Invariant measures via inverse limits of finite structures

Nathanael Ackerman[a], Cameron Freer[b,*], Jaroslav Nešetřil[c], Rehana Patel[d]

[a] *Department of Mathematics, Harvard University, One Oxford Street, Cambridge, MA 02138, USA*
[b] *Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 77 Massachusetts Ave. 32-G480, Cambridge, MA 02139, USA*
[c] *Department of Applied Mathematics and Institute of Theoretical Computer Science (IUUK and ITI), Charles University, Malostranské nám.25, 11800 Praha 1, Czech Republic*
[d] *Franklin W. Olin College of Engineering, Olin Way, Needham, MA 02492, USA*

Abstract

Building on recent results regarding symmetric probabilistic constructions of countable structures, we provide a method for constructing probability measures, concentrated on certain classes of countably infinite structures, that are invariant under all permutations of the underlying set that fix all constants. These measures are constructed from inverse limits of measures on certain finite structures. We use this construction to obtain invariant probability measures concentrated on the classes of countable models of certain first-order theories, including measures that do not assign positive measure to the isomorphism class of any single model. We also characterize those transitive Borel G-spaces admitting a G-invariant probability measure, when G is an arbitrary countable product of symmetric groups on a countable set.

1. Introduction

Symmetric probabilistic constructions of mathematical structures have a long history, dating back to the countable random graph model of Erdős-Rényi [1], a construction that with probability 1 yields (up to isomorphism) the Rado graph, i.e., the countable universal ultrahomogeneous graph. In this paper, we build on recent developments that have extended the range of such constructions. In particular, we consider when a symmetric probabilistic construction can produce many different countable structures, with no isomorphism class occurring with positive probability. We also consider probabilistic constructions with respect to various notions of partial symmetry.

One natural notion of a symmetric probabilistic construction is via an *invariant measure* — namely, a probability measure on a class of countably infinite structures that is invariant under all permutations of the underlying set of elements. When such an invariant measure assigns probability 1 to a given class of structures (as the Erdős-Rényi construction does to the isomorphism class of the Rado graph), we say that it is *concentrated* on such structures, and that the given class *admits* an invariant measure.

For several decades, most known examples of such invariant measures were variants of the Erdős-Rényi random graph, for instance, an analogous construction that

*Corresponding author. 77 Massachusetts Ave., 32-G480, Cambridge, MA 02139, USA, phone: +1-716-713-8228

Email addresses: `nate@math.harvard.edu` (Nathanael Ackerman), `freer@mit.edu` (Cameron Freer), `nesetril@kam.ms.mff.cuni.cz` (Jaroslav Nešetřil), `rehana.patel@olin.edu` (Rehana Patel)

produces the countable universal bipartite graph. In recent years, a number of other important classes of structures have been shown to admit invariant measures, most notably the collection of countable metric spaces whose completion is Urysohn space, by Vershik [2], [3], and Henson's universal ultrahomogeneous K_n-free graphs by Petrov and Vershik [4]. Both constructions are considerably more complicated than the Erdős-Rényi construction. By extending the methods of [4], Ackerman, Freer, and Patel [5] have completely characterized those countable structures in a countable language whose isomorphism class admits an invariant measure.

In the present paper we extend the construction of [5]. Our new construction is more streamlined than the one in [5], and also broader in its consequences. Both constructions involve building continuum-sized structures from which invariant measures are obtained by sampling, but the one in [5] produces an explicit structure with underlying set the real numbers, necessitating various book-keeping devices, which we avoid here.

As a first application of the present more general construction, we describe certain first-order theories having the property that there is an invariant probability measure that is concentrated on the class of models of the theory but that assigns measure 0 to the isomorphism class of each particular model. We thereby obtain new examples of classes of structures admitting invariant measures, and new examples of invariant measures concentrated on collections of structures that were previously known to admit invariant measures.

Towards our second application, we consider measures that are invariant under the action of certain subgroups of the full permutation group S_∞ on the underlying set. Note that any random construction of a countably infinite structure with constants faces a fundamental obstacle to having an S_∞-invariant distribution, as described in [5]. Namely, if the distribution were S_∞-invariant, then the probability that any given constant symbol in the language is interpreted as a particular element would have to be the same as for any other element, leading to a contradiction, as a countably infinite set of identical reals cannot sum to 1. In other words, if a structure admits an S_∞-invariant measure, then it cannot be in a language having constant symbols. Furthermore, if a measure concentrated on the isomorphism class of the structure is invariant under a given permutation, then that permutation must fix all elements that interpret constant symbols.

With that obstacle in mind, we may ask, more generally, which structures admit measures that are invariant under all permutations of the underlying set of the structure and that fix the restriction of the structure to a particular sublanguage. We answer this question in the case of a unary sublanguage, i.e., where the sublanguage consists entirely of unary relations. By results in descriptive set theory, this is equivalent to describing all those transitive Borel G-spaces admitting a G-invariant probability measure when G is a countable product of symmetric groups on a countable (finite or infinite) set. This constitutes the second application of our construction.

In the special case of undirected graphs, our methods for producing invariant measures can be viewed as constructing dense graph limits, in the sense of Lovász and Szegedy [6] and others; for details, see [7]. In fact, by results of Aldous [8], Hoover [9], Kallenberg [10], and Vershik [11] in work on the probability theory of exchangeable arrays, an invariant measure on graphs is necessarily the distribution of a particular sampling procedure from *some* continuum-sized limit structure. For more details on this connection, see Diaconis and Janson [12] and Austin [13].

Our work also has connections to a recent study of Borel models of size continuum by Baldwin, Laskowski, and Shelah [14], building on work of Shelah [15, Theorem VII.3.7].

Their continuum-sized structures, like ours, are constructed from inverse limits; however, our methods differ from theirs in several respects and, unlike [14], our focus is on the consequences of these constructions for invariant measures.

1.1. Outline of the paper

In Section 2, we provide preliminaries for our constructions, including definitions and basic results from the model theory of infinitary logic and from descriptive set theory.

We then pause, in Section 3, to provide a toy construction, for graphs, that will motivate the more technical aspects of our main construction.

In Section 4, we present our main technical construction, in which we build a special kind of continuum-sized structure from inverse limits.

In the following sections, we provide two applications of this main construction. First, in Section 5, we use it to provide new constructions of invariant probability measures concentrated on the class of models of certain first-order theories, but assigning positive measure to no single isomorphism class.

Second, in Section 6, we use the main construction to characterize those structures that are invariant under automorphism groups that fix the restrictions of the structures to unary sublanguages. As noted, this amounts to characterizing those transitive Borel G-spaces that admit a G-invariant probability measure, when G is a countable product of symmetric groups on a countable (finite or infinite) set.

2. Preliminaries

In this section, we describe some notation, and introduce several basic notions regarding infinitary logic, transitive G-spaces, and model-theoretic structures and their automorphisms that we will use throughout the paper.

The set $\mathbb{N}^{<\omega}$ is defined to be the collection of finite sequences of natural numbers. For $x, y \in \mathbb{N}^{<\omega}$ we write $x \preceq y$ when x is an initial segment of y. The set $\mathbb{N}^{\omega}$ is the collection of countably infinite sequences of natural numbers. For $x \in \mathbb{N}^{\omega}$, we write $x|_n$ to denote the length-n initial segment of x in $\mathbb{N}^n$, and similarly for elements of $\mathbb{N}^{<\omega}$ of length at least n.

Suppose $j \in \mathbb{N}$. For $x_0, \ldots, x_j, y_0, \ldots, y_j \in \mathbb{N}^{<\omega}$, we write

$$(x_0, \ldots, x_j) \sqsubseteq (y_0, \ldots, y_j)$$

when $x_i \preceq y_i$ for $0 \leq i \leq j$.

We write $a^\wedge b$ to denote the concatenation of $a, b \in \mathbb{N}^{<\omega}$, though we often omit the symbol $^\wedge$ when concatenating explicit sequences. Occasionally we will use exponential notation for repeated numerals; e.g., $0^4 2^2$ denotes $000022 \in \mathbb{N}^{<\omega}$. Define the **projection function**

$$\pi \colon \mathbb{N}^{<\omega} \to \mathbb{N}^{<\omega}$$

by

$$\pi(a^\wedge b) = a$$

when $a \in \mathbb{N}^{<\omega}$ and $b \in \mathbb{N}$, and

$$\pi(\langle\,\rangle) = \langle\,\rangle,$$

where $\langle\,\rangle$ denotes the empty string. Write the composition of projection with itself as $\pi^2 := \pi \circ \pi$. We will use this notation in §4.2.

Define $\mathbb{R}_+ := \{x \in \mathbb{R} : x > 0\}$ and $\mathbb{Q}_{\geq 0} := \{x \in \mathbb{Q} : x \geq 0\}$.

A probability measure on $\mathbb{R}$ is said to be **non-degenerate** when every non-empty open set has positive measure and **atomless** when every singleton has measure 0.

We say that a probability measure μ on an arbitrary measure space S is **concentrated** on a measurable set $X \subseteq S$ when $\mu(X) = 1$. Given a measurable action of a group G on S, we say that μ is **G-invariant** if $\mu(X) = \mu(g \cdot X)$ for every $g \in G$ and measurable $X \subseteq S$.

2.1. Model theory of infinitary logic

We now briefly recall notation for finitary and infinitary formulas. For more details on such formulas and on the corresponding notion of satisfiability (denoted by $\models$), see [16] and [17, §1.1]. Throughout this paper, L will be a countable language, i.e., a countable collection of relation, constant, and function symbols. Fix an implicit set of countably infinitely many variables. Then $\mathcal{L}_{\omega,\omega}(L)$ is the set of all (finitary) first-order formulas (in that set of variables) with relation, constant, and function symbols from L. The set $\mathcal{L}_{\omega_1,\omega}(L)$ of infinitary L-formulas is the smallest set containing $\mathcal{L}_{\omega,\omega}(L)$ and closed under countable conjunctions, existential quantification, and negation, and such that each formula has only finitely many free variables. In particular, $\mathcal{L}_{\omega_1,\omega}(L)$ is closed under taking subformulas. A sentence is a formula having no free variables, and a theory is an arbitrary collection of sentences.

Let $k \in \mathbb{N}$ and let $x_1, \ldots, x_k$ be distinct variables. A (complete) **quantifier-free L-type** q with free variables $x_1, \ldots, x_k$ is a countable collection of quantifier-free formulas of $\mathcal{L}_{\omega_1,\omega}(L)$ whose set of free variables is contained in $\{x_1, \ldots, x_k\}$, and such that for any quantifier-free $\mathcal{L}_{\omega_1,\omega}(L)$-formula ψ whose free variables are among $x_1, \ldots, x_k$, either

$$\models (\forall x_1, \ldots, x_k)\Big(\bigwedge_{\varphi \in q} \varphi \to \psi\Big) \qquad \text{or} \qquad \models (\forall x_1, \ldots, x_k)\Big(\bigwedge_{\varphi \in q} \varphi \to \neg\psi\Big).$$

Note that any collection q of formulas which has this property with respect to all *atomic* formulas $\psi \in \mathcal{L}_{\omega,\omega}(L)$ is already a complete quantifier-free L-type.

Note that we will consider quantifier-free types to entail a fixed ordering of their free variables. This will be important because for a quantifier-free type q with k-many free variables, and a set X of size k with a specified ordering $<$, we will sometimes write $q(X)$ to represent the statement that $q(\ell_1, \ldots, \ell_k)$ holds, where $\ell_1 < \cdots < \ell_k$ are the elements of X.

We say that a quantifier-free type with free variables $x_1, \ldots, x_k$ is **non-constant** when it implies that none of $x_1, \ldots, x_k$ instantiates a constant symbol, and is **non-redundant** when it implies

$$\bigwedge_{1 \leq i < j \leq k} (x_i \neq x_j).$$

Suppose L_0 is a sublanguage of L, i.e., each of the sets of relation, constant, and function symbols of L_0 is a subset of the corresponding set for L. Then the **restriction** $q|_{L_0}$ of a quantifier-free L-type to L_0 is defined to be set of atomic L_0-formulas and their negations that are implied by $\bigwedge_{\varphi \in q} \varphi$.

An L-theory T is **quantifier-free complete** when it is consistent and for every quantifier-free L-sentence φ, exactly one of $T \models \varphi$ or $T \models \neg\varphi$ holds.

We will later make use of the notion of a **Scott sentence**: a sentence of $\mathcal{L}_{\omega_1,\omega}(L)$ which characterizes a given countable structure up to isomorphism among other countable L-structures. For more details, see [16, Corollary VII.6.9]. We will also use the notion of an admissible set, and in particular the admissible set HF of hereditarily finite sets; again see [16].

For a structure $\mathcal{M}$ with underlying set M, a natural number $k \in \mathbb{N}$, and a k-tuple $\overline{a} = (a_1, \dots, a_k) \in M^k$, we will sometimes abuse notation and write either $\overline{a} \in M$ or $\overline{a} \in \mathcal{M}$ to mean that $a_1, \dots, a_k \in M$. We will also sometimes write $a_1 \cdots a_k$ to denote such a tuple.

Suppose $\mathcal{M}$ is an L-structure. When U is a relation symbol in L, we write $U^{\mathcal{M}}$ to denote the set of tuples $\overline{a} \in \mathcal{M}$ such that $\mathcal{M} \models U(\overline{a})$. Similarly, we write $c^{\mathcal{M}}$ for the instantiation in $\mathcal{M}$ of a constant symbol $c \in L$ and $f^{\mathcal{M}}$ to denote the function on $\mathcal{M}$-tuples corresponding to the function symbol $f \in L$. Given a sublanguage $L_0 \subseteq L$, we write $\mathcal{M}|_{L_0}$ to denote the restriction of $\mathcal{M}$ to L_0.

2.2. Definitional expansions

Fundamental to our main construction is a special sort of sentence. We define the **pithy Π_2 sentences** of $\mathcal{L}_{\omega_1,\omega}(L)$ to be those $\mathcal{L}_{\omega_1,\omega}(L)$-sentences that are of the form

$$(\forall \overline{x})(\exists y)\varphi(\overline{x}, y),$$

where $\varphi \in \mathcal{L}_{\omega_1,\omega}(L)$ is quantifier-free with free variables precisely $\overline{x}, y$, and where the tuple $\overline{x}$ of variables is possibly empty. We say that a theory $T \subseteq \mathcal{L}_{\omega_1,\omega}(L)$ is pithy Π_2 when each sentence in T is.

In Sections 5 and 6 we will make use of the following technical result, which produces a definitional expansion of the empty theory to a pithy Π_2 theory Σ_A in which every formula in a desired admissible set A is equivalent to a quantifier-free formula; we call Σ_A the **definitional expansion for** A. This result is a straightforward extension of the standard *Morleyization* method.

Lemma 2.1. *For every admissible set $A \supseteq L$, there is an expanded language $L_A \subseteq A$ and a pithy Π_2 theory $\Sigma_A \subseteq \mathcal{L}_{\omega_1,\omega}(L_A) \cap A$ such that*

(i) for every formula $\varphi \in \mathcal{L}_{\omega_1,\omega}(L) \cap A$, there is some atomic formula $R_\varphi \in L_A$ such that

$$\Sigma_A \models (\forall \overline{x})\Big[\big((\forall w)R_\varphi(\overline{x}, w) \leftrightarrow \varphi(\overline{x})\big) \wedge \big((\exists w)R_\varphi(\overline{x}, w) \leftrightarrow \varphi(\overline{x})\big)\Big],$$

where $\overline{x}$ is the tuple of free variables of φ,

(ii) every L-structure has a unique expansion to an L_A-structure that satisfies Σ_A, and

(iii) Σ_A implies that every atomic formula of $\mathcal{L}_{\omega,\omega}(L_A) \backslash \mathcal{L}_{\omega,\omega}(L)$ is equivalent to some formula of $\mathcal{L}_{\omega_1,\omega}(L) \cap A$.

Proof. Consider the countable language $L_A := L \cup \{R_\psi : \psi \in A\}$, where each relation symbol R_ψ is a distinct element of $A \setminus L$ and has arity one more than the number of free variables in ψ.

