. *Trees and connected graphs with finite cut-sets.*

A tree is an acyclic connected graph. No tree can have trivial algebraic closure because there exists a unique finite path between any two distinct vertices of the tree. Similarly, no connected graph with a cut-vertex (a vertex whose removal disconnects the graph) can have trivial algebraic closure. More generally, if a connected graph contains a finite cut-set (a finite set whose removal disconnects the graph), then it cannot have trivial algebraic closure.

5.2.8. *Rational Urysohn space.*

A rational metric space is a metric space all of whose distances are rational. The class of all finite rational metric spaces, considered in the language with one relation symbol for each rational distance, is a Fraïssé class. Its Fraïssé limit is known as the *rational Urysohn space*, denoted $\mathbb{QU}$ (for details see [CV06]). The completion of $\mathbb{QU}$ is the Urysohn space, the universal ultrahomogeneous complete separable metric space.

The space $\mathbb{QU}$ admits an invariant measure, as can be seen from our results, since the class of finite rational metric spaces has strong amalgamation. Vershik, in [Ver02b] and [Ver04], has earlier constructed invariant measures concentrated on a collection of countable metric spaces whose completions are also Urysohn space. For a construction of several related invariant measures, see [AFNP16].

6. Applications and further observations

We conclude the paper with some observations and applications of our results. We describe, in §6.1, some of the theory of dense graph limits and its connections to our setting.

Our main theorem, Theorem 1.1, completely characterizes those single orbits of S_∞ on which an invariant measure can be concentrated. In §6.2, we ask which other Borel subsets of Str_L, consisting of multiple orbits, are such that some invariant measure is concentrated on them, and we make some observations based on our machinery.

Finally, in §6.3, we note a corollary of our result for sentences of $\mathcal{L}_{\omega_1,\omega}$ that have exactly one model (countable or otherwise).

6.1. **Invariant measures and dense graph limits.**

As remarked in the introduction, our constructions in the case of graphs can be viewed within the framework of the theory of dense graph limits. Here we describe this connection and some of its consequences.

6.1.1. *Invariant measures via graphons and W-random graphs.*

We now describe how invariant measures arise in the context of dense graph limits. We begin with some definitions from [LS06]; for more details, see also [DJ08], and [Lov12].

A **graphon** is defined to be a symmetric measurable function $W\colon [0,1]^2 \to [0,1]$. In what follows, we will take all graphons to be Borel measurable. Let L_G be the language of graphs, i.e., a language consisting of a single binary relation symbol E, representing the edges. Let T_G be the theory in the language L_G that says that E is symmetric and irreflexive. A graph may be considered to be an L_G-structure that satisfies T_G. An invariant measure on graphs is then precisely an invariant measure on Str_{L_G} that is concentrated on the set of models of T_G in Str_{L_G}.

Given a graphon W, the W**-random graph** $\mathbb{G}(\mathbb{N}, W)$ can be thought of as a random element of Str_{L_G}, defined as follows. Let $\{X_k\}_{k\in\mathbb{N}}$ be an independent sequence of random variables uniformly distributed on the unit interval. Then for $i, j \in \mathbb{N}$ with $i < j$, let $E(i, j)$ hold with independent probability $W(X_i, X_j)$; for each i, require that $E(i, i)$ not hold; and for each $i > j$, let $E(i, j)$ hold if and only if $E(j, i)$ does. For example, when W is a constant function p where $0 < p < 1$, then $\mathbb{G}(\mathbb{N}, W)$ is essentially the Erdős-Rényi graph $\mathbb{G}(\mathbb{N}, p)$, described in §1.1. Notice that for any graphon W, the distribution of $\mathbb{G}(\mathbb{N}, W)$ is an invariant measure on graphs.

Not only is the distribution of $\mathbb{G}(\mathbb{N}, W)$ invariant for an arbitrary graphon W, but so are the mixtures, i.e., convex combinations, of such distributions. Conversely, Aldous [Ald81] and Hoover [Hoo79] showed, in the context of exchangeable random arrays, that *every* invariant measure on graphs is such a mixture, thereby completely characterizing the invariant measures on graphs. This characterization has also arisen in the theory of dense graph limits; for details see [DJ08] and [Aus08a].