Let Σ_A be the countable $\mathcal{L}_{\omega_1,\omega}(L_A)$-theory consisting of the following Π_2 sentences:

- $(\forall \overline{x}, w)[R_P(\overline{x}, w) \leftrightarrow P(\overline{x})]$ for P a relation symbol in L of arity $|\overline{x}|$,

- $(\forall \overline{x}, w)[R_c(y, w) \leftrightarrow c = y]$ for c a constant symbol in L,
- $(\forall \overline{x}, w)[R_f(\overline{x}, y, w) \leftrightarrow f(\overline{x}) = y]$ for f a function symbol in L of arity $|\overline{x}|$,
- $(\forall \overline{x}, w)[R_{\neg\psi}(\overline{x}, w) \leftrightarrow \neg R_\psi(\overline{x}, w)]$,
- $(\forall \overline{x}, w)[R_{\bigwedge_{i\in I}\psi_i}(\overline{x}, w) \leftrightarrow \bigwedge_{i\in I} R_{\psi_i}(\overline{z}_i, w)]$,
- $(\forall \overline{x}, w)[R_{(\exists y)\varphi}(\overline{x}, w) \leftrightarrow (\exists y)R_\varphi(\overline{x}, y, w)]$, and
- $(\forall \overline{x}, w)[R_{(\exists y)\psi}(\overline{x}, w) \leftrightarrow (\exists y)R_\psi(\overline{x}, w)]$,

where $\overline{x}$ is a tuple containing precisely the free variables of $\psi \in A$, where $\bigwedge_{i\in I}\psi_i \in A$, where the tuple $\overline{z}_i \subseteq \overline{x}$ contains precisely the free variables of ψ_i for each $i \in I$, and where the free variables of $\varphi \in A$ are precisely the variables in $\overline{x}y$, with $y \notin \overline{x}$.

Note that $(\forall \overline{x})[\varphi(\overline{x}) \leftrightarrow \psi(\overline{x})]$ is equivalent to $(\forall \overline{x})[\varphi(\overline{x}) \rightarrow \psi(\overline{x})] \wedge (\forall \overline{x})[\psi(\overline{x}) \rightarrow \varphi(\overline{x})]$. Hence Σ_A is equivalent to a theory all of whose axioms are either Π_1 or pithy Π_2. Further, every Π_1 sentence is equivalent to some pithy Π_2 sentence. Hence we may assume without loss of generality that Σ_A itself is a pithy Π_2 theory.

Observe that $\Sigma_A \subseteq A$ and that

$$\Sigma_A \models (\forall \overline{x})\Big[\big((\forall w)R_\varphi(\overline{x}, w) \leftrightarrow \varphi(\overline{x})\big) \wedge \big((\exists w)R_\varphi(\overline{x}, w) \leftrightarrow \varphi(\overline{x})\big)\Big],$$

for all $\varphi \in \mathcal{L}_{\omega_1,\omega}(L_A) \cap A$, where $\overline{x}$ is the tuple of free variables of φ.

Note that in the definition of Σ_A, we included the dummy variable w in order to ensure that for every $\psi \in A$, there is a universal formula that is equivalent to ψ in every model of Σ_A, even for quantifier-free ψ. This is needed in order for Σ_A to itself be pithy Π_2, which often is not required in the usual first-order Morleyization procedure [18, Theorem 2.6.6].

An immediate generalization of [18, Theorem 2.6.5] to countable fragments of $\mathcal{L}_{\omega_1,\omega}(L)$ shows that every L-structure $\mathcal{M}$ has a unique expansion to an L_A-structure that satisfies Σ_A. Finally, Σ_A implies that every atomic formula of $\mathcal{L}_{\omega,\omega}(L_A) \setminus \mathcal{L}_{\omega,\omega}(L)$ is equivalent to some formula of $\mathcal{L}_{\omega_1,\omega}(L) \cap A$. □

We will make use of Lemma 2.1 in the proof of Proposition 6.14.

For a first-order theory $T \subseteq \mathcal{L}_{\omega,\omega}(L)$, we define the **pithy Π_2 expansion of T** to be the $\mathcal{L}_{\omega,\omega}(L_{\mathrm{HF}})$-theory

$$\Sigma_{\mathrm{HF}} \cup \{(\forall x)R_\varphi(x) \,:\, \varphi \in T\},$$

where HF denotes the hereditarily finite sets. We will make use of this notion in Lemma 5.2 and Theorem 5.3.

2.3. Fraïssé limits and trivial definable closure

Suppose that the countable language L is **relational**, i.e., does not contain constant or function symbols. The **age** of an L-structure $\mathcal{M}$ is defined to be the class of all finite L-structures isomorphic to a substructure of $\mathcal{M}$.

A countable L-structure $\mathcal{M}$ is said to be **ultrahomogeneous** when any partial isomorphism between finite substructures of $\mathcal{M}$ can be extended to automorphism of $\mathcal{M}$. Any two ultrahomogeneous countably infinite L-structures have the same age if and only if they are isomorphic. The age of any ultrahomogeneous countably infinite L-structure is a class that contains countably infinitely many isomorphism types and that satisfies the so-called hereditary property, joint embedding property, and amalgamation property. Conversely, any class of finite L-structures that is closed under isomorphism, contains countably infinitely many isomorphism types, and that satisfies these three properties is the age of some ultrahomogeneous countably infinite L-structure, in fact a unique such structure (up to isomorphism), called its **Fraïssé limit**; such a class of finite structures is called an **amalgamation class**. An amalgamation class is called a **strong amalgamation class** when it further satisfies the strong amalgamation property — namely, when any two elements of the class can be amalgamated over any finite common substructure in a non-overlapping way.

It is a standard fact that the first-order theory of any Fraïssé limit in a finite relational language has an axiomatization consisting of pithy Π_2 sentences that are first-order. These axioms are often referred to as *(one-point) extension axioms*. For more details, see, e.g., [18, §7.1].

Let $\mathcal{M}$ be an L-structure and let M be its underlying set. Suppose $X \subseteq M$. The **definable closure** of X in $\mathcal{M}$, written $\mathrm{dcl}(X)$, is the set of all elements of M that are fixed by every automorphism of $\mathcal{M}$ fixing X pointwise. We say that $\mathcal{M}$ has **trivial definable closure** when $\mathrm{dcl}(\overline{a}) = \overline{a}$ for all finite tuples $\overline{a} \in \mathcal{M}$. An ultrahomogeneous countably infinite structure $\mathcal{M}$ in a relational language has trivial definable closure if and only if its age has the strong amalgamation property (again see [18, §7.1]).

2.4. Transitive G-spaces

Let $(G, e, \cdot)$ be a Polish group. We now recall the notion of a *transitive Borel G-space*.

Definition 2.2. A **Borel G-space** $(X, \circ)$ consists of a Borel space X along with a Borel map $\circ\colon G \times X \to X$ such that

- $(g \cdot h) \circ x = g \circ (h \circ x)$ for every $g, h \in G$ and $x \in X$, and
- $e \circ x = x$ for every $x \in X$.

For Borel G-spaces $(X, \circ_X)$ and $(Y, \circ_Y)$, a **map** τ between $(X, \circ_X)$ and $(Y, \circ_Y)$ is a Borel map $\tau\colon X \to Y$ for which $\tau(g \circ_X x) = g \circ_Y \tau(x)$ for all $g \in G$ and $x \in X$.

A Borel G-space $(X, \circ)$ is a **universal** Borel G-space when every other Borel G-space maps injectively into it.

Definition 2.3. A Borel G-space $(X, \circ)$ is **transitive** when for every $x, y \in X$ there is some $g \in G$ such that $g \circ x = y$, i.e., the action $\circ$ has a single orbit. Equivalently, there is no proper subspace $Y \subseteq X$ such that $(Y, \circ)$ is also a Borel G-space.

Note that in particular, any orbit of a Borel G-space is itself a transitive Borel G-space under the restricted action.

The main result of Section 6 is a classification of *transitive* Borel G-spaces for certain groups G.

2.5. Structures and automorphisms

We consider three types of countable structures: those with underlying set $\mathbb{N}$, those with a fixed countable set of constants disjoint from $\mathbb{N}$, and those with underlying set $\mathbb{N}$ whose restriction to a sublanguage is some fixed structure.

2.5.1. The Borel space of countable structures

We now define the Borel space Str_L and its associated *logic action*. These notions will be used throughout the paper, and especially in Sections 3, 5, and 6.

Definition 2.4. Let L be a countable language. Define Str_L to be the set of L-structures with underlying set $\mathbb{N}$.

Definition 2.5. Let L be a countable language. Then for every $\mathcal{L}_{\omega_1,\omega}(L)$-formula φ, define

$$[\![\varphi(\ell_1,\ldots,\ell_j)]\!] := \{\mathcal{M} \in \mathrm{Str}_L \, : \, \mathcal{M} \models \varphi(\ell_1,\ldots,\ell_j)\}$$

for all $\ell_1,\ldots,\ell_j \in \mathbb{N}$, where $j \in \mathbb{N}$ is the number of free variables (possibly 0) of φ.

When Str_L is equipped with the σ-algebra consisting of all such sets $[\![\varphi(\ell_1,\ldots,\ell_j)]\!]$, it becomes a standard Borel space; for details, see [19, §2.5]. Note that when we say that a probability measure is concentrated on some class of models of an L-theory, we mean that the measure is concentrated on the restriction of that class to Str_L.

Definition 2.6. For a non-empty set A, we write S_A to denote the symmetric group on A. For $n \in \mathbb{N}$, we write S_n to denote $S_{\{0,\ldots,n-1\}}$, and we will use S_∞ to denote $S_{\mathbb{N}}$, the symmetric group on $\mathbb{N}$.

Definition 2.7 ([19, §2.5]). Let L be a countable language. Define the Borel S_∞-action

$$\circledast_L \colon S_\infty \times \mathrm{Str}_L \to \mathrm{Str}_L$$

to be such that for all $g \in S_\infty$ and $\mathcal{M} \in \mathrm{Str}_L$,

$$g \circledast_L \mathcal{M} \models \varphi(\ell_1,\ldots,\ell_j)$$

if and only if

$$\mathcal{M} \models \varphi\big(g^{-1}(\ell_1),\ldots,g^{-1}(\ell_j)\big)$$

for all $\mathcal{L}_{\omega_1,\omega}(L)$-formulas φ and all $\ell_1,\ldots,\ell_j \in \mathbb{N}$, where j is the number of free variables of φ.

2.5.2. Countable structures with a fixed set of constants

We now define the analogous notions for the situation where we instantiate constants by elements other than ones from $\mathbb{N}$. We will need these notions in Section 4.

Definition 2.8. Let L be a countable language and let C be the set of its constant symbols (possibly empty). Let C_0 be a countable set (empty when C is empty) that is disjoint from $\mathbb{N}$, and suppose $\mathscr{C}_0 \colon C \to C_0$ is a surjective function. Then define $\mathrm{Str}_{\mathscr{C}_0,L}$ to be the set of L-structures with underlying set $\mathbb{N} \cup C_0$ in which the instantiation of c is $\mathscr{C}_0(c)$, for each constant symbol $c \in C$. In particular, no element of $\mathbb{N}$ instantiates any constant symbol of L.

Note that when L has no constant symbols, then $C = C_0 = \emptyset$ and $\mathscr{C}_0$ is the empty function, and we have $\mathrm{Str}_{\mathscr{C}_0,L} = \mathrm{Str}_L$.

Definition 2.9. Let L be a countable language with C its set of constant symbols, and let C_0 and $\mathscr{C}_0$ be as in Definition 2.8. Then for every $\mathcal{L}_{\omega_1,\omega}(L)$-formula φ, define

$$[\![\varphi(\ell_1,\ldots,\ell_j)]\!]_{\mathscr{C}_0} := \{\mathcal{M} \in \mathrm{Str}_{\mathscr{C}_0,L} \,:\, \mathcal{M} \models \varphi(\ell_1,\ldots,\ell_j)\}$$

for all $\ell_1,\ldots,\ell_j \in \mathbb{N}$, where $j \in \mathbb{N}$ is the number of free variables (possibly 0) of φ.

When $\mathrm{Str}_{\mathscr{C}_0,L}$ is equipped with the σ-algebra consisting of all such sets $[\![\varphi(\ell_1,\ldots,\ell_j)]\!]_{\mathscr{C}_0}$, it likewise becomes a standard Borel space.

Definition 2.10. Let L be a countable language with C its set of constant symbols, and let C_0 and $\mathscr{C}_0$ be as in Definition 2.8. Define $S_\infty^{C_0} \subseteq S_{\mathbb{N}\cup C_0}$ to be the subgroup consisting of all permutations of $\mathbb{N}\cup C_0$ fixing C_0 pointwise. Define the Borel $S_\infty^{C_0}$-action

$$\circledast_{\mathscr{C}_0,L} \colon S_\infty^{C_0} \times \mathrm{Str}_{\mathscr{C}_0,L} \to \mathrm{Str}_{\mathscr{C}_0,L}$$

to be such that for all $g \in S_\infty^{C_0}$ and $\mathcal{M} \in \mathrm{Str}_{\mathscr{C}_0,L}$,

$$g \circledast_{\mathscr{C}_0,L} \mathcal{M} \models \varphi(\ell_1,\ldots,\ell_j)$$

if and only if

$$\mathcal{M} \models \varphi\big(g^{-1}(\ell_1),\ldots,g^{-1}(\ell_j)\big)$$

for all $\mathcal{L}_{\omega_1,\omega}(L)$-formulas φ and all $\ell_1,\ldots,\ell_j \in \mathbb{N}$, where j is the number of free variables of φ.

Note that any permutation of $\mathbb{N}$ extends uniquely to a permutation of $\mathbb{N}\cup C_0$ that fixes C_0 pointwise, and every such permutation of $\mathbb{N}\cup C_0$ restricts to a permutation of $\mathbb{N}$, and hence $S_\infty \cong S_\infty^{C_0}$.

2.5.3. Relativized notions via sublanguages

Finally, we consider structures with underlying set $\mathbb{N}$ whose restriction to a sublanguage is some fixed structure. We will make use of such structures in Section 6.

Definition 2.11. Let L be a countable language and let $\mathcal{M}$ be an L-structure with underlying set $\mathbb{N}$. We write $\mathrm{Aut}(\mathcal{M})$ to denote the **automorphism group** of $\mathcal{M}$, i.e., the subgroup of S_∞ consisting of all permutations of $\mathbb{N}$ that preserve every relation, constant, and function of $\mathcal{M}$.

Definition 2.12. Let L be a countable language and let L_0 be a sublanguage of L. Let $\mathcal{M}_0$ be an L_0-structure on $\mathbb{N}$. Define $\mathrm{Str}_{L_0,L}^{\mathcal{M}_0}$ to be the collection of those structures in Str_L whose restriction to L_0 is $\mathcal{M}_0$, i.e.,

$$\mathrm{Str}_{L_0,L}^{\mathcal{M}_0} := \{\mathcal{M} \in \mathrm{Str}_L \,:\, \mathcal{M}|_{L_0} = \mathcal{M}_0\}.$$

Note that when L has no constant symbols, L_0 is the empty language, $\mathcal{M}_0$ is the empty structure, and $\mathscr{C}_0$ is the empty function, we have $\mathrm{Str}_{L_0,L}^{\mathcal{M}_0} = \mathrm{Str}_{\mathscr{C}_0,L} = \mathrm{Str}_L$. If L does have constant symbols, but L_0 and $\mathcal{M}_0$ are empty, then we still have $\mathrm{Str}_{L_0,L}^{\mathcal{M}_0} = \mathrm{Str}_L$.

Definition 2.13. Let L be a countable language and let L_0 be a sublanguage of L. Let $\mathcal{M}_0$ be an L_0-structure on $\mathbb{N}$. Then for every $\mathcal{L}_{\omega_1,\omega}(L)$-formula φ, define

$$[\![\varphi(\ell_1,\ldots,\ell_j)]\!]_{\mathcal{M}_0} := \{\mathcal{M} \in \mathrm{Str}^{\mathcal{M}_0}_{L_0,L} \,:\, \mathcal{M} \models \varphi(\ell_1,\ldots,\ell_j)\}$$

for all $\ell_1,\ldots,\ell_j \in \mathbb{N}$, where $j \in \mathbb{N}$ is the number of free variables (possibly 0) of φ.

When $\mathrm{Str}^{\mathcal{M}_0}_{L_0,L}$ is equipped with the σ-algebra consisting of all such sets $[\![\varphi(\ell_1,\ldots,\ell_j)]\!]_{\mathcal{M}_0}$, it also becomes a standard Borel space.

Definition 2.14 ([19, §2.7]). Let L be a countable language and let L_0 be a sublanguage of L. Let $\mathcal{M}_0$ be an L_0-structure on $\mathbb{N}$. Define the **relativized logic action**

$$\circledast^{\mathcal{M}_0}_L \colon \mathrm{Aut}(\mathcal{M}_0) \times \mathrm{Str}^{\mathcal{M}_0}_{L_0,L} \to \mathrm{Str}^{\mathcal{M}_0}_{L_0,L}$$

to be the restriction of the action $\circledast_L \colon S_\infty \times \mathrm{Str}_L \to \mathrm{Str}_L$.

3. Toy construction

We now provide a toy construction of invariant measures via limits of finite structures, where the measure is concentrated on the isomorphism class of a single graph. This is a simplification of a special case of the main construction of this paper, which we present in order to illustrate several motivating ideas, in a considerably easier setting. This toy construction is also a variant of a special case of the main construction of [5], where it is shown that whenever a countably infinite structure $\mathcal{M}$ in a countable language L has trivial definable closure, there is an S_∞-invariant measure on Str_L concentrated on the isomorphism class of $\mathcal{M}$.

All graphs in this section will be simple graphs, i.e., undirected unweighted graphs with no loops or multiple edges. Model-theoretically, such a graph is considered to be a structure in the *language of graphs*, i.e., a language consisting of a single binary relation symbol (interpreted as the edge relation), in which the edge relation is symmetric and irreflexive.

This toy construction applies only to the special case where the target structure is an ultrahomogeneous countably infinite graph having trivial definable closure. Admittedly, there are not many such structures: only a small number of parametrized classes of countably infinite graphs are ultrahomogeneous (see [20]), and fewer still have trivial definable closure (see, e.g., [5]) — and even those have been treated before (essentially in [4]). However, this toy construction serves to illustrate some of the key ideas of the main construction. In fact, the case of graphs is particularly simple, because it allows us to make use of results from the theory of dense graph limits.

Roughly speaking, given a target countably infinite ultrahomogeneous graph, we will build a sequence of finite graphs such that subgraphs sampled from them (in an appropriate sense) look more and more like induced "typical" subgraphs of the target. Then the distribution of an appropriate limit of the random graphs resulting from this sequence of sampling procedures will constitute the invariant measure concentrated on the isomorphism class of our target.

Our construction of the sequence of finite graphs resembles a directed system of finite graphs. This motivates our main construction in Section 4, which is built from directed systems in a more precise sense.

A key notion in the toy construction will be that of "duplication", whereby a sequence of elements branches into multiple copies that stand in parallel relationship to each other. This notion, too, will be essential in the main construction.

Suppose $\mathcal{M}$ is a countably infinite graph with underlying set M that is a Fraïssé limit whose age has the strong amalgamation property; recall that for relational languages, this property is equivalent to $\mathcal{M}$ having trivial definable closure.

The strong amalgamation property implies an important property that we call *duplication of quantifier-free types*: given any finite subset $A \subseteq M$ and any element $s \in M \setminus A$, there is some $s' \in M \setminus A$ such that the quantifier-free type of $A \cup \{s\}$ is the same as the quantifier-free type of $A \cup \{s'\}$.

As a consequence of this duplication property, for any $s_1, \ldots, s_n \in M$, we can find sets $S_1, \ldots, S_n \subseteq \mathcal{M}$ of arbitrary finite sizes such that each $s_i \in S_i$, and such that for any tuple $s'_1, \ldots, s'_n$ satisfying $s'_i \in S_i$ for $1 \leq i \leq n$, the quantifier-free type of $s'_1, \ldots, s'_n$ is the same as the quantifier-free type of $s_1, \ldots, s_n$. We call the sequence $S_1, \ldots, S_n$ a **branching** of $s_1, \ldots, s_n$, and say that each s_i branches into $|S_i|$-many **offshoots**.

3.1. Convergence and graph limits

As above, let $\mathcal{M}$ be an arbitrary countably infinite ultrahomogeneous graph whose age has the strong amalgamation property. We will construct a probability measure on countably infinite graphs with underlying set $\mathbb{N}$ that is invariant under arbitrary permutations of $\mathbb{N}$ and is concentrated on the isomorphism class of $\mathcal{M}$. We will do so by constructing a sequence $\langle \mathcal{M}_i \rangle_{i \in \mathbb{N}}$ of graphs of increasingly large finite size, and considering the corresponding sequence of infinite random graphs $\left\langle \mathbb{G}(\mathbb{N}, \mathcal{M}_i) \right\rangle_{i \in \mathbb{N}}$.

Definition 3.1. Let G be a finite graph. The **infinite random graph induced from G with replacement**, written $\mathbb{G}(\mathbb{N}, G)$, is a countably infinite random graph with underlying set $\mathbb{N}$ with edges defined as follows. Let $\langle x_i \rangle_{i \in \mathbb{N}}$ be a sequence of elements of G uniformly independently sampled with replacement. Then distinct $j, k \in \mathbb{N}$ have an edge between them in $\mathbb{G}(\mathbb{N}, G)$ precisely when x_j and x_k have an edge between them in G.

This sampling procedure has arisen independently a number of times; see [7, §10.1] for some of its history. The form we use can be concisely described using the theory of dense graph limits, or *graphons*; see [7, §11.2.2] for details. That work describes, given a graphon, a distribution on countably infinite graphs built from that graphon, called the *countable random graph model.* This distribution corresponds to the distribution of $\mathbb{G}(\mathbb{N}, G)$ in Definition 3.1 in the case where the graphon in question is the step-function built from G ([7, §7.1]). Note, however, that this distribution does not cohere with the definition in [7, §10.1] of $\mathbb{G}(k, G)$ for finite k bounded by the number of vertices of G, which involves sampling without replacement.

Our goal is to find a sequence of finite graphs $\langle \mathcal{M}_i \rangle_{i \in \mathbb{N}}$ as above, such that the sequence of random variables $\langle \mathbb{G}(\mathbb{N}, \mathcal{M}_i) \rangle_{i \in \mathbb{N}}$ converges in distribution to a random graph that is almost surely isomorphic to $\mathcal{M}$, and whose distribution is invariant under permutations of $\mathbb{N}$. The invariance will be automatic, as each $\mathbb{G}(\mathbb{N}, \mathcal{M}_n)$ is obtained via i.i.d. sampling, as described in the definition. In order to show the convergence, we will use results from the theory of graphons.

Given a graph G, we write $v(G)$ to denote the number of vertices of G.

Definition 3.2. Let F, G be finite graphs. Let $k = v(F)$ and $n = v(G)$. Then $t_{\text{full}}(F, G)$, the **full homomorphism density**, is defined to be the fraction of maps from F to G that preserve both adjacency and non-adjacency, i.e.,

$$t_{\text{full}}(F, G) = \frac{\text{Full}(F, G)}{n^k},$$

where $\text{Full}(F, G)$ is the number of homomorphisms from F to G that also preserve non-adjacency.

The value $t_{\text{full}}(F, G)$ may also be described in terms of the following random procedure. First consider an independent random selection of $v(F)$-many vertices of G chosen uniformly with replacement, each labeled with the corresponding element of F. (In particular, some vertices of G may be labeled by multiple vertices of F.) Then $t_{\text{full}}(F, G)$ is the probability that the graph with labels from F induced by the sampling procedure is a labeled copy of F, preserving both edges and non-edges.

This notion of a full homomorphism occurs in the graph homomorphism literature, e.g., in [21, §1.10.10]. Note, however, that t_{full} is somewhat different from the various densities that are typically used in the study of graph limits, namely, the density t of homomorphisms, t_{inj} of injective homomorphisms, and t_{ind} of induced injective homomorphisms, i.e., embeddings; for details see [7, §5.2.2].

Definition 3.3. We say that a sequence of finite graphs $\langle G_i \rangle_{i \in \mathbb{N}}$ is **unbounded** when $\lim_{i \to \infty} v(G_i) = \infty$.

The following definition of a type of convergence is slightly nonstandard as it uses t_{full}, but is equivalent to the more usual definitions in the literature on dense graph limits, which involve the other density notions, as described in the discussion in the beginning of [7, §11.1].

Definition 3.4. An unbounded sequence of finite graphs $\langle G_i \rangle_{i \in \mathbb{N}}$ is **convergent** when the sequence of induced subgraph densities

$$\big\langle t_{\text{full}}(F, G_i) \big\rangle_{i \in \mathbb{N}}$$

converges for every finite graph F.

Theorem 3.5 ([7, Theorem 11.7]). *Let $\langle G_i \rangle_{i \in \mathbb{N}}$ be an unbounded sequence of finite graphs that is convergent. Then $\big\langle \mathbb{G}(\mathbb{N}, G_i) \big\rangle_{i \in \mathbb{N}}$ converges in distribution to a countably infinite random graph whose distribution is an S_∞-invariant measure.*

In fact, every such S_∞-invariant measure is ergodic, as shown by Aldous [22, Lemma 7.35]; for an argument involving graph limits, see [23, Proposition 3.6].

Corollary 3.6. *Let $\langle G_i \rangle_{i \in \mathbb{N}}$ be an unbounded sequence of finite graphs. Suppose the limiting probability*

$$\lim_{i \to \infty} \mathbb{P}\big(\mathbb{G}(\mathbb{N}, G_i) \models q(0, \ldots, \ell - 1)\big)$$

exists for every quantifier-free type q in the language of graphs, where ℓ is the number of free variables of q. Then $\big\langle \mathbb{G}(\mathbb{N}, G_i) \big\rangle_{i \in \mathbb{N}}$ converges in distribution to an S_∞-invariant measure on countably infinite graphs.

Proof. By Theorem 3.5, it suffices to show that $\langle t_{\text{full}}(F, G_i)\rangle_{i\in\mathbb{N}}$ converges for every finite graph F.

Let F be an arbitrary finite graph with underlying set $\{0, \ldots, n-1\}$, where $n = v(F)$. Let q_F be the unique non-redundant quantifier-free type with n-many free variables such that

$$F \models q_F(0, \ldots, n-1).$$

Note that, for each $j \in \mathbb{N}$,

$$t_{\text{full}}(F, G_j) = \mathbb{P}\big(\mathbb{G}(\mathbb{N}, G_j) \models q_F(0, \ldots, n-1)\big).$$

Hence $\langle t_{\text{full}}(F, G_i)\rangle_{i\in\mathbb{N}}$ converges, as

$$\Big\langle \mathbb{P}\big(\mathbb{G}(\mathbb{N}, G_i) \models q_F(0, \ldots, n-1)\big)\Big\rangle_{i\in\mathbb{N}}$$

converges by hypothesis. □

3.2. Construction

Because $\mathcal{M}$ is a Fraïssé limit in a finite relational language, as discussed in §2.3 we may take its first-order theory T to be axiomatized by pithy Π_2 extension axioms, so that

$$T = \{(\forall \overline{x})(\exists y)\varphi_i(\overline{x}, y) \,:\, i \in \mathbb{N}\},$$

where each φ_i is quantifier-free; we may further assume that for each $i \in \mathbb{N}$ there are infinitely many indices $j \in \mathbb{N}$ such that $\varphi_i = \varphi_j$. We will consider, in successive stages, each such formula $\varphi_i(\overline{x}, y)$ and every tuple $\overline{a} \in \mathcal{M}$ of the same length as $\overline{x}$, and will look for **witnesses** in $\mathcal{M}$ to $(\exists y)\varphi_i(\overline{a}, y)$, i.e., instantiations $b \in \mathcal{M}$ of y that make $\varphi_i(\overline{a}, b)$ hold in $\mathcal{M}$.

Our construction proceeds in stages, at each of which we build a finite structure larger than that in the previous stage. We will think of the structure that we build at stage n as consisting of $(n+1)$-many slices, each built at a substage. In the first substage of stage n, we add a slice that consists of new witnesses to the formula under consideration (or one new element, if no witnesses are needed). In the remaining substages, we branch each element of each old slice into some number of offshoots.

Specifically, we divide each stage n into $(n+1)$-many distinct substages indexed by pairs (n, k), where $0 \le k \le n$. The substage $(n, 0)$ involves adding witnesses to extension axioms for everything from stage $n-1$ (as one often does when iteratively building a Fraïssé limit). The substages (n, k), for $0 < k \le n$, consist of successively branching elements. By duplicating ever larger portions, we cause the structure to asymptotically stabilize.

More precisely, at substage (n, k) we will define a structure $\mathcal{M}_n^k$ and a set $B(n, n-k)$. The intuition is that $B(n, n)$ consists of new witnesses, while $B(n, n-k)$, for $k > 0$, consists of all elements of $\mathcal{M}_n^k$ that are offshoots of elements that first appear at substage $(n-k, 0)$. In particular, the underlying set of $\mathcal{M}_n^k$ will be

$$\bigcup_{i=0}^{n-1-k} B(n-1, i) \;\cup\; \bigcup_{i=n-k}^{n} B(n, i), \qquad (\star)$$

because at substage (n,k), the newly-constructed set $B(n,n-k)$ contains all elements of $B(n-1,n-k)$.

Substage $(0,0)$: Let $\mathcal{M}_0^0$ be any finite substructure of the Fraïssé limit $\mathcal{M}$, and let $B(0,0)$ be its underlying set.

Substage $(n,0)$, for $n>0$: Let ℓ_n be one less than the number of free variables in the formula φ_n. Let A be the set of those $\overline{a} \subseteq \mathcal{M}_{n-1}^{n-1}$ of length ℓ_n such that $\mathcal{M}_{n-1}^{n-1} \not\models \bigvee_{b\in\overline{a}} \varphi_n(\overline{a},b)$. We now define $B(n,n)$ and $\mathcal{M}_n^0$. Consider whether or not A is empty.

If A is non-empty, then for each $\overline{a} \in A$ choose a distinct element $d_{\overline{a}} \in \mathcal{M}$ that satisfies $\mathcal{M}_n^0 \models \varphi_n(\overline{a}, d_{\overline{a}})$. We can always find such a collection of witnesses, because our formulas are realized in the Fraïssé limit $\mathcal{M}$. Furthermore, because $\mathcal{M}$ has strong amalgamation, by duplication of quantifier-free types, we may assume that for any distinct tuples $\overline{a}, \overline{a}' \in \mathcal{M}_{n-1}^{n-1}$, the elements $d_{\overline{a}}$ and $d_{\overline{a}'}$ are distinct. Define $B(n,n) = \{d_{\overline{a}} : \overline{a} \in A\}$ and let $\mathcal{M}_n^0$ be any substructure of $\mathcal{M}$ extending $\mathcal{M}_{n-1}^{n-1}$ by the elements of $B(n,n)$.

If A is empty, then let $B(n,n)$ consist of an arbitrary single element of $\mathcal{M}$ not in $\mathcal{M}_{n-1}^{n-1}$, and set $\mathcal{M}_n^0$ to be the (unique) substructure of $\mathcal{M}$ extending $\mathcal{M}_{n-1}^{n-1}$ by the element of $B(n,n)$.

Substage (n,k) for $0<k\le n$: Let $\alpha_n := 2^{n-1}\,|B(n,n)|$. Let $\mathcal{M}_n^k$ be any substructure of $\mathcal{M}$ that extends $\mathcal{M}_n^{k-1}$ to some structure in which each element of $B(n-1,n-k)$ branches into precisely α_n-many offshoots, and these are the only new elements. Let $B(n,n-k)$ be the set of those elements of $\mathcal{M}_n^k$ that are an offshoot of some element of $B(n-1,n-k)$. By the definition of α_n, we have

$$\frac{|B(n,n)|}{|\mathcal{M}_n^k|} \le 2^{-(n-1)}.$$

This concludes the construction.

For notational convenience, we will henceforth refer to $\mathcal{M}_n^n$ as $\mathcal{M}_n$.

In the verification, we will need a particular projection map. Let $\widetilde{\pi}$ be the following map from the union of the underlying sets of all $\mathcal{M}_n$, for $n \in \mathbb{N}$, to itself. The map $\widetilde{\pi}$ takes each element of $B(n,n-k)$ to the element of $B(n-k,n-k)$ of which it is a k-fold offshoot (i.e., an offshoot's offshoot's offshoot, etc., k levels deep), for $n \in \mathbb{N}$ and $0 \le k < n$, and the identity map on each $B(n,n)$. This is well-defined because if an element of the domain is in both $B(n,k)$ and $B(m,\ell)$, then $k=\ell$. (Note that $\widetilde{\pi}$ is not the same as the projection map π defined in Section 2, though it will play a similar role here to that of π in the main construction in Section 4.)