An analogous theory to that of graphons has been developed for other combinatorial structures such as partial orders [Jan11] and permutations [HKM+13]. The standard recipe described in [Aus08a] extends this machinery to the general case of countable relational languages of bounded arity. When L has bounded arity, our notion of Borel L-structure, from §3.1, can be viewed as a specialization of certain structures that occur in the standard recipe. In particular, any Borel L_G-structure that is a model of T_G corresponds to a graphon, as we will now see. Recall that because L_G is relational, every Borel L_G-structure is samplable.

6.1.2. *Borel L_G-structures and random-free graphons.*

Borel L_G-structures that are models of T_G (i.e., graphs) are closely related to a particular class of graphons. Here we describe this relationship and use it to deduce a corollary about W-random graphs whose distributions are concentrated on single countable graphs.

A graphon W is said to be **random-free** [Jan13, §10] if for a.e. $(x, y) \in [0, 1]^2$ we have $W(x, y) \in \{0, 1\}$. (See also the *simple arrays* of [Kal99] and $0-1$ valued graphons in [LS10].) When W is random-free, the W-random graph process amounts, in the language of [PV10], to "randomization in vertices" but not "randomization in edges".

We now describe a correspondence between Borel L_G-structures satisfying T_G and random-free graphons. Let α be an arbitrary Borel measurable bijection from the open interval $(0, 1)$ to $\mathbb{R}$, and let m_α be the distribution of $\alpha(U)$ where U is uniformly distributed on $[0, 1]$. Given a Borel L_G-structure $\mathcal{P}$ that satisfies T_G, define the random-free graphon $W_\mathcal{P}$ as follows. For $(x, y) \in (0, 1)^2$ let

$$W_\mathcal{P}(x, y) = 1 \qquad \text{if and only if} \qquad \mathcal{P} \models E\big(\alpha(x), \alpha(y)\big),$$

and for (x, y) on the boundary of $[0, 1]^2$ let $W_\mathcal{P}(x, y) = 0$. The distribution of $\mathbb{G}(\mathbb{N}, W_\mathcal{P})$ is precisely $\mu_{(\mathcal{P}, m_\alpha)}$, as defined in Definition 3.4. Conversely, given a graphon W that is Borel and random-free, one can easily build a Borel L_G-structure $\mathcal{P}_W$ satisfying T_G such that the distribution of $\mathbb{G}(\mathbb{N}, W)$ is $\mu_{(\mathcal{P}_W, m_\alpha)}$.

By Corollary 3.22, if a countable graph admits an invariant measure, then it admits one of the form $\mu_{(\mathcal{P}, m)}$, where $\mathcal{P}$ is a Borel L_G-structure. In particular, the corresponding random-free graphon $W_\mathcal{P}$ is such that the distribution of $\mathbb{G}(\mathbb{N}, W_\mathcal{P})$ is an invariant measure concentrated on the given graph. This leads to the following corollary.

Corollary 6.1. *Let $\mathcal{M}$ be a countable infinite graph. Suppose there is some graphon W such that the distribution of $\mathbb{G}(\mathbb{N}, W)$ is concentrated on $\mathcal{M}$. Then there is a random-free graphon W' such that the distribution of $\mathbb{G}(\mathbb{N}, W')$ is also concentrated on $\mathcal{M}$.*

Proof. The distribution of $\mathbb{G}(\mathbb{N}, W)$ is an invariant measure concentrated on $\mathcal{M}$. Therefore by Theorem 1.1, the graph $\mathcal{M}$ must have trivial definable closure. By Corollary 3.22, there is a Borel L_G-structure $\mathcal{P}$ such that $\mu_{(\mathcal{P},m)}$ is concentrated on $\mathcal{M}$ whenever m is a continuous nondegenerate probability measure on $\mathbb{R}$. As above, let $\alpha\colon (0,1) \to \mathbb{R}$ be a Borel bijection and let $W_{\mathcal{P}}$ be the random-free graphon induced by the given correspondence. Then the distribution of $\mathbb{G}(\mathbb{N}, W_{\mathcal{P}})$ is $\mu_{(\mathcal{P},m_\alpha)}$, and hence is also concentrated on $\mathcal{M}$. □

In fact, for an arbitrary countable relational language L, our procedure for sampling from a Borel L-structure essentially arises in [Aus08a] as a standard recipe in which all but the first "ingredient" are deterministic maps. In this setting, one can prove an analogue of Corollary 6.1 for arbitrary countable infinite L-structures.