3.3. *Verification*

We now show that the sequence of random graphs $\langle \mathbb{G}(\mathbb{N}, \mathcal{M}_i)\rangle_{i\in\mathbb{N}}$ converges in distribution to a random graph that is almost surely isomorphic to our original graph $\mathcal{M}$. We show this in two parts: convergence to such a random graph, whose distribution is an invariant measure on countable graphs, and concentration of this invariant measure on the desired isomorphism class.

Proposition 3.7. *The sequence of random graphs $\big\langle \mathbb{G}(\mathbb{N}, \mathcal{M}_i)\big\rangle_{i\in\mathbb{N}}$ converges in distribution to a countably infinite random graph whose distribution is an S_∞-invariant measure.*

Proof. Note that $\langle \mathcal{M}_i\rangle_{i\in\mathbb{N}}$ is an unbounded sequence of finite graphs. Hence by Corollary 3.6, it suffices to show that

$$\Big\langle \mathbb{P}\big(\mathbb{G}(\mathbb{N}, \mathcal{M}_i) \models q(0, \dots, \ell - 1)\big)\Big\rangle_{i\in\mathbb{N}}$$

is Cauchy for every quantifier-free type q in the language of graphs, where ℓ is the number of free variables of q.

Fix such a q and ℓ. For each $n \in \mathbb{N}$, define

$$\delta_{n+1} := \mathbb{P}\big(\mathbb{G}(\mathbb{N}, \mathcal{M}_n) \models q(0, \dots, \ell - 1)\big) - \mathbb{P}\big(\mathbb{G}(\mathbb{N}, \mathcal{M}_{n+1}) \models q(0, \dots, \ell - 1)\big).$$

We will show that δ_{n+1} decays exponentially in n for fixed ℓ.

Let G_n be a sample from $\mathbb{G}(\mathbb{N}, \mathcal{M}_n)$, and let $\langle a_i\rangle_{i\in\mathbb{N}}$ be the random sequence of vertices (with replacement) chosen from $\mathcal{M}_n$ in the course of the sampling procedure. Likewise, let G_{n+1} be a sample from $\mathbb{G}(\mathbb{N}, \mathcal{M}_{n+1})$ with vertex sequence $\langle b_i\rangle_{i\in\mathbb{N}}$. Observe that

$$\mathbb{P}\big(\mathbb{G}(\mathbb{N}, \mathcal{M}_n) \models q(0, \dots, \ell - 1)\big) = \mathbb{P}\big(G_n \models q(a_0, \dots, a_{\ell-1})\big)$$

and

$$\mathbb{P}\big(\mathbb{G}(\mathbb{N}, \mathcal{M}_{n+1}) \models q(0, \dots, \ell - 1)\big) = \mathbb{P}\big(G_{n+1} \models q(b_0, \dots, b_{\ell-1})\big).$$

Let $E_{n+1,\ell}$ be the event that for each i such that $0 \le i \le \ell-1$, the projection $\widetilde{\pi}(b_i) \in \mathcal{M}_n$. By our construction, the conditional probability $\mathbb{P}\big(G_{n+1} \models q(b_0, \dots, b_{\ell-1}) \;\big|\; E_{n+1,\ell}\big)$ satisfies

$$\mathbb{P}\big(G_{n+1} \models q(b_0, \dots, b_{\ell-1}) \;\big|\; E_{n+1,\ell}\big) \;=\; \mathbb{P}\big(G_n \models q(a_0, \dots, a_{\ell-1})\big).$$

Therefore $\delta_{n+1} \le 1 - \mathbb{P}(E_{n+1,\ell})$.

By construction of $\mathcal{M}_{n+1}$, we have

$$\mathbb{P}(E_{n+1,\ell}) = \Big(1 - \frac{|B(n+1, n+1)|}{|\mathcal{M}_{n+1}|}\Big)^{\ell}$$

Recall that at the end of the construction we observed that

$$\frac{|B(n+1, n+1)|}{|\mathcal{M}_{n+1}|} \le 2^{-n},$$

and so $\mathbb{P}(E_{n+1,\ell}) \ge (1 - 2^{-n})^{\ell}$. Using Bernoulli's inequality, we obtain the bound $\mathbb{P}(E_{n+1,\ell}) \ge 1 - \ell\, 2^{-n}$, and so $\delta_{n+1} \le \ell\, 2^{-n}$, as desired. □

Let $\mu_{\mathcal{M}}$ denote the distribution of the limit of $\big\langle \mathbb{G}(\mathbb{N}, \mathcal{M}_i)\big\rangle_{i\in\mathbb{N}}$. Proposition 3.7 demonstrates that $\mu_{\mathcal{M}}$ is an S_∞-invariant measure on Str_L, where L is the language of graphs. We now show that $\mu_{\mathcal{M}}$ assigns measure 1 to the isomorphism class of $\mathcal{M}$. We begin with a combinatorial lemma.

Recall that for each $j \in \mathbb{N}$, we have defined $\ell_j \in \mathbb{N}$ to be one less than the number of free variables in the quantifier-free formula φ_j. For each $n, j \in \mathbb{N}$, define

$$\gamma_{n,j} := \mathbb{P}\Big(\mathbb{G}(\mathbb{N}, \mathcal{M}_n) \models (\exists y)\, \varphi_j\big(0 \cdots (\ell_j - 1), y\big)\Big).$$

Before proving our main bound on this quantity, we need a technical lemma.

Lemma 3.8. *Let $k \in \mathbb{N}$ and suppose $0 < C < 2^k$. Then*

$$\prod_{i=k}^{\infty} (1 - C\ 2^{-i}) \geq (1 - C\ 2^{-k})^2.$$

Proof. By our hypothesis on C, each term of the product is positive. In particular, we have

$$\log\Big(\prod_{i=k}^{\infty}(1 - C\ 2^{-i})\Big) = \sum_{i=k}^{\infty} \log(1 - C\ 2^{-i})$$

By the concavity of the function $\log(1-t)$, we have $\log(1-t) \geq t \log(1-t_0)/t_0$ for $t_0 \geq t > 0$. Setting $t_0 = C\,2^{-k}$ and $t = C\,2^{-i}$ where $i \geq k$, we obtain

$$\begin{aligned}
\sum_{i=k}^{\infty} \log(1 - C\ 2^{-i}) &\geq \sum_{i=k}^{\infty} C\,2^{-i} \log(1 - C\ 2^{-k})/(C\,2^{-k}) \\
&= \log(1 - C\ 2^{-k}) \sum_{i=k}^{\infty} 2^{-i+k} \\
&= 2 \log(1 - C\ 2^{-k}).
\end{aligned}$$

Therefore $\prod_{i=k}^{\infty}(1 - C\ 2^{-i}) \geq (1 - C\ 2^{-k})^2$ by the monotonicity of log. ☐

Lemma 3.9. *For all $j \in \mathbb{N}$,*

$$\lim_{n \to \infty} \gamma_{n,j} = 1.$$

Proof. By our enumeration of formulas φ_j, observe that for all $n \in \mathbb{N}$ there is some $j \geq 1$ such that $\gamma_{n,j} = \gamma_{n,0}$. Hence it suffices to prove the claim for all $j \geq 1$.

We may assume that j is large enough that $\ell_j < 2^{j-1}$, as each formula is enumerated infinitely often, and $\gamma_{n,j}$ depends only on n and on φ_j, not on j.

Fix such a $j \geq 1$. As in the proof of Proposition 3.7, for $n \in \mathbb{N}$ let G_n be a sample from $\mathbb{G}(\mathbb{N}, \mathcal{M}_n)$, and let $\langle a_i \rangle_{i \in \mathbb{N}}$ be the random sequence of vertices (with replacement) chosen from $\mathcal{M}_n$ in the course of the sampling procedure.

Analogously, for $n > j$, define D_{n,j,ℓ_j} to be the event that for each i such that $0 \leq i \leq \ell_j - 1$, the projection $\widetilde{\pi}(a_i) \in \mathcal{M}_{j-1}$. Recall that $\mathbb{P}(E_{h,\ell_j}) \geq 1 - \ell_j\ 2^{-(h-1)}$ for each $h \in \mathbb{N}$, and so

$$\begin{aligned}
\mathbb{P}(D_{n,j,\ell_j}) &\geq \mathbb{P}(E_{n,\ell_j}) \cdot \mathbb{P}(E_{n-1,\ell_j}) \cdots \mathbb{P}(E_{j,\ell_j}) \\
&\geq (1 - \ell_j\ 2^{-(n-1)}) \cdot (1 - \ell_j\ 2^{-(n-2)}) \cdots (1 - \ell_j\ 2^{-(j-1)}).
\end{aligned}$$

Taking $C = \ell_j$ and $k = j - 1$ in Lemma 3.8, we obtain

$$\prod_{i=j-1}^{\infty} (1 - \ell_j\ 2^{-i}) \geq (1 - \ell_j\ 2^{-(j-1)})^2,$$

and so

$$\mathbb{P}(D_{n,j,\ell_j}) \geq (1 - \ell_j\ 2^{-(j-1)})^2,$$

as each term in the infinite product is between 0 and 1.

Now let F_{n,ℓ_j} be the event that the elements $a_0, \ldots, a_{\ell_j - 1}$ of $\mathcal{M}_n$ are distinct. Observe that, because $\langle \mathcal{M}_i \rangle_{i \in \mathbb{N}}$ is an unbounded sequence of graphs,

$$\lim_{n \to \infty} \mathbb{P}(F_{n,\ell_j}) = 1.$$

Because of the way witnesses are chosen at substage (j, j), for $n > j$ if events D_{n,j,ℓ_j} and F_{n,ℓ_j} hold, then

$$\mathbb{G}(\mathbb{N}, \mathcal{M}_n) \models (\exists y)\varphi_j\big(0 \cdots (\ell_j - 1), y\big).$$

Therefore,

$$\mathbb{P}\Big(\mathbb{G}(\mathbb{N}, \mathcal{M}_n) \models (\exists y)\, \varphi_j\big(0 \cdots (\ell_j - 1), y\big)\Big) \geq \mathbb{P}(D_{n,j,\ell_j}) \cdot \mathbb{P}(F_{n,\ell_j}).$$

For $n > j$, define $\zeta(n, j)$ to be the greatest $k < n$ such that $\varphi_k = \varphi_j$. We then have

$$\begin{aligned} \mathbb{P}\Big(\mathbb{G}(\mathbb{N}, \mathcal{M}_n) \models (\exists y)\, \varphi_j\big(0 \cdots (\ell_j - 1), y\big)\Big) \quad &\geq \quad \mathbb{P}(D_{n,\zeta(n,j),\ell_j}) \cdot \mathbb{P}(F_{n,\ell_j}) \\ &\geq \quad (1 - \ell_j\ 2^{-(\zeta(n,j)-1)})^2 \cdot \mathbb{P}(F_{n,\ell_j}). \end{aligned}$$

Because each formula is enumerated infinitely often, $\lim_{n \to \infty} \zeta(n, j) = \infty$. Hence

$$\lim_{n \to \infty} \mathbb{P}\Big(\mathbb{G}(\mathbb{N}, \mathcal{M}_n) \models (\exists y)\, \varphi_j\big(0 \cdots (\ell_j - 1), y\big)\Big) = 1,$$

as desired. □

Proposition 3.10. *The S_∞-invariant measure $\mu_{\mathcal{M}}$ is concentrated on the isomorphism class of $\mathcal{M}$.*

Proof. For each $j \in \mathbb{N}$, we have

$$\mu_{\mathcal{M}}\Big(\big[\!\big[(\exists y)\varphi_j\big(0 \cdots (\ell_j - 1), y\big)\big]\!\big]\Big) = 1,$$

by Lemma 3.9. Therefore, by the S_∞-invariance of $\mu_{\mathcal{M}}$, we have

$$\mu_{\mathcal{M}}\Big(\big[\!\big[(\forall \overline{x})(\exists y)\varphi_j\big(\overline{x}, y\big)\big]\!\big]\Big) = 1,$$

where $\overline{x}$ is an ℓ-tuple of distinct variables. But T consists solely of sentences of the form $(\forall \overline{x})(\exists y)\varphi_j\big(\overline{x}, y\big)$. Hence $\mu_{\mathcal{M}}$ is concentrated on the class of models of T. Because T is $\aleph_0$-categorical and $\mathcal{M} \models T$, the measure $\mu_{\mathcal{M}}$ is concentrated on the isomorphism class of $\mathcal{M}$. □

4. Inverse limit construction

We now give the key technical construction of the paper. This will take a theory with certain properties and produce a probability measure, invariant under permutations of the non-constant elements in the underlying set, that is concentrated on the class of models of the theory. This construction is a variant of the one in [5] and will be the crucial tool used in later sections.

4.1. Setup

Before providing the construction itself, we describe the main conditions it requires. We begin by fixing the following languages, theories, and quantifier-free types.

First let $\langle L_i \rangle_{i\in\mathbb{N}}$ be an increasing sequence of countable languages having no function symbols, but possibly both constant and relation symbols, and let $L_\infty := \bigcup_{i\in\mathbb{N}} L_i$, so that

$$L_0 \subseteq L_1 \subseteq L_2 \subseteq \cdots \subseteq L_\infty.$$

Further assume that all constant symbols appearing in any L_i are already in the language L_0; call this set of constant symbols C.

Now fix an increasing sequence $\langle T_i \rangle_{i\in\mathbb{N}}$ of countable pithy Π_2 theories that are quantifier-free complete and satisfy $T_i \in \mathcal{L}_{\omega_1,\omega}(L_i)$ for each $i \in \mathbb{N}$. Let $T_\infty := \bigcup_{i\in\mathbb{N}} T_i$, so that

$$T_0 \subseteq T_1 \subseteq T_2 \subseteq \cdots \subseteq T_\infty.$$

For each $i \in \mathbb{N}$, let $Q_i = \langle q^i_j \rangle_{j\in\mathbb{N}}$ be any sequence of complete non-constant quantifier-free L_i-types that are consistent with T_i and that satisfy the following four conditions. Let k^i_j denote the number of free variables of q^i_j.

(W) For each $i, j \in \mathbb{N}$ and every sentence $(\forall \overline{x})(\exists y)\psi(\overline{x}, y) \in T_i$ for which $|\overline{x}| = k^i_j$, there is some $e_{i,j,\psi} \in \mathbb{N}$ such that $q^\natural := q^i_{e_{i,j,\psi}}$ is a quantifier-free type with one more free variable than q^i_j and such that

$$\models \quad (\forall \overline{x}, y)\big(q^\natural(\overline{x}, y) \rightarrow q^i_j(\overline{x})\big) \qquad \text{and}$$
$$\models \quad (\forall \overline{x}, y)\big(q^\natural(\overline{x}, y) \rightarrow \psi(\overline{x}, y)\big).$$

(D) For each $i, j \in \mathbb{N}$ and variable y such that q^i_j is a non-redundant quantifier-free type satisfying

$$\models (\forall \overline{x}, y)\big(q^i_j(\overline{x}, y) \rightarrow \bigwedge_{c\in C} (y \neq c)\big),$$

where $|\overline{x}| + 1 = k^i_j$, there is some $f_{i,j}$ such that the quantifier-free type $q^\natural := q^i_{f_{i,j}}$ has $(k^i_j + 1)$-many free variables, is non-redundant, and satisfies

$$\models (\forall \overline{x}, y, z)\big(q^\natural(\overline{x}, y, z) \rightarrow \big(q^i_j(\overline{x}, y) \wedge q^i_j(\overline{x}, z)\big)\big).$$

(E) For each $i, j \in \mathbb{N}$ there is some $j' \in \mathbb{N}$ such that

$$\models (\forall \overline{x})\big(q^{i+1}_{j'}(\overline{x}) \rightarrow q^i_j(\overline{x})\big),$$

where $|\overline{x}| = k^i_j = k^{i+1}_{j'}$.

(C) For each $i, j \in \mathbb{N}$ and quantifier-free type p such that

$$\models (\forall \overline{x}, \overline{w})\big(q^i_j(\overline{x}) \to p(\overline{w})\big),$$

where $|\overline{x}| = k^i_j$ and $\overline{w}$ is a subtuple of variables of $\overline{x}$ with $|\overline{w}|$ equal to the number of free variables of p, there is some $h_p \in \mathbb{N}$ such that $q^\natural := q^i_{h_p}$ satisfies

$$\models (\forall \overline{w})\big(q^\natural(\overline{w}) \leftrightarrow p(\overline{w})\big).$$

Condition (W) ensures that for each quantifier-free type in Q_i and pithy Π_2 sentence in the theory T_i, the sequence Q_i contains some extension of the quantifier-free type that *witnesses* the formula.

Condition (D) requires that for every non-redundant quantifier-free type in Q_i and every free variable of that quantifier-free type which it requires to not be instantiated by a constant, there is some other quantifier-free type in Q_i that *duplicates* that variable. In particular, by repeated use of (D), we can show that for any non-redundant $q^i_j \in Q_i$, any $h \in \mathbb{N}$, and any $k^* \leq k^i_j$ such that

$$(\forall \overline{x}, \overline{y})\big(q^i_j(\overline{x}, \overline{y}) \to \bigwedge_{z \in \overline{y}} \bigwedge_{c \in C} (z \neq c)\big)$$

where $|\overline{x}| = k^i_j - k^*$ and $|\overline{y}| = k^*$, there is an $f^*_{i,j,k^*} \in \mathbb{N}$ such that the quantifier-free type $q^\natural := q^i_{f^*_{i,j,k^*}}$ has $(k_{i,j} + h\,k^*)$-many free variables, and for all functions $\beta\colon \{1, \dots, k^*\} \to \{0, \dots, h\}$, we have

$$\models (\forall \overline{x},\, y^0_1 \cdots y^h_1 \,\cdots\, y^0_{k^*} \cdots y^h_{k^*})\ \big(q^\natural(\overline{x}, y^0_1 \cdots y^h_1 \,\cdots\, y^0_{k^*} \cdots y^h_{k^*}) \to q^i_j(\overline{x}, y^{\beta(1)}_1 \dots y^{\beta(k^*)}_{k^*})\big),$$

where the y^w_ℓ are new distinct variables, for $1 \leq \ell \leq k^*$ and $0 \leq w \leq h$.

In summary, $q^\natural$ is a quantifier-free type that duplicates, $(h+1)$-fold, all variables of q^i_j. Furthermore, for every tuple of variables from $q^\natural$ that contains exactly one duplicate of each variable of q^i_j, the resulting restriction of $q^\natural$ to those variables is precisely q^i_j with corresponding variables substituted. We call such a $q^\natural$ an **iterated duplicate**. Recall our assumption that each T_i is quantifier-free complete, hence consistent, and that each element of Q_i is consistent with T_i. Therefore, as a consequence of iterated duplication, each T_i must have models with infinitely many elements that do not instantiate constant symbols.

Condition (E) says that for every quantifier-free type in Q_i and larger language, we can find an *extension* of that quantifier-free type to that language.

Condition (C) says that the quantifier-free types of Q_i are *closed* under implication.

There will be one further condition, which we will not always require to hold. However, when it does hold, it will guarantee that the construction assigns measure 0 to every isomorphism class of models of the target theory.

(S) For some $\ell \in \mathbb{N}$ (called the *order* of splitting), every $i \in \mathbb{N}$, and every non-redundant quantifier-free L_i-type $q^i_j \in Q_i$ with $k^i_j \geq \ell$, there is some $e \in \mathbb{N}$ and some quantifier-free L_e-type $q^\natural \in Q_e$ with $2k^i_j$ many free variables, such that for each $\beta\colon \{1,\dots,k^i_j\} \to \{0,1\}$, we have

$$\models (\forall x^0_1 x^1_1 \dots x^0_{k^i_j} x^1_{k^i_j})\big(q^\natural(x^0_1 x^1_1 \dots x^0_{k^i_j} x^1_{k^i_j}) \to q^i_j(x^{\beta(1)}_1 \dots x^{\beta(k^i_j)}_{k^i_j})\big),$$

where $x^0_1 x^1_1 \cdots x^0_{k^i_j} x^1_{k^i_j}$ is a tuple of distinct free variables, and for each $i_1,\dots,i_\ell \in \mathbb{N}$ such that $1 \le i_1 < i_2 < \cdots < i_\ell \le k^i_j$, and each $\gamma_0,\gamma_1\colon \{1,\dots,\ell\} \to \{0,1\}$, there are distinct non-redundant $p_0,p_1 \in Q_j$ such that for $w \in \{0,1\}$,

$$\models (\forall x^0_1 x^1_1 \dots x^0_{k^i_j} x^1_{k^i_j})\big(q^\natural(x^0_1 x^1_1 \dots x^0_{k^i_j} x^1_{k^i_j}) \to p_w(x^{\gamma_w(1)}_{i_1} \dots x^{\gamma_w(\ell)}_{i_\ell})\big).$$

We call $q^\natural$ a *splitting* of q^i_j of order ℓ.

If there is such an ℓ then we say that $\langle Q_i\rangle_{i\in\mathbb{N}}$ has **splitting of quantifier-free types of order ℓ**.

The intuition is that if (S) is satisfied then for every non-redundant quantifier-free type in $\bigcup_{i\in\mathbb{N}} Q_i$ in at least ℓ-many free variables, there is some larger language in which we can duplicate the quantifier-free type so that every quantifier-free subtype with ℓ-many free variables splits into at least two distinct quantifier-free types. In other words, for each quantifier-free subtype with ℓ-many free variables, if we consider all ways in which it is duplicated (i.e., all the quantifier-free types where no two distinct free variables are duplicates of the same variable), then that collection of quantifier-free subtypes always has at least two elements. We will use this condition to show that any given quantifier-free ℓ-type is realized with probability 0.

We now turn to the construction itself.

4.2. Construction

The aim is to construct a continuum-sized measurable space with certain properties. This will proceed via the inverse limit of a system of finite structures in an increasing system of languages with associated measures. We will build this system of structures in stages, each of which will interleave four tasks. The first task is to enlarge the underlying set and update the measures so that they assign mass to the new set in a way that is compatible with our earlier choices. The second task is to add elements to ensure that ever more of our pithy Π_2 theory is realized, and adjust the mass accordingly. The third task is to make sure the quantifier-free type of the entire structure up until this point is duplicated. This will ensure that the end result is a continuum-sized structure. Finally, the fourth task is to ensure that if there is splitting of quantifier-free types of some order ℓ, then the appropriate quantifier-free type splits as we enlarge the language. This will ensure that we obtain a continuum-sized structure and a measure such that under a certain sampling procedure, the probability of any particular quantifier-free type with ℓ-many free variables being realized is 0, and hence sampling from our structure will not assign positive measure to the isomorphism class of any single structure.

The construction proceeds in stages indexed by $n \in \mathbb{N} \cup \{\infty\}$. At each finite stage, the structure we construct will have underlying set equal to the union of a fixed countable set C of elements that instantiate the constant symbols with some finite subset of $\mathbb{N}^{<\omega} = \bigcup_{i\in\mathbb{N}} \mathbb{N}^i$. Recall the infinitary theory $T_\infty = \bigcup_{i\in\mathbb{N}} T_i$. in the language $L_\infty = \bigcup_{i\in\mathbb{N}} L_i$. Fix an enumeration $\langle \varphi_i(\overline{x}_i, y)\rangle_{i\in\mathbb{N}}$ of all (quantifier-free) $\mathcal{L}_{\omega_1,\omega}(L_\infty)$-formulas such that each formula occurs infinitely often, and such that for each $i \in \mathbb{N}$,

$$(\forall \overline{x}_i)(\exists y)\varphi_i(\overline{x}_i, y) \in T_\infty$$

and the formula φ_i has precisely $(|\overline{x}_i| + 1)$-many free variables; let ξ_i denote $|\overline{x}_i|$. Let $\langle \overline{a_i} \rangle_{i \in \mathbb{N}}$ be an enumeration with repetition of finite tuples of elements of $\mathbb{N}^{<\omega}$ such that for all $i \in \mathbb{N}$, we have $|\overline{a_i}| = \xi_i$ and for every $\overline{a} \in (\mathbb{N}^{<\omega})^{\xi_i}$, there are infinitely many j such that $\varphi_j = \varphi_i$ and $\overline{a_j} = \overline{a}$. Also fix an arbitrary non-degenerate probability measure m^* on $\mathbb{N}$, i.e., such that no element has measure 0.

At the end of each finite Stage $n \in \mathbb{N}$, we will have constructed

- a finite set $X_n \subseteq \mathbb{N}^{2n}$ such that $\pi^2(X_n) \supseteq X_{n-1}$ (when $n \geq 1$),
- a measure m_n on X_n,
- some natural number $\alpha_n > \alpha_{n-1}$ (when $n \geq 1$),
- the complete non-redundant quantifier-free L_{α_n}-type of X_n (chosen from Q_{α_n}), and
- an L_{α_n}-structure $\mathcal{X}_n$.

In fact, X_0 will be empty and $\alpha_0 = 0$. We will define an L_0-structure $\mathcal{X}_0$, whose underlying set will be precisely a set of instantiations of the constant symbols in L_0. Call this set of instantiations C_0.

For all $n \in \mathbb{N}$, the L_{α_n}-structure $\mathcal{X}_n$ will have underlying set $X_n \cup C_0$, and hence is determined by the quantifier-free L_{α_n}-type of X_n. We call X_n the **constantless** part of $\mathcal{X}_n$.

For convenience of various indices, Stage 1 will not add anything essential to those objects constructed in Stage 0.

For $n \geq 2$ we will divide Stage n into substages $n.i$, indexed by $i \in \{0, 1, 2, 3\}$, each devoted to a different task: $n.0$ (adding mass), $n.1$ (adding witnesses), $n.2$ (duplication of quantifier-free types), and $n.3$ (expanding the language).

At the end of Stage $n.i$, for $i \in \{0, 1, 2\}$, we will have constructed

- a finite set X_n^i,
- a measure m_n^i on X_n^i,
- the complete non-redundant quantifier-free $L_{\alpha_{n-1}}$-type of X_n^i (chosen from $Q_{\alpha_{n-1}}$), and
- an $L_{\alpha_{n-1}}$-structure $\mathcal{X}_n^i$.

As with the major stages, each $L_{\alpha_{n-1}}$-structure $\mathcal{X}_n^i$ will have underlying set $X_n^i \cup C_0$, and hence will be determined by the quantifier-free $L_{\alpha_{n-1}}$-type of X_n^i. We similarly call X_n^i the constantless part of $\mathcal{X}_n^i$. Because each Substage $n.3$ completes Stage n, we write $\mathcal{X}_n$, X_n, and m_n rather than $\mathcal{X}_n^3$, X_n^3, and m_n^3, respectively.

Furthermore, the sets will satisfy

- $X_n^0 \subseteq X_n^1 \subseteq \mathbb{N}^{2n-2}$,
- $X_n^2 \subseteq \mathbb{N}^{2n-1}$ and $\pi(X_n^2) \supseteq X_n^1$, and
- $X_n \subseteq \mathbb{N}^{2n}$ and $\pi(X_n) \supseteq X_n^2$.

Finally, at the end of Stage ∞, we will have constructed an L_∞-structure $\mathcal{X}_\infty$ defined by the quantifier-free L_∞-type of each finite subset of the *infinite* constantless part $X_\infty \subseteq \mathbb{N}^\omega$ of $\mathcal{X}_\infty$, and a probability measure m_∞ on X_∞. The structure $\mathcal{X}_\infty$ may be viewed as a sort of inverse limit of the structures $\mathcal{X}_n$ for $0 \leq n < \infty$, with elements "glued together" in accordance with the projection map π.

We will also, at the end of each (sub)stage, verify that the new choices cohere with those made earlier. Specifically, they will satisfy the following *existence* and *duplication* properties for every $j \in \mathbb{N}$:

($\mathscr{E}$) If $\varphi_{j+1} \in \mathcal{L}_{\omega_1,\omega}(L_{\alpha_j})$, then for every tuple $\overline{s} = s_1, \ldots, s_{|\overline{a_{j+1}}|}$ of (not necessarily distinct) elements from X_j such that $\overline{a_{j+1}} \sqsubseteq \overline{s}$, and every $\ell_1, \ldots, \ell_{|\overline{a_{j+1}}|} \in \mathbb{N}^2$ such that $s_1{}^\wedge\ell_1, \ldots, s_{|\overline{a_{j+1}}|}{}^\wedge\ell_{|\overline{a_{j+1}}|} \in X_{j+1}$, we have

$$\mathcal{X}_{j+1} \models (\exists y)\varphi_{j+1}(s_1{}^\wedge\ell_1, \ldots, s_{|\overline{a_{j+1}}|}{}^\wedge\ell_{|\overline{a_{j+1}}|}, y).$$

($\mathscr{D}$) For all $g \in \mathbb{N}$, all distinct $s_1, \ldots, s_g \in \mathbb{N}^{2j}$, all $\ell_1, \ldots, \ell_g \in \mathbb{N}^2$, and all quantifier-free L_{α_j}-types r with g-many free variables, if $s_1, \ldots, s_g \in X_j$ and $s_1{}^\wedge\ell_1, \ldots, s_g{}^\wedge\ell_g \in X_{j+1}$ then

$$\mathcal{X}_j \models r(s_1, \ldots, s_g)$$

if and only if

$$\mathcal{X}_{j+1} \models r(s_1{}^\wedge\ell_1, \ldots, s_g{}^\wedge\ell_g).$$

Furthermore, for any $s \in X_j$, we have

$$m_j(s) = m_{j+1}\big((\pi^2)^{-1}(s) \cap X_{j+1}\big)$$

and

$$\lim_{i\to\infty} m_i(X_i) = 1.$$

In this sense, mass is preserved via projection throughout the construction.

We now make the construction precise.

Stage 0: Defining the mass on $\mathbb{N}$ and the quantifier-free type of the constants.

We begin by defining the constantless part $X_0 := \emptyset$. Let $\alpha_0 := 0$. Let m_0 be the unique measure on X_0, i.e., which satisfies $m_0(\emptyset) = 0$.

Choose an arbitrary element of Q_0 having no free variables. Because T_0 is quantifier-free complete, there is only one such choice of quantifier-free L_0-type (up to equivalence). This quantifier-free type describes which relations hold of any finite tuple of elements instantiating constant symbols. In particular, this determines when two constant symbols must be instantiated by the same element. Let $\mathcal{X}_0$ be an L_0-structure in which X_0 has this quantifier-free type, which amounts to choosing a set of instantiations of the constant symbols, related in this way. Let C_0 denote this set of instantiations, and let $\mathscr{C}_0$ be the map that assigns each constant symbol of L_0 to its instantiation in $\mathcal{X}_0$.

Stage 1: Same as stage 0.

Let $X_1 := X_0 = \emptyset$, let $\alpha_1 := 1$, and let m_1 be the unique measure on X_1. Let $\mathcal{X}_1$ be the unique L_1-structure whose reduct to L_0 is $\mathcal{X}_0$.

Stage $n.0$ (for $1 < n < \infty$): Adding mass.

Having already determined the $L_{\alpha_{n-1}}$-structure $\mathcal{X}_{n-1}$ and the measure m_{n-1}, we now define an $L_{\alpha_{n-1}}$-structure $\mathcal{X}_n^0$ extending $\mathcal{X}_{n-1}$, and the associated measure m_n^0. We will define the structure $\mathcal{X}_n^0$ by choosing its constantless part $X_n^0 \supseteq X_{n-1}$ and the quantifier-free $L_{\alpha_{n-1}}$-type of X_n^0.

This substage adds new elements of $\mathbb{N}^{2(n-1)}$ to the support of m_{n-1} so as to ensure that the eventual measure m_∞ will be a probability measure.

If there is an $x \in X_{n-1}$ with $x = n^\frown \overline{b}$ for some $\overline{b} \in \mathbb{N}^{2n-3}$, then let $\mathcal{X}_n^0 := \mathcal{X}_{n-1}$ be the same $L_{\alpha_{n-1}}$-structure, and let $m_n^0 := m_{n-1}$.

Otherwise let $X_n^0 := X_{n-1} \cup \{n^\frown 0^{2n-3}\}$ and fix some ordering on it. Let $q(\overline{x}, y) \in Q_{\alpha_{n-1}}$ be a quantifier-free type with $|X_n^0|$-many free variables such that if q^* is the quantifier-free type of X_{n-1} (considered as an increasing tuple in the corresponding ordering) in $\mathcal{X}_{n-1}$, then

$$\models (\forall \overline{x}, y)\big(q(\overline{x}, y) \to q^*(\overline{x})\big).$$

Note that such a q exists in $Q_{\alpha_{n-1}}$ by condition (D). Define the quantifier-free $L_{\alpha_{n-1}}$-type of X_n^0 in $\mathcal{X}_n^0$ (where X_n^0 is considered as an increasing tuple in that ordering) to be q. Finally, let $m_n^0(z) = m_{n-1}(z)$ for all $z \in X_{n-1}$ and $m_n^0(n^\frown 0^{2n-3}) = m^*(n)$, where m^* is the non-degenerate probability measure on $\mathbb{N}$ that we fixed before the construction.