The best-known graphons W for which $\mathbb{G}(\mathbb{N}, W)$ is isomorphic to the Rado graph are the constant functions $W \equiv p$ for $0 < p < 1$, i.e., those given by the Erdős–Rényi construction. However, these are not the only such graphons. Petrov and Vershik [PV10] were the first to describe invariant measures concentrated on the Rado graph that correspond to random-free graphons. Figure 2 is a visualization of a random-free graphon W, built essentially by the methods of [PV10] and the present paper, for which $\mathbb{G}(\mathbb{N}, W)$ is a.s. isomorphic to the Rado graph,

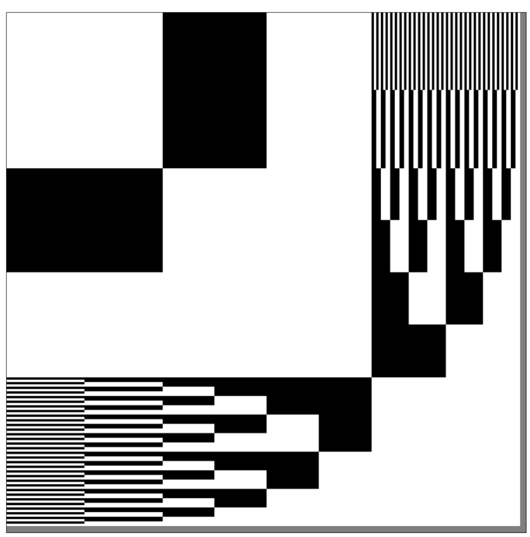

Figure 2. An illustration of a random-free graphon W such that $\mathbb{G}(\mathbb{N}, W)$ is a.s. isomorphic to the Rado graph. (The thin grey strips on the right and bottom represent regions not drawn in detail — not values of the graphon between 0 and 1.)

6.2. **Multiple isomorphism classes.**

In this paper, we have focused on the problem of identifying those countable infinite L-structures $\mathcal{M}$ such that some invariant measure is concentrated on the isomorphism class of $\mathcal{M}$ in Str_L, i.e., on the orbit under the logic action of any structure in Str_L isomorphic to $\mathcal{M}$. But it is natural to investigate those larger subsets of Str_L, consisting of the union of multiple orbits, on which an invariant measure may be concentrated. For example, Austin [Aus08a, Question 3.27] asks for a characterization of first-order theories T such that any invariant measure concentrated on the set of models of T in Str_L must come from a standard recipe having a property akin to being random-free.

There are clearly invariant measures on Str_L that are not concentrated on any single structure, as can be seen by taking mixtures of invariant measures concentrated on different structures. But if there is a countable set of structures on which such an invariant measure is concentrated, then we can see by conditioning that there must be some invariant measure concentrated on one of these structures.

Lemma 6.2. *Let L be a countable language, and let T be a theory of $\mathcal{L}_{\omega_1,\omega}(L)$ that has at most countably many countable infinite models (up to isomorphism). Suppose μ_T is an invariant measure on Str_L that is concentrated on the set of models of T in Str_L. Then there is a countable model $\mathcal{M}$ of T such that some invariant measure $\mu_{\mathcal{M}}$ is concentrated on the isomorphism class of $\mathcal{M}$ in Str_L.*

Proof. Because μ_T is countably additive and T has only countably many countable infinite models, there must be some countable infinite structure $\mathcal{M} \models T$ such that its isomorphism class $\widetilde{\mathcal{M}} := \{\mathcal{N} \in \mathrm{Str}_L : \mathcal{N} \cong \mathcal{M}\}$ has positive μ_T-measure. Recall that $\widetilde{\mathcal{M}}$ is a Borel set. Let $\mu_{\mathcal{M}}$ be μ_T conditioned on this positive measure set, i.e.,

$$\mu_{\mathcal{M}}(A) := \mu_T\big(A \,|\, \widetilde{\mathcal{M}}\big) = \mu_T\big(A \cap \widetilde{\mathcal{M}}\big)/\mu_T\big(\widetilde{\mathcal{M}}\big)$$

for every Borel set $A \subseteq \mathrm{Str}_L$. Then $\mu_{\mathcal{M}}$ is a probability measure on Str_L concentrated on the isomorphism class of $\mathcal{M}$.