In summary, at stage $n.0$, if no element of X_{n-1} is a sequence beginning with n, then we add one such sequence to our set, adjust the measure accordingly, and define the larger quantifier-free type appropriately. Note that $\mathcal{X}_{n-1}$ is a substructure of $\mathcal{X}_n^0$, and so the quantifier-free type of any tuple in $\mathcal{X}_{n-1}$ is the same as its quantifier-free type in $\mathcal{X}_n^0$. Furthermore, it is clear from the definition of m_n^0 that the measures m_n^0 and m_{n-1} agree on elements in the intersection of their domains.

Stage $n.1$ (for $1 < n < \infty$): Adding witnesses.

We now extend $\mathcal{X}_n^0$ to $\mathcal{X}_n^1$, in particular defining the quantifier-free $L_{\alpha_{n-1}}$-type of its constantless part $X_n^1 \supseteq X_n^0$ so as to ensure that certain subtuples have witnesses to appropriate formulas, and define the associated measure m_n^1.

Call $\varphi_n(\overline{a}_n, y)$ **valid** for Stage n when the following hold:

- $(\forall \overline{x}_n)(\exists y)\varphi_n(\overline{x}_n, y) \in \bigcup_{1 \le i \le n-1} T_{\alpha_i}$.
- At least one tuple $\overline{b}$ of elements of X_n^0 satisfies $\overline{a_n} \sqsubseteq \overline{b}$.

If $\varphi_n(\overline{a}_n, y)$ is not valid for Stage n then do nothing. Otherwise let V be the set of all $\overline{b} \in X_n^0$ such that $\overline{a_n} \sqsubseteq \overline{b}$ and

$$\mathcal{X}_n^0 \models \neg \bigvee_{d \in \overline{b}} \varphi_n(\overline{b}, d).$$

For each $\overline{b} \in V$, let $n_{\overline{b}} \in \mathbb{N}$ be such that for all $x \in X_n^0$ we have $n_{\overline{b}} \not\preceq x$. Then let $X_n^1 := X_n^0 \cup \{n_{\overline{b}}{}^\frown 0^{2n-3} : \overline{b} \in V\}$. Fix some ordering of X_n^0, and let q^* be the quantifier-free $L_{\alpha_{n-1}}$-type of X_n^0 (considered as an increasing tuple under this ordering) in $\mathcal{X}_n^0$. Choose a quantifier-free $L_{\alpha_{n-1}}$-type $q \in Q_{\alpha_{n-1}}$ such that if q holds of X_n^1 under some ordering, then

$$q^*(X_n^0) \wedge \bigwedge_{\overline{b} \in V} \varphi_n(\overline{b}, n_{\overline{b}}{}^\frown 0^{2n-3})$$

holds, where X_n^0 occurs in its ordering. Note that the formula $(\forall \overline{x}_n)(\exists y)\varphi_n(\overline{x}_n, y)$ is in T and q^* is consistent with T. Hence by condition (W) we can always find such a q. Declare q to be the quantifier-free $L_{\alpha_{n-1}}$-type of X_n^1 in $\mathcal{X}_n^1$ (under that ordering). In other words, we require that either there is a new witness or some witness already existed.

Finally, let m_n^1 agree with m_n^0 on X_n^0 and set $m_n^1(n_{\overline{b}}{}^\wedge 0^{2n-3}) := m^*(n_{\overline{b}})$ for $n_{\overline{b}}{}^\wedge 0^{2n-3} \in X_n^1 \setminus X_n^0$.

At this substage, we have ensured that if φ_n is valid (for stage n) then there are witnesses in $\mathcal{X}_n^1$ to $(\exists y)\varphi_n(\overline{b}, y)$ for all appropriate elements $\overline{b}$ of X_n^0. We will use this fact to verify property $(\mathscr{E})$ at the end of stage n.

Again, $\mathcal{X}_n^0$ is a substructure of $\mathcal{X}_n^1$, and so the quantifier-free type of any tuple in $\mathcal{X}_n^0$ is the same as its quantifier-free type in $\mathcal{X}_n^1$. Likewise, it is clear from the definition of m_n^1 that the measures m_n^1 and m_n^0 agree on elements in the intersection of their domains.

Stage $n.2$ (for $1 < n < \infty$): Duplication of Quantifier-Free Types.

Having defined $\mathcal{X}_n^1$ in the previous substage, we now define $\mathcal{X}_n^2$, in which we duplicate the quantifier-free type of X_n^1 in $\mathcal{X}_n^1$. We will define the structure $\mathcal{X}_n^2$ by choosing its constantless part $X_n^2 \supseteq X_n^1$ and the quantifier-free $L_{\alpha_{n-1}}$-type of X_n^2. We also define the associated measure m_n^2.

Let $\Lambda_n \in \mathbb{N}$ be large enough that if n balls are placed uniformly independently in Λ_n-many boxes, the probability of two or more balls landing in the same box is less than 2^{-n}.

Define

$$X_n^2 := \bigcup_{1 \le j \le \Lambda_n} \{x^\wedge j \,:\, x \in X_n^1\},$$

and fix an ordering of X_n^2. Fix an ordering $\langle x_i \rangle_{1 \le i \le |X_n^1|}$ of the elements of X_n^1, and let $q \in Q_{\alpha_{n-1}}$ be the quantifier-free type of this tuple. Choose a quantifier-free type $q^* \in Q_{\alpha_{n-1}}$ such that whenever q^* holds of X_n^2 (under its ordering) and any subset $\{y_i \,:\, 1 \le i \le |X_n^1|\} \subseteq X_n^2$ satisfies $\pi(y_i) = x_i$ for all i such that $1 \le i \le |X_n^1|$, then

$$q\big(\langle y_i \rangle_{1 \le i \le |X_n^1|}\big)$$

holds. Recall that our assumption (D) of duplication of quantifier-free types implies the existence of iterated duplicates. Hence there is such a q^*, as it is precisely an iterated duplicate of q. Declare q^* to be the quantifier-free type of X_n^2 (under its ordering) in $\mathcal{X}_n^2$.

Suppose that $g \in \mathbb{N}$ and $s_1, \dots, s_g \in X_n^1$ are distinct. Further suppose that $\ell_1, \dots, \ell_g \in \mathbb{N}$, such that $s_1{}^\wedge \ell_1, \dots, s_g{}^\wedge \ell_g \in X_n^2$. Then note that for any quantifier-free $L_{\alpha_{n-1}}$-type r with g-many free variables

$$\mathcal{X}_n^1 \models r(s_1, \dots, s_g)$$

if and only if

$$\mathcal{X}_n^2 \models r(s_1{}^\wedge \ell_1, \dots, s_g{}^\wedge \ell_g).$$

This is the analogue, for the situation of moving from substage $n.1$ to $n.2$, of property $(\mathscr{D})$.

Finally, for each $x \in X_n^2$, define $m_n^2(x) := m_n^1(\pi(x))/\Lambda_n$. In other words, for each $y \in X_n^1$, its mass is divided evenly between its Λ_n-many extensions.

Stage $n.3$ (for $1 < n < \infty$): Expanding the Language.

Having defined $\mathscr{X}_n^2$ in the previous substage, we now define $\mathscr{X}_n$ itself, some $\alpha_n > \alpha_{n-1}$, and the associated measure m_n. We will define by $\mathscr{X}_n$ via its constantless part $X_n \supseteq X_n^2$ and the quantifier-free L_{α_n}-type of X_n. We do this in a way that ensures that if, for some $\ell \in \mathbb{N}$ such that $\ell \leq |X_n^2|$, there is splitting of quantifier-free types of order ℓ, then for the least such ℓ, as we enlarge the language we split all non-redundant quantifier-free types with ℓ-many free variables.

Fix some ordering on X_n^2 and let p_{n-1} be the quantifier-free $L_{\alpha_{n-1}}$-type of X_n^2 (considered as an increasing tuple under that ordering) in $\mathscr{X}_n^2$.

Case (a): If either there is no splitting of quantifier-free types of order ℓ for any $\ell \in \mathbb{N}$, or there is a splitting, but the least such order ℓ is greater than $|X_n^2|$, then let $\alpha_n := \alpha_{n-1} + 1$, let $X_n := \{x^\wedge 0 : x \in X_n^2\}$ and let $p_n \in Q_{\alpha_n}$ be any non-redundant quantifier-free L_{α_n}-type with $|X_n^2|$-many free variables such that

$$\models (\forall \overline{x})\big(p_n(\overline{x}) \to p_{n-1}(\overline{x})\big)$$

where $|\overline{x}| = |X_n^2|$. We know that such a quantifier-free type exists by condition (E). Then declare p_n to be the quantifier-free L_{α_n}-type of X_n (considered as an increasing ordered tuple under the order induced from X_n^2) in $\mathscr{X}_n$. For $x \in X_n$, define $m_n(x) := m_n^2(\pi(x))$, since every element has just one extension.

Case (b): If, however, there is splitting of some order, i.e., condition (S) holds, and the least such order $\ell \in \mathbb{N}$ is no greater than $|X_n^2|$, then let $q^\natural$ be some splitting of p_{n-1} of order ℓ. Let α_n be the $e \in \mathbb{N}$ such that $q^\natural$ is a quantifier-free L_e-type.

Define

$$X_n := \{x^\wedge 0 : x \in X_n^2\} \cup \{x^\wedge 1 : x \in X_n^2\}.$$

and declare that $q^\natural$ is the quantifier-free L_{α_n}-type of X_n in $\mathscr{X}_n$, where X_n is considered as the tuple

$$x_1{}^\wedge 0, x_1{}^\wedge 1, \dots, x_{|X_n^2|}{}^\wedge 0, x_{|X_n^2|}{}^\wedge 1$$

where $x_1, \dots, x_{|X_n^2|}$ is increasing in the chosen order of X_n^2.

Finally, for each $x \in X_n$, define $m_n(x) := m_n^2(\pi(x))/2$. In other words, each element of X_n^2 has its mass divided evenly between its two extensions. This concludes case (b).

Now, regardless of the case, we verify property $(\mathscr{D})$ for stage n. Suppose that $g \in \mathbb{N}$ and $s_1, \dots, s_g \in X_n^2$ are distinct. Further suppose that $\ell_1, \dots, \ell_g \in \mathbb{N}$, such that $s_1{}^\wedge \ell_1, \dots, s_g{}^\wedge \ell_g \in X_n$. Then note that for any quantifier-free $L_{\alpha_{n-1}}$-type r with g-many free variables

$$\mathscr{X}_n^2 \models r(s_1, \dots, s_g)$$

if and only if

$$\mathscr{X}_n \models r(s_1{}^\wedge \ell_1, \dots, s_g{}^\wedge \ell_g).$$

Note that this property, composed with the analogous property verified at the end of substage $n.2$, guarantees that $(\mathscr{D})$ holds.

Finally, for $n > 0$ note that, by property $(\mathscr{D})$, if $\varphi_n \in \mathcal{L}_{\omega_1,\omega}(L_{\alpha_{n-1}})$, then for every tuple $s_1, \ldots, s_{|a_n|} \in X_n$ with $\overline{a_n} \sqsubseteq \big(\pi^2(s_1), \ldots, \pi^2(s_{|a_n|})\big)$, there is an element $t \in X^1_n$ such that

$$\mathcal{X}^1_n \models \varphi_n\big(\pi^2(s_1), \ldots, \pi^2(s_{|a_n|}), t\big).$$

Hence if $t^* \in (\pi^2)^{-1}(t) \cap X_n$, then

$$\mathcal{X}_n \models \varphi_n(s_1, \ldots, s_{|a_n|}, t^*).$$

This verifies property $(\mathscr{E})$.

Stage ∞: Defining the Limiting Structure.

To complete the construction, we define the L_∞-structure $\mathcal{X}_\infty$ via its constantless part X_∞ and the quantifier-free L_∞-type of every finite subset of X_∞. We also define the measure m_∞.

Let

$$X_\infty := \{x \in \mathbb{N}^\omega \,:\, (\forall i \in \mathbb{N})\,(x|_i \in X_i)\},$$

and for each $n \in \mathbb{N}$ and each $y \in X_n$ define

$$m_\infty\big(\{x \in X_\infty \,:\, x|_n = y\}\big) := m_n(y).$$

Consider X_∞ endowed with the topology inherited as a subspace of $\mathbb{N}^\omega$ (itself under the product topology of $\mathbb{N}$ as a discrete set). Then X_∞ is the countable disjoint union $\bigcup_{\ell \in \mathbb{N}} Y_\ell$, where for each $\ell \in \mathbb{N}$,

$$Y_\ell := \{\ell^\wedge a \,:\, a \in \mathbb{N}^\omega \text{ and } \ell^\wedge a \in X_\infty\}$$

is a compact topological space having a basis of clopen sets, under the topology inherited as a subspace of $\mathbb{N}^\omega$. Hence m_∞ can be extended in a unique way to a countably additive measure on X_∞.

For every $j, n \in \mathbb{N}$ and every $s_1, \ldots, s_j \in X_\infty$, there are some $n' \geq n$ and $t_1, \ldots t_j \in X_{n'}$ such that for all distinct $i, i' \leq j$,

- $t_i = s_i|_{2n'}$ and
- $t_i \neq t_{i'}$.

Let $q \in Q_{\alpha_n}$ be the quantifier-free L_{α_n}-type such that

$$\mathcal{X}_n \models q(t_1, \ldots, t_j).$$

Then declare that

$$\mathcal{X}_\infty \models q(s_1, \ldots, s_j)$$

holds. This choice of quantifier-free type is well-defined because of property $(\mathscr{D})$ at all earlier stages. This ends the construction.

4.3. Invariant measures via the construction

We now verify properties of $\mathcal{X}_\infty$ and m_∞ that will allow us to produce the desired invariant measure.

Proposition 4.1. *The measure m_∞ on X_∞ is a non-degenerate atomless probability measure.*

Proof. The measures m_n for $n \in \mathbb{N}$ cohere under projection and agree with m^*, in the sense that

$$m_\infty(Y_n) = m^*(n).$$

But m^* is a probability measure, and so m_∞ is as well.

For $n \in \mathbb{N}$ and $a \in X_n$, let

$$B_a := \{s \in X_\infty : s|_n = a\}.$$

The collection of sets of the form B_a form a basis for the topological space X_∞. Furthermore, for all $n \in \mathbb{N}$ and $a \in X_n$,

$$m_\infty(B_a) = m_n(a) > 0.$$

Hence m_∞ is non-degenerate.

For each $n \in \mathbb{N}$, define $\Gamma_n := \max\{m_n(a) : a \in X_n\}$; in substage $n.2$, we duplicate every element of X_{n-1}, and so

$$\Gamma_n \leq \Gamma_{n-1}/2.$$

Consider a singleton $\{b\} \subseteq X_\infty$. Then

$$m_\infty(\{b\}) \leq m_\infty(B_{b|_n}) \leq \Gamma_n$$

for each $n \in \mathbb{N}$, and so $m_\infty(\{b\}) = 0$. Hence m_∞ is atomless. $\square$

We will show that $\mathcal{X}_\infty$ is an uncountable Borel model such that when we sample countably infinitely many elements from X_∞ independently according to the probability measure m_∞, the induced substructure is almost surely a model of T_∞.

Proposition 4.2. *The structure $\mathcal{X}_\infty$ is a Borel L_∞-structure.*

Proof. Fix $n \in \mathbb{N}$. Let ψ be a quantifier-free L_{α_n}-formula, and let ℓ be the number of free variables of ψ. Then define the set of its instantiating tuples:

$$\Psi := \{a_1 \cdots a_\ell \in X_\infty : \mathcal{X}_\infty \models \psi(a_1, \ldots, a_\ell)\}.$$

Also define, for each $n' \geq n$,

$$P_{n'} := \big\{a_1 \cdots a_\ell \in X_\infty : \mathcal{X}_{n'} \models \psi(a_1|_{2n'}, \ldots, a_\ell|_{2n'})\big\}$$

and

$$I_{n'} := \big\{a_1 \cdots a_\ell \in X_\infty : a_i = a_j \text{ iff } a_i|_{2n'} = a_j|_{2n'}, \text{ whenever } 1 \leq i \leq j \leq \ell\big\}.$$

Note that for each $n' \geq n$, both $P_{n'}$ and $I_{n'}$ are open sets. We then have

$$\Psi = \bigcup_{n' \geq n} \big(P_{n'} \cap I_{n'}\big),$$

and so Ψ is an open set. As ψ was arbitrary, $\mathcal{X}_\infty$ is a Borel L_∞-structure. $\square$

A natural procedure for sampling substructures of $\mathcal{X}_\infty$ using m_∞ will yield the desired invariant measure.

Because T_0 is quantifier-free complete, all models of T_∞ have the same number of elements that instantiate constant symbols, and the theory of equality between constants is fixed (as encoded in $\mathscr{C}_0$).

Let μ be an arbitrary atomless probability measure on $\mathcal{X}_\infty$. We begin by describing a sampling procedure that uses μ to determine an invariant measure μ° on $\mathrm{Str}_{\mathscr{C}_0,L_\infty}$: First sample a countably infinite sequence of elements $\langle x_i\rangle_{i\in\mathbb{N}}$ from X_∞ independently according to μ. If there exist distinct $i, j \in \mathbb{N}$ such that $x_i = x_j$, then declare that all atomic relations hold among all tuples; however, this occurs with probability 0, as μ is atomless. Otherwise, for each quantifier-free L_∞-formula ψ, declare that

$$\psi(n_1, \ldots, n_\ell)$$

holds if and only if

$$\mathcal{X}_\infty \models \psi(x_{n_1}, \ldots, x_{n_\ell})$$

for all $n_1, \ldots, n_\ell \in \mathbb{N}$, where ℓ is the number of free variables of ψ. The distribution of this random L_∞-structure is a probability measure on $\mathrm{Str}_{\mathscr{C}_0,L_\infty}$; this is our desired μ°. (As with the measures described via sampling in Section 3, such probability measures are ergodic, as Kallenberg showed by extending the argument of Aldous [22, Lemma 7.35] to languages of unbounded arity in [22, Lemma 7.22] and [22, Lemma 7.28 (iii)].) Note that μ° is $S_\infty^{C_0}$-invariant, as $\langle x_i\rangle_{i\in\mathbb{N}}$ is i.i.d. Because μ is atomless, μ° is concentrated on the class of structures with underlying set $\mathbb{N} \cup C_0$ that are isomorphic to countably infinite substructures of $\mathcal{X}_\infty$.

Proposition 4.3. *The $S_\infty^{C_0}$-invariant probability measure m_∞° on $\mathrm{Str}_{\mathscr{C}_0,L_\infty}$ is concentrated on the class of models of T_∞.*

Proof. By Proposition 4.1, the measure m_∞ is atomless, and so m_∞° is an $S_\infty^{C_0}$-invariant probability measure on $\mathrm{Str}_{\mathscr{C}_0,L_\infty}$ that is concentrated on the class of countably infinite substructures of $\mathcal{X}_\infty$.

Now let $\mathcal{M}$ be a sample from m_∞°, say via the m_∞-i.i.d. sequence $\langle x_i\rangle_{i\in\mathbb{N}}$ of elements of X_∞. Fix an arbitrary $\eta \in T_\infty$. We will show that $\mathcal{M} \models \eta$ almost surely. Because η is pithy Π_2, we may write it in the form $(\forall \overline{z})(\exists y)\psi(\overline{z}, y)$ for some quantifier-free L_∞-formula ψ. Let $\ell = |\overline{z}|$ and let $n \in \mathbb{N}$ be such that $\psi \in L_{\alpha_n}$. Fix an arbitrary tuple $\overline{b} := b_1 \cdots b_\ell \in \mathbb{N}$. We must show that there is some $d \in \mathbb{N}$ such that

$$\mathcal{M} \models \psi(\overline{b}, d) \quad \text{a.s.}$$

Let $\overline{b^*}$ be the random tuple $x_{b_1} \cdots x_{b_\ell}$. Let $j > n$ be any index of ψ (i.e., such that $\psi = \varphi_j$) satisfying $\overline{a_j} \sqsubseteq \overline{b^*}$. This is possible because of our choice of repetitive enumeration.

By our construction in stage $j.2$, there is some $e \in X_j^2$ such that

$$\mathcal{X}_j^2 \models \psi(x_{b_1}|_{2j-2} \cdots x_{b_\ell}|_{2j-2}, e) \quad \text{a.s.}$$

As in the proof of Proposition 4.1, let

$$B_e := \{s \in X_\infty \,:\, s|_{2j-2} = e\}.$$

By our construction, for any $e^* \in B_e$,

$$\mathcal{X}_\infty \models \psi(x_{b_1} \cdots x_{b_\ell}, e^*) \quad \text{a.s.}$$

However, $m_\infty(B_e) > 0$, and so there is some $h \in \mathbb{N}$ such that $x_h \in B_e \cap \mathcal{M}$, almost surely. Hence

$$\mathcal{M} \models \psi(\overline{b}, h) \quad \text{a.s.}$$

Again by Proposition 4.1, the measure m_∞ is non-degenerate. □

We now show that if the collection of quantifier-free types has splitting of some order, the resulting construction assigns measure 0 to any particular isomorphism class of models of the theory T_∞.

Theorem 4.4. *Suppose that $\langle Q_i \rangle_{i \in \mathbb{N}}$ has splitting of some order. Then there is an $S_\infty^{C_0}$-invariant probability measure on $\mathrm{Str}_{\mathscr{C}_0, L}$ that is concentrated on the class of models of T_∞ and is such that no single isomorphism class has positive measure.*

Proof. Let $\ell \in \mathbb{N}$ be least such that $\langle Q_i \rangle_{i \in \mathbb{N}}$ has splitting of order ℓ. Let m' be the $S_\infty^{C_0}$-invariant probability measure obtained in Proposition 4.3. Define M to be the collection of isomorphism classes of countably infinite models of T_∞ to which m' assigns positive measure.

Suppose, to obtain a contradiction, that $M \neq \emptyset$. Then by the countable additivity of m', there can be at most countably many elements of M. Hence among the quantifier-free L_∞-types with ℓ-many free variables, at most countably many are realized in some structure in M. In particular, at most countably many *non-constant* quantifier-free L_∞-types with ℓ-many free variables are realized in some structure in M. Then by countable additivity, there must be some non-constant quantifier-free L_∞-type p with ℓ-many free variables that is realized in a positive fraction of models, i.e., such that

$$m'\big([\![(\exists \overline{x}) p(\overline{x})]\!]_{\mathscr{C}_0}\big) > 0,$$

where $|x| = \ell$.

We then have

$$0 < m'\big([\![(\exists \overline{x}) p(\overline{x})]\!]_{\mathscr{C}_0}\big) = m'\Big(\bigcup_{\overline{t} \in \mathbb{N}^\ell} [\![p(\overline{t})]\!]_{\mathscr{C}_0}\Big) \leq \sum_{\overline{t} \in \mathbb{N}^\ell} m'\big([\![p(\overline{t})]\!]_{\mathscr{C}_0}\big),$$

where the equality is because p is non-constant. Hence there is some $\overline{t} \in \mathbb{N}^\ell$ such that $m'\big([\![p(\overline{t})]\!]_{\mathscr{C}_0}\big) > 0$, by the countable additivity of m'.

For every $i \in \mathbb{N} \cup \{\infty\}$, define $\eta_i := m'\big([\![p|_{L_{\alpha_i}}(\overline{t})]\!]_{\mathscr{C}_0}\big)$. Because

$$L_0 \subseteq L_1 \subseteq \cdots \subseteq L_\infty,$$

we have $\eta_i \geq \eta_j$ whenever $0 \leq i < j \leq \infty$.

Let $g \geq \ell$ be arbitrary. We will show that

$$\eta_g \leq 2^{-g} + (1 - 2^{-\ell})^{g-\ell}.$$

This will imply that $\eta_\infty \leq \inf_i (1 - 2^{-\ell})^{2i} = 0$, and so $m'\big([\![p(\overline{t})]\!]_{\mathscr{C}_0}\big) = 0$, a contradiction.

There are two (overlapping) ways that an ℓ-tuple of elements of X_∞ sampled independently according to m' can fail to satisfy $p|_{L_{\alpha_g}}$: either (1) the restriction of the tuple to $\mathbb{N}^{2g}$ satisfies a redundant quantifier-free type, in which case the tuple might not satisfy $p|_{L_{\alpha_g}}$, or (2) its restriction to $\mathbb{N}^{2g}$ is non-redundant but satisfies some quantifier-free type other than $p|_{L_{\alpha_g}}$.

By our choice of Λ_g in stage $g.2$, we know that for any assignment of mass to X_g^1, the probability of an independently selected ℓ-tuple having two elements selected from the same element of X_g^2 is no more than 2^{-g}, as $g \geq \ell$. Hence the probability that (1) occurs is bounded by 2^{-g}.

Because the mass of every element is split evenly between those elements descending from it via iterated duplication, the probability that a given non-redundant ℓ-tuple of X_g^2 is selected independently according to m_g^2 is 2^ℓ times the probability that any of such duplicated elements are selected independently according to m_g.

Let ζ_g be the probability that a given ℓ-tuple, independently selected from X_g according to m_g, has quantifier-free type $p|_{L_{\alpha_g}}$ conditioned on the fact each element of the ℓ-tuple is distinct (i.e., ζ_g is a bound on the probability that (2) occurs, so that $\eta_g \leq 2^{-g} + \zeta_g$). By the splitting of quantifier-free types in stage $g.3$, we know that for every ℓ-tuple in X_g^2 there are at least two quantifier-free L_{α_g}-types of duplicates of the ℓ-tuple.

Hence we have

$$\zeta_g \leq (1 - 2^{-\ell}) \cdot \zeta_{g-1} \leq (1 - 2^{-\ell})^{g-\ell}.$$

In total, we have $\eta_g \leq 2^{-g} + (1 - 2^{-\ell})^{g-\ell}$. □

5. Approximately $\aleph_0$-categorical theories

In this section, we introduce several conditions on first-order theories that together allow us to apply Theorem 4.4. These will give us an invariant probability measure that is concentrated on the class of models of a theory, but does not assign positive measure to any single isomorphism class of models. We then give examples of first-order theories satisfying these conditions.

Key among these conditions is a property that we call *approximate* $\aleph_0$-*categoricity*.

Definition 5.1. Let L be a countable language. A first-order theory $T \subseteq \mathcal{L}_{\omega,\omega}(L)$ is **approximately** $\aleph_0$**-categorical** when there is a sequence of languages $\langle L_i \rangle_{i\in\mathbb{N}}$, called a **witnessing sequence**, such that

- $L_i \subseteq L_{i+1}$ for all $i \in \mathbb{N}$,
- $L = \bigcup_{i\in\mathbb{N}} L_i$, and
- $T \cap \mathcal{L}_{\omega,\omega}(L_i)$ is $\aleph_0$-categorical for each $i \in \mathbb{N}$.

In particular, any approximately $\aleph_0$-categorical theory is the countable union of $\aleph_0$-categorical first-order theories (in different languages).

We now give criteria under which the class of models of an approximately $\aleph_0$-categorical theory admits an invariant probability measure that assigns measure 0 to any single isomorphism class of models.

Recall the notion of a pithy Π_2 expansion from §2.2. Note that any model of a first-order L-theory T has a unique expansion to a model of its pithy Π_2 expansion. Furthermore, any invariant measure concentrated on a Borel set $X \subseteq \mathrm{Str}_L$ can be expanded uniquely to an invariant measure concentrated on

$$\{\mathcal{M}^* \in \mathrm{Str}_{L_{\mathrm{HF}}} : \mathcal{M}^*|_L \in X\}.$$

Lemma 5.2. *Let L be a countable language, and suppose that T is an approximately $\aleph_0$-categorical $\mathcal{L}_{\omega,\omega}(L)$-theory with witnessing sequence $\langle L_i \rangle_{i\in\mathbb{N}}$. Then the pithy Π_2 expansion T^* of T is also approximately $\aleph_0$-categorical.*

Proof. For each $i \in \mathbb{N}$, the L_i-theory $T \cap \mathcal{L}_{\omega,\omega}(L_i)$ is $\aleph_0$-categorical by hypothesis. For each i, let L_i^* be the language of the pithy Π_2 expansion T_i^* of $T \cap \mathcal{L}_{\omega,\omega}(L_i)$. Then each T_i^* is $\aleph_0$-categorical. Note that $T^* \cap \mathcal{L}_{\omega,\omega}(L_i^*) = T_i^*$ for each $i \in \mathbb{N}$, and $\langle L_i^* \rangle_{i\in\mathbb{N}}$ is a nested sequence whose union is the language of T^*. Hence T^* is approximately $\aleph_0$-categorical with witnessing sequence $\langle L_i^* \rangle_{i\in\mathbb{N}}$. $\square$

The following result is now straightforward from Theorem 4.4.

Theorem 5.3. *Let L be a countable relational language, and suppose that T is an approximately $\aleph_0$-categorical $\mathcal{L}_{\omega,\omega}(L)$-theory with witnessing sequence $\langle L_i \rangle_{i\in\mathbb{N}}$. For each $i \in \mathbb{N}$, let Q_i be any enumeration of the quantifier-free L_i-types that are consistent with $T \cap \mathcal{L}_{\omega,\omega}(L_i)$. Further suppose that*

- *for each $i \in \mathbb{N}$, the age of the unique countable model (up to isomorphism) of $T \cap \mathcal{L}_{\omega,\omega}(L_i)$ has the strong amalgamation property, and*
- *the sequence $\langle Q_i \rangle_{i\in\mathbb{N}}$ has splitting of some order.*

Then there is an S_∞-invariant probability measure on Str_L that is concentrated on the class of models of T but that assigns measure 0 to each isomorphism class of models.

Proof. By Lemma 5.2, the pithy Π_2 expansion T^* of T is approximately $\aleph_0$-categorical. Note that for each $i \in \mathbb{N}$, every element of Q_i is consistent with the pithy Π_2 expansion of $T \cap \mathcal{L}_{\omega,\omega}(L_i)$. We may therefore run the construction of §4.2, under the assumption that conditions (W), (D), (E), and (C) hold of $\langle Q_i \rangle_{i\in\mathbb{N}}$. Under the further assumption that (S) holds of $\langle Q_i \rangle_{i\in\mathbb{N}}$, we may apply Theorem 4.4 to obtain an invariant measure on $\mathrm{Str}_{L_{\mathrm{HF}}}$ that is concentrated on the class of models of T^* but that assigns measure 0 to each isomorphism class. The restriction of this invariant measure to Str_L will give us an invariant measure with the desired properties. We now show that these five conditions hold of $\langle Q_i \rangle_{i\in\mathbb{N}}$.

Condition (D) follows from our first hypothesis, and (S) from our second.

Conditions (E) and (C) hold of $\langle Q_i \rangle_{i\in\mathbb{N}}$ because for each $i \in \mathbb{N}$, the set Q_i contains *every* quantifier-free L_i-type that is consistent with $T \cap \mathcal{L}_{\omega,\omega}(L_i)$.

Finally, we show condition (W). Note that any pithy Π_2 sentence

$$(\forall \overline{x})(\exists y)\psi(\overline{x}, y) \in T$$

is an L_n-formula for some $n \in \mathbb{N}$. Hence as Q_n is consistent with $T \cap \mathcal{L}_{\omega,\omega}(L_n)$, for any quantifier-free L_n-type $q \in Q_n$, there is some $q' \in Q_n$ extending q such that for every tuple $\overline{z}$ of free variables of q having size $|\overline{x}|$,

$$\models (\forall \overline{w})\big(q'(\overline{w}) \to (\exists y)\psi(\overline{z}, y)\big)$$

holds, where $|\overline{w}|$ is the number of free variables of q'. Therefore condition (W) holds of $\langle Q_i \rangle_{i\in\mathbb{N}}$. $\square$

In particular, a theory satisfying the hypotheses of Theorem 5.3 is not itself $\aleph_0$-categorical, as it must have uncountably many countable models. We now use this theorem to give examples of an invariant measure that is concentrated on the class of models of a first-order theory but but that assigns measure 0 to each isomorphism class of models.

5.1. Kaleidoscope theories

Here we show a simple way in which a Fraïssé limit whose age has the strong amalgamation property gives rise to an approximately $\aleph_0$-categorical theory, which we call its corresponding *Kaleidoscope* theory, whose countable models consist of countably infinitely many copies of the Fraïssé limit combined in an appropriate way. Furthermore, we show that if such a Fraïssé limit satisfies the mild condition that for some finite size its age has at least two non-equal structures of that size (not necessarily non-isomorphic), then its Kaleidoscope theory satisfies the hypotheses of Theorem 5.3.