Moreover, $\mu_{\mathcal{M}}$ is invariant, as we now show. Suppose $g \in S_\infty$, and let A be an arbitrary Borel subset of Str_L. Because $\widetilde{\mathcal{M}}$ is an S_∞-invariant subset of Str_L, we have

$$\mu_T\big(g(A) \cap \widetilde{\mathcal{M}}\big) = \mu_T\big(g(A) \cap g(\widetilde{\mathcal{M}})\big),$$

and because μ_T is an invariant measure, we have

$$\mu_T\big(g(A \cap \widetilde{\mathcal{M}})\big) = \mu_T\big(A \cap \widetilde{\mathcal{M}}\big).$$

Since $g(A) \cap g\big(\widetilde{\mathcal{M}}\big) = g\big(A \cap \widetilde{\mathcal{M}}\big)$, we have $\mu_{\mathcal{M}}(g(A)) = \mu_{\mathcal{M}}(A)$, as desired. ☐

One may ask, more specifically, given a samplable Borel L-structure $\mathcal{P}$ and a continuous nondegenerate probability measure m on $\mathbb{R}$, the minimum number of isomorphism classes on whose union the measure $\mu_{(\mathcal{P},m)}$ is concentrated. When $\mathcal{P}$ strongly witnesses a theory T of $\mathcal{L}_{\omega_1,\omega}(L)$ having just one countable infinite model up to isomorphism, then there is just one isomorphism class by design. However, if $\mathcal{P}$ strongly witnesses a pithy Π_2 theory T of $\mathcal{L}_{\omega_1,\omega}(L)$ that has nonisomorphic countable infinite models, then the situation is more complicated. In this case, still $\mathcal{P} \models T$ by Lemma 3.9, but the induced invariant measure might be concentrated on a union of multiple isomorphism classes of models of T, but not on any single such class.

However, as we state in Corollary 6.3, this is not possible if the measure is concentrated on a *countable* union of isomorphism classes. By countable additivity, any invariant measure on Str_L that is concentrated on a union of countably many isomorphism classes, but not on a single class, must be non-ergodic. But every measure of the form $\mu_{(\mathcal{P},m)}$ is ergodic, as we now explain.

The ergodic invariant measures on graphs are precisely those induced by sampling from a single graphon, rather than a mixture of such (see [DJ08, Corollary 5.4] and [LS12, Proposition 3.6]). Aldous had earlier characterized the ergodic invariant measures on hypergraphs in a similar way (see [Ald81, Proposition 3.3] or [Kal05, Lemma 7.35]). This characterization has a generalization to countable infinite languages, e.g., via the setting of Kallenberg's extension of the Aldous–Hoover theorem ([Kal05, Lemma 7.22] and [Kal05, Lemma 7.28]). In particular, it can be shown that measures of the form $\mu_{(\mathcal{P},m)}$ are ergodic, and so the following corollary holds.

Corollary 6.3. *Let $\mathcal{P}$ be a samplable Borel L-structure, and suppose m is a continuous nondegenerate probability measure on $\mathbb{R}$. If $\mu_{(\mathcal{P},m)}$ is concentrated on some countable union of isomorphism classes in Str_L, then in fact $\mu_{(\mathcal{P},m)}$ is concentrated on a single isomorphism class.*

In other words, for any samplable Borel L-structure $\mathcal{P}$, the measure $\mu_{(\mathcal{P},m)}$, as defined in §3.1, is concentrated on either one or uncountably many isomorphism classes. For an investigation of some circumstances with continuum-many isomorphism classes, see [AFNP16].

6.3. Continuum-sized models of Scott sentences.

We conclude with a somewhat unexpected corollary of the machinery that we have developed. A countable structure $\mathcal{M}$ is said to be *absolutely characterizable* when its Scott sentence $\sigma_{\mathcal{M}}$ has no uncountable models, and hence characterizes $\mathcal{M}$ up to isomorphism among all structures, not just among countable structures (see [KK04, §1.3]). Our results imply that there is no invariant measure concentrated on such a structure.

Corollary 6.4. *Let L be a countable language and let $\mathcal{M} \in \mathrm{Str}_L$. Suppose that $\sigma_{\mathcal{M}}$, the Scott sentence of $\mathcal{M}$, has no continuum-sized models. Then there is no invariant measure on Str_L that is concentrated on the isomorphism class of $\mathcal{M}$.*

Proof. Suppose there exists an invariant measure concentrated on $\mathcal{M}$. Then by Theorem 4.1, $\mathcal{M}$ has trivial definable closure. Let $\overline{\mathcal{M}}$ be the canonical structure of $\mathcal{M}$ and $L_{\overline{\mathcal{M}}}$ be the canonical language. By Lemmas 2.13 and 2.15, $\overline{\mathcal{M}}$ also has trivial definable closure.

By Proposition 2.17, there is a pithy Π_2 $\mathcal{L}_{\omega_1,\omega}(L_{\overline{\mathcal{M}}})$-theory $T_{\overline{\mathcal{M}}}$ all of whose countable models are isomorphic to $\overline{\mathcal{M}}$. Hence by Theorem 3.19 there exists a (continuum-sized) Borel $L_{\overline{\mathcal{M}}}$-structure $\mathcal{Q}$ strongly witnessing $T_{\overline{\mathcal{M}}}$. But then $\mathcal{Q} \models T_{\overline{\mathcal{M}}}$, by Lemma 3.9. By Lemma 2.12, using the interdefinition given in Lemma 2.13 between $\mathcal{M}$ and $\overline{\mathcal{M}}$, there is a (continuum-sized) L-structure interdefinable with $\mathcal{Q}$, which has the same $\mathcal{L}_{\omega_1,\omega}(L)$-theory as $\mathcal{M}$, and which hence satisfies $\sigma_{\mathcal{M}}$. □

Finally, this shows that if the Scott sentence $\sigma_{\mathcal{M}}$ of a countable infinite structure $\mathcal{M}$ has no continuum-sized models (e.g., if $\mathcal{M}$ is absolutely characterizable), then $\mathcal{M}$ must have nontrivial definable closure.

Acknowledgements

We would like to thank Gregory Cherlin for pointing us to [PV10], and Lionel Levine, Gábor Lippner, Ben Rossman, Gerald Sacks, and Peter Winkler for helpful initial discussions. We are grateful to Anatoly Vershik and Jaroslav Nešetřil for generously sharing their insights over several conversations. Thanks to John Truss for inviting C. F. and R. P. to the Workshop on Homogeneous Structures at the University of Leeds in 2011, and to Jaroslav Nešetřil for inviting C. F. to the Workshop on Graph and Hypergraph Limits at the American Institute of Mathematics in 2011; both provided stimulating environments in which to think about this work. We also thank Dan Roy for helping to improve the proofs of certain results, and Willem Fouché, Alex Kruckman, M. Malliaris, Peter Orbanz, Max Weiss, and Carol Wood for comments on a draft. We thank Alexander Kechris for sharing with us the argument of Kechris–Marks incorporated into Theorem 4.1 and the observation Corollary 4.2. Finally, we would like to thank the anonymous referees for several comments which have significantly improved the presentation of the paper.

C. F. was partially supported by NSF grants DMS-0901020 and DMS-0800198 and ARO grant W911NF-13-1-0212, and this research was partially done while he was a visiting fellow at the Isaac Newton Institute for the Mathematical Sciences in the program *Semantics and Syntax*. His work on this publication was also made possible through the support of grants from the John Templeton Foundation and Google. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

References

[Ack16] N. L. Ackerman, *A characterization of quasitrivial n-semigroups*, Algebra Universalis, to appear (2016).

[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.

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

[Aus08a] T. Austin, *On exchangeable random variables and the statistics of large graphs and hypergraphs*, Probab. Surv. **5** (2008), 80–145.

[Aus08b] ———, *Razborov flag algebras as algebras of measurable functions*, ArXiv e-print 0801.1538 (2008).

[AZ86] G. Ahlbrandt and M. Ziegler, *Quasi-finitely axiomatizable totally categorical theories*, Ann. Pure Appl. Logic **30** (1986), no. 1, 63–82.

[Bar75] J. Barwise, *Admissible sets and structures: An approach to definability theory*, Perspectives in Mathematical Logic, Springer-Verlag, Berlin, 1975.

[BK96] H. Becker and A. S. Kechris, *The descriptive set theory of Polish group actions*, London Math. Soc. Lecture Note Ser., vol. 232, Cambridge Univ. Press, Cambridge, 1996.

[Cam03] P. J. Cameron, *Homogeneous permutations*, Electron. J. Combin. **9** (2002/03), no. 2, Research paper 2.

[Cam90] ———, *Oligomorphic permutation groups*, London Math. Soc. Lecture Note Ser., vol. 152, Cambridge Univ. Press, Cambridge, 1990.

[Che87] G. Cherlin, *Homogeneous directed graphs. The imprimitive case*, Logic Colloquium '85 (Orsay, 1985), Stud. Logic Found. Math., vol. 122, North-Holland, Amsterdam, 1987, pp. 67–88.

[Che98] ———, *The classification of countable homogeneous directed graphs and countable homogeneous n-tournaments*, vol. 131, Mem. Amer. Math. Soc., no. 621, Amer. Math. Soc., Providence, RI, 1998.

[Che11] ______, *Two problems on homogeneous structures, revisited*, Model-theoretic Methods in Finite Combinatorics, Contemp. Math., vol. 558, Amer. Math. Soc., Providence, RI, 2011, pp. 319–415.

[CSS99] G. Cherlin, S. Shelah, and N. Shi, *Universal graphs with forbidden subgraphs and algebraic closure*, Adv. in Appl. Math. **22** (1999), no. 4, 454–491.

[CT07] G. Cherlin and L. Tallgren, *Universal graphs with a forbidden near-path or 2-bouquet*, J. Graph Theory **56** (2007), no. 1, 41–63.

[CV06] P. J. Cameron and A. M. Vershik, *Some isometry groups of the Urysohn space*, Ann. Pure Appl. Logic **143** (2006), no. 1-3, 70–78.

[DJ08] P. Diaconis and S. Janson, *Graph limits and exchangeable random graphs*, Rend. Mat. Appl. (7) **28** (2008), no. 1, 33–61.

[DK03] M. Droste and D. Kuske, *On random relational structures*, J. Combin. Theory Ser. A **102** (2003), no. 2, 241–254.

[DM12] I. Dolinka and D. Mašulović, *Properties of the automorphism group and a probabilistic construction of a class of countable labeled structures*, J. Combin. Theory Ser. A **119** (2012), no. 5, 1014–1030.

[EKR76] P. Erdős, D. J. Kleitman, and B. L. Rothschild, *Asymptotic enumeration of K_n-free graphs*, Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973), Tomo II, Atti dei Convegni Lincei, vol. 17, Accad. Naz. Lincei, Rome, 1976, pp. 19–27.

[ER59] P. Erdős and A. Rényi, *On random graphs. I*, Publ. Math. Debrecen **6** (1959), 290–297.

[Gai64] H. Gaifman, *Concerning measures in first order calculi*, Israel J. Math. **2** (1964), 1–18.

[Gao07] S. Gao, *Complexity ranks of countable models*, Notre Dame J. Formal Logic **48** (2007), no. 1, 33–48.

[Gil59] E. N. Gilbert, *Random graphs*, Ann. Math. Statist. **30** (1959), 1141–1144.

[Hen71] C. W. Henson, *A family of countable homogeneous graphs*, Pacific J. Math. **38** (1971), 69–83.

[Hen72] ______, *Countable homogeneous relational structures and $\aleph_0$-categorical theories*, J. Symbolic Logic **37** (1972), 494–500.

[HKM$^+$13] C. Hoppen, Y. Kohayakawa, C. G. Moreira, B. Ráth, and R. Menezes Sampaio, *Limits of permutation sequences*, J. Combin. Theory Ser. B **103** (2013), no. 1, 93–113.

[HN14] J. Hubička and J. Nešetřil, *Homomorphism and embedding universal structures for restricted classes*, ArXiv e-print 0909.4939 (2014).

[Hod93] W. Hodges, *Model theory*, Encyclopedia of Mathematics and its Applications, vol. 42, Cambridge University Press, Cambridge, 1993.

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

[Jan11] S. Janson, *Poset limits and exchangeable random posets*, Combinatorica **31** (2011), no. 5, 529–563.

[Jan13] ______, *Graphons, cut norm and distance, couplings and rearrangements*, New York J. Math. NYJM Monographs, vol. 4, State Univ. of New York, Univ. at Albany, 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.

[Kal02] ______, *Foundations of modern probability*, 2nd ed., Probability and its Applications, Springer-Verlag, New York, 2002.

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

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

[Kei71] H. J. Keisler, *Model theory for infinitary logic: Logic with countable conjunctions and finite quantifiers*, Studies in Logic and the Foundations of Mathematics, vol. 62, North-Holland Publishing Co., Amsterdam, 1971.

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

[Kom99] P. Komjáth, *Some remarks on universal graphs*, Discrete Math. **199** (1999), no. 1-3, 259–265.

[KPR87] P. G. Kolaitis, H. J. Prömel, and B. L. Rothschild, *K_{l+1}-free graphs: asymptotic structure and a* 0–1 *law*, Trans. Amer. Math. Soc. **303** (1987), no. 2, 637–671.

[Kra69] P. H. Krauss, *Representation of symmetric probability models*, J. Symbolic Logic **34** (1969), 183–193.

[Lac84] A. H. Lachlan, *Countable homogeneous tournaments*, Trans. Amer. Math. Soc. **284** (1984), no. 2, 431–461.

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

[LS06] L. Lovász and B. Szegedy, *Limits of dense graph sequences*, J. Combin. Theory Ser. B **96** (2006), no. 6, 933–957.

[LS10] ———, *Regularity partitions and the topology of graphons*, An Irregular Mind, Bolyai Soc. Math. Stud., vol. 21, János Bolyai Math. Soc., Budapest, 2010, pp. 415–446.

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

[LW80] A. H. Lachlan and R. E. Woodrow, *Countable ultrahomogeneous undirected graphs*, Trans. Amer. Math. Soc. **262** (1980), no. 1, 51–94.

[Mac11] D. Macpherson, *A survey of homogeneous structures*, Discrete Math. **311** (2011), no. 15, 1599–1634.

[Mar02] D. Marker, *Model theory*, Graduate Texts in Math., vol. 217, Springer-Verlag, New York, 2002.

[MN13] A. Montalbán and A. Nies, *Borel structures: a brief survey*, Effective Mathematics of the Uncountable, Lect. Notes Log., vol. 41, Assoc. Symbol. Logic, La Jolla, CA, 2013, pp. 124–134.

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

[Rad64] R. Rado, *Universal graphs and universal functions*, Acta Arith. **9** (1964), 331–340.

[Raz07] A. A. Razborov, *Flag algebras*, J. Symbolic Logic **72** (2007), no. 4, 1239–1282.

[Sch79] J. H. Schmerl, *Countable homogeneous partially ordered sets*, Algebra Universalis **9** (1979), no. 3, 317–321.

[SK66] D. Scott and P. Krauss, *Assigning probabilities to logical formulas*, Aspects of inductive logic (J. Hintikka and P. Suppes, eds.), Studies in Logic and the Foundations of Mathematics, North Holland, Amsterdam, 1966, pp. 219–259.

[Ste85] C. I. Steinhorn, *Borel structures and measure and category logics*, Model-theoretic Logics, Perspect. Math. Logic, Springer, New York, 1985, pp. 579–596.

[Usv08] A. Usvyatsov, *Generic separable metric structures*, Topology Appl. **155** (2008), no. 14, 1607–1617.

[Ver02a] A. M. Vershik, *Classification of measurable functions of several arguments, and invariantly distributed random matrices*, Funktsional. Anal. i Prilozhen. **36** (2002), no. 2, 12–27, 95.

[Ver02b] ———, *A random metric space is a Urysohn space*, Dokl. Akad. Nauk **387** (2002), no. 6, 733–736.

[Ver04] ———, *Random metric spaces and universality*, Uspekhi Mat. Nauk **59** (2004), no. 2(356), 65–104.

[Wil91] D. Williams, *Probability with martingales*, Cambridge Univ. Press, Cambridge, 1991.

Department of Mathematics, Harvard University, One Oxford Street, Cambridge, MA 02138

E-mail address: nate@math.harvard.edu

Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 32 Vassar Street, Cambridge, MA 02139

E-mail address: freer@mit.edu

Franklin W. Olin College of Engineering, 1000 Olin Way, Needham, MA 02492

E-mail address: rehana.patel@olin.edu