Definition 5.4. Suppose L is a countable relational language. Let $\langle L^j \rangle_{j\in\mathbb{N}}$ be an infinite sequence of pairwise disjoint copies of L such that $L^0 = L$, and for $i \in \mathbb{N}$, define $L_i := \bigcup_{0\le j\le i} L^j$.

Lemma 5.5. *Let L be a countable relational language, and let A be a strong amalgamation class of L-structures. For each $i \in \mathbb{N}$, define A_i to be the class of all finite L_i-structures $\mathcal{M}$ such that for $0 \le j \le i$, the reduct $\mathcal{M}|_{L^j}$ (when considered as an L-structure) is in A. Then each A_i is a strong amalgamation class.*

Proof. Each A_i satisfies the strong amalgamation property: Suppose $\mathcal{M}, \mathcal{N} \in A_i$ have a common substructure $\mathcal{O} \in A_i$. For each j such that $0 \le j \le i$, let $\mathcal{X}^j$ be a strong amalgam of $\mathcal{M}|_{L^j}$ and $\mathcal{N}|_{L^j}$ over $\mathcal{O}|_{L^j}$. Because $\mathcal{X}^0, \dots, \mathcal{X}^i$ are in disjoint languages and have the same underlying set, there is an L_i-structure $\mathcal{X}$ on this underlying set such that for $0 \le j \le i$, we have $\mathcal{X}|_{L^j} = \mathcal{X}^j$. Hence $\mathcal{X} \in A_i$ is a strong amalgam of $\mathcal{M}, \mathcal{N}$ over $\mathcal{O}$.

Each A_i is a class containing countably many isomorphism types, for which the hereditary property holds trivially. Further, the joint embedding property holds by a similar argument to that above. Thus each A_i is a strong amalgamation class. □

Definition 5.6. Using the notation of Lemma 5.5, for each $i \in \mathbb{N}$, let T_i be the theory of the Fraïssé limit of A_i, and notice that $T_i \subseteq T_{i+1}$. The theory $T_\infty := \bigcup_{i\in\mathbb{N}} T_i$ in the language $L_\infty := \bigcup_{i\in\mathbb{N}} L_i = \bigcup_{j\in\mathbb{N}} L^j$ is therefore consistent. The theory T_∞ is said to be the **Kaleidoscope theory** built from A.

Proposition 5.7. *Let L be a countable relational language, and let A be a strong amalgamation class of L-structures. Let T_∞, in the language L_∞, be the Kaleidoscope theory built from A, as above. Then T_∞ is approximately $\aleph_0$-categorical.*

Furthermore, suppose that for some $n \in \mathbb{N}$, the age A has at least two non-equal elements of size n on the same underlying set. (Note that we do not require these elements to be non-isomorphic.) Then there is an S_∞-invariant probability measure on Str_{L_∞} that is concentrated on the class of models of T_∞ but that assigns measure 0 to each isomorphism class of models.

Proof. For each $i \in \mathbb{N}$, let A_i be as defined in Lemma 5.5; then A_i is the age of a model of T_i, which is an $\aleph_0$-categorical L_i-theory. Therefore T_∞ is an approximately $\aleph_0$-categorical L_∞-theory with witnessing sequence $\langle L_i \rangle_{i\in\mathbb{N}}$.

We will apply Theorem 5.3 to obtain the desired invariant measure. We must show its two hypotheses: the strong amalgamation property for the age of each $T_\infty \cap \mathcal{L}_{\omega,\omega}(L_i)$, and that $\langle Q_i \rangle_{i\in\mathbb{N}}$ (as defined in Theorem 5.3) has splitting of some order.

For any $i \in \mathbb{N}$, because A_i is the age of the unique model of $T_i = T_\infty \cap \mathcal{L}_{\omega,\omega}(L_i)$, we may apply Lemma 5.5 to see that A_i is a strong amalgamation class as well.

We now show that $\langle Q_i \rangle_{i\in\mathbb{N}}$ has splitting of order n. Fix $j \in \mathbb{N}$, and let $q \in Q_j$ be a non-redundant quantifier-free L_j-type with k-many free variables, for some $k > n$. It suffices to find, for some $j' > j$, a quantifier-free type $q^\natural \in Q_{j'}$ with free variables $\overline{x} := x_1^0, x_1^1, \ldots, x_k^0, x_k^1$ such that the restriction $q^\natural$ to L_j is an iterated duplicate of q, and for any $2n$-tuple $y_1 \cdots y_n z_1 \cdots z_n$ of distinct free variables of $q^\natural$, we have

$$q^\natural|_{y_1,\ldots,y_n} \neq q^\natural|_{z_1,\ldots,z_n},$$

which ensures that $q^\natural$ is a splitting of q. We construct $q^\natural$ in the following manner.

In languages $L^0, \ldots, L^j$, the quantifier-free type $q^\natural$ describes an iterated duplicate of q; each of the remaining languages $L^{j+1}, \ldots, L^{j'}$, corresponds to a particular way of choosing a $2n$-tuple of variables from the $2k$-tuple $\overline{x}$, and describes a pair of different n-element structures on this $2n$-tuple. Let q^* be the quantifier-free L_j-type with free variables $\overline{x}$ that is an iterated duplicate of q. Let B_0 and B_1 be two non-equal elements of A of size n on the same underlying set $\{0, \ldots, n-1\}$, and let $p_0, p_1 \in Q_0$ be quantifier-free L-types such that

$$B_i \models p_i(0, \ldots, n-1)$$

for $i \in \{0, 1\}$. By the joint embedding property of A, let $p \in Q_0$ be any quantifier-free L-type with $2n$-many free variables $v_1, \ldots, v_n, w_1, \ldots, w_n$ such that

$$p(\overline{v}, \overline{w}) \rightarrow p_0(\overline{v}) \wedge p_1(\overline{w}),$$

where $\overline{v} := v_1 \cdots v_n$ and $\overline{w} := w_1 \cdots w_n$.

Enumerate all $2n$-tuples of distinct variables of $\overline{x}$. Assign each such tuple $\overline{u}$ a distinct value

$$j_{\overline{u}} \in [j+1, \ldots, j'],$$

where $j' := j + (2k)(2k-1) \cdots (2k-2n+1)$. For each such tuple $\overline{u}$, choose a quantifier-free L-type $q_{\overline{u}}$ with free variables $\overline{x}$ such that

$$\models (\forall \overline{x})\big(q_{\overline{u}}(\overline{x}) \rightarrow p(\overline{u})\big).$$

Let $q^\natural$ be a quantifier-free $L_{j'}$-type with free variables $\overline{x}$ that implies $q^*(\overline{x})$ and that also implies, for each such tuple $\overline{u}$, that $q_{\overline{u}}^{L^{j_{\overline{u}}}}(\overline{x})$ holds, where $q_{\overline{u}}^{L^{j_{\overline{u}}}}$ describes in language $L^{j_{\overline{u}}}$ what $q_{\overline{u}}$ describes in L. Note that we can find such a $q^\natural$ because the restrictions of T_∞ to each copy of L do not interact with each other. Finally, because $p(\overline{v}, \overline{w}) \neq p(\overline{w}, \overline{v})$, for any $2n$-tuple $y_1 \cdots y_n z_1 \cdots z_n$ of distinct free variables of $q^\natural$, we have that

$$q^\natural|_{y_1,\ldots,y_n} \neq q^\natural|_{z_1,\ldots,z_n}.$$

Therefore $\langle Q_i \rangle_{i\in\mathbb{N}}$ has splitting of order n. □

A key example of this construction is provided by what we call the *Kaleidoscope random graphs*, which are the countable models of the Kaleidoscope theory built from the class of finite graphs (in the language of graphs). There are continuum-many Kaleidoscope random graphs (up to isomorphism). Each Kaleidoscope random graph G can be thought of as countably many random graphs (i.e., Rado graphs), each with a different color for its edge-set, overlaid on the same vertex-set in such a way that for every finite substructure F of G and any chosen finite set of colors, there is an extension of F by a single vertex v of G satisfying any given assignment of edges and non-edges in those colors between v and the vertices of F.

The invariant measures provided by Proposition 5.7 are fundamentally different from those obtained in [5]. No measure provided by Proposition 5.7 is concentrated on the isomorphism class of a single structure, nor is any such measure concentrated on a class of structures having trivial definable closure. To see this, consider such a measure, and suppose $n \in \mathbb{N}$ is such that the age A has at least two elements of size n. Then for a structure sampled from the invariant measure, with probability 1 the tuple $0, \ldots, n-1$ has a quantifier-free type different from that of every other n-tuple in the structure. Hence the structures sampled from such a measure almost surely do not have trivial definable closure. As a consequence of this and the main result of [5], for almost every structure sampled from this measure, there is no invariant measure concentrated on the isomorphism class of just that structure.

5.2. Urysohn space

The *Urysohn space* $\mathbb{U}$ is the universal ultrahomogeneous Polish space. In other words, up to isomorphism (i.e., bijective isometry), $\mathbb{U}$ is the unique complete separable metric space that is *universal*, in that $\mathbb{U}$ contains an isomorphic copy of every complete separable metric space, and *ultrahomogeneous*, in that every isomorphism between two finite subsets of $\mathbb{U}$ can be extended to an isomorphism of the entire space $\mathbb{U}$.

Although Urysohn's work predates that of Fraïssé [24], his construction of $\mathbb{U}$ can be viewed as a continuous generalization of the Fraïssé method. Hušek [25] describes Urysohn's original construction [26] and its history, and Katětov's more recent generalizations [27]. For further background, see the introductory remarks in Hubička–Nešetřil [28] and Cameron–Vershik [29]. For perspectives from model theory and descriptive set theory, see, e.g., Ealy–Goldbring [30], Melleray [31], Pestov [32], and Usvyatsov [33].

Vershik [2], [3] has demonstrated how Urysohn space, in addition to being the universal ultrahomogeneous Polish space, also can be viewed as the *generic* Polish space, and as a *random* Polish space. Namely, Vershik shows that $\mathbb{U}$ is the generic complete separable metric space, in the sense of Baire category, and he provides symmetric random constructions of $\mathbb{U}$ by describing a wide class of invariant measures concentrated on the class of metric spaces whose completion is $\mathbb{U}$. As with the constructions in [4] and [5], these measures are determined by sampling from certain continuum-sized structures.

Here we construct an approximately $\aleph_0$-categorical theory whose models are those countable metric spaces (encoded in an infinite relational language) that have Urysohn space as their completion. Hence our invariant probability measure concentrated on the class of models of this theory can be thought of as providing yet another symmetric random construction of Urysohn space.

Before describing the theory itself, we provide a relational axiomatization of metric spaces using infinitely many binary relations, where the distance function is implicit in these relations. Let L_{MS} be the language consisting of a binary relation d_q for every $q \in \mathbb{Q}_{\geq 0}$. Given a metric space with distance function $\mathbf{d}$, the intended interpretation will be that $d_q(x, y)$ holds when $\mathbf{d}(x, y) \leq q$. More explicitly, we have, for all $q, r \in \mathbb{Q}_{\geq 0}$,

- $(\forall x)(\forall y)\ \big(d_q(x, y) \rightarrow d_r(x, y)\big)$ when $r \geq q$,
- $(\forall x)(\forall y)\ \big(d_q(x, y) \leftrightarrow d_q(y, x)\big)$,
- $(\forall x)(\forall y)(\forall z)\ \big((d_q(x, y) \wedge d_r(y, z)) \rightarrow d_{q+r}(x, z)\big)$, and

- $(\forall x)\ d_0(x,x)$.

Let T_{MS} denote this theory in the language L_{MS}.

The following result is immediate.

Proposition 5.8. *For every metric space $\mathcal{S} = (S, \mathbf{d}_S)$, the L_{MS}-structure $\mathcal{M}_{\mathcal{S}}$ with underlying set S and sequence of relations $\langle d_q^{\mathcal{M}_S} \rangle_{q \in \mathbb{Q}_{\geq 0}}$ defined by*

$$d_q^{\mathcal{M}_S}(x,y) \quad \text{if and only if} \quad \mathbf{d}_S(x,y) \leq q$$

is a model of T_{MS}.

Conversely, if $\mathcal{N}$ is a model of T_{MS} with underlying set N, and

$$\mathbf{d}_{\mathcal{N}}(x,y) := \inf\{q \in \mathbb{Q}_{\geq 0} : \mathcal{N} \models d_q(x,y)\},$$

then $\mathcal{P}_{\mathcal{N}} := (N, \mathbf{d}_{\mathcal{N}})$ is a metric space.

We will use the maps $\mathcal{S} \mapsto \mathcal{M}_{\mathcal{S}}$ and $\mathcal{N} \mapsto \mathcal{P}_{\mathcal{N}}$ that are implicit in Proposition 5.8 throughout our discussion of Urysohn space.

Note that when a model $\mathcal{N}$ of T_{MS} further satisfies, for each $q \in \mathbb{Q}_{\geq 0}$, the infinitary axioms

- $(\forall x)\ \Big((\bigwedge_{p>q} d_p(x,y)) \to d_q(x,y) \Big)$ and
- $(\forall x)(\forall y)\ \big(d_0(x,y) \to (x = y)\big)$,

then $\mathcal{N} = \mathcal{M}_{\mathcal{S}}$ for some metric space $\mathcal{S}$. However, we will not be able to ensure that these axioms hold in our construction, each stage of which involves a language that has only a finite number of relations of the form d_q.

Proposition 5.9. *For any finite sublanguage L of L_{MS}, every model of the restriction $T_{\mathrm{MS}} \cap \mathcal{L}_{\omega,\omega}(L)$ of T_{MS} can be extended to a model of T_{MS}.*

Proof. Let L be a finite sublanguage of L_{MS}, and let $\mathcal{N}$ be a model of $T_{\mathrm{MS}} \cap \mathcal{L}_{\omega,\omega}(L)$ with underlying set N. Define

$$\mathbb{Q}_L := \{q \in \mathbb{Q}_{\geq 0} : d_q \in L\}.$$

Let $p := \max \mathbb{Q}_L$. For every pair of distinct elements $x, y \in \mathcal{N}$, define

$$\delta^*_{\mathcal{N}}(x,y) := \min\big(2p,\ \inf\{q \in \mathbb{Q}_L : \mathcal{N} \models d_q(x,y)\}\big),$$

and for all $x \in \mathcal{N}$ set

$$\delta^*_{\mathcal{N}}(x,x) := 0.$$

Finally, define

$$\delta_{\mathcal{N}}(x,y) := \inf\{\delta^*_{\mathcal{N}}(x,z_1) + \delta^*_{\mathcal{N}}(z_1,z_2) + \cdots + \delta^*_{\mathcal{N}}(z_n,y) : n \geq 1 \text{ and } z_1, \ldots, z_n \in \mathcal{N}\}.$$

Although $(N, \delta_{\mathcal{N}})$ need not be a metric space, the L_{MS}-structure $\mathcal{M}_{(N,\delta_{\mathcal{N}})}$, given by the map defined in Proposition 5.8, is a model of T_{MS}. By construction, if $\mathcal{N} \models d_q(x,y)$, then $\delta_{\mathcal{N}}(x,y) \leq q$. However, if $\mathcal{N} \models \neg d_q(x,y)$, then by the triangle inequality

$\delta_{\mathcal{N}}(x,y) > q$. Hence $(N, \delta_{\mathcal{N}})$ is consistent with the above "intended interpretation" of the relations in $\mathcal{N}$. In particular, $\mathcal{M}_{(N,\delta_{\mathcal{N}})}$ is an expansion of $\mathcal{N}$ to L_{MS} that is a model of T_{MS}. □

We now describe an important class of examples of countable metric spaces whose completions are (isomorphic to) the full Urysohn space.

Definition 5.10. Let D be a countable dense subset of $\mathbb{R}_+$. Consider the class $\mathscr{S}$ of finite metric spaces $\mathcal{S}$ whose non-zero distances occur in D, and let $\mathscr{F} := \{\mathcal{M}_{\mathcal{S}} : \mathcal{S} \in \mathscr{S}\}$. Note that $\mathscr{F}$ is an amalgamation class. Define $D\mathbb{U}$ to be $\mathcal{P}_{\mathcal{N}}$, where $\mathcal{N}$ is the Fraïssé limit of $\mathscr{F}$.

It is a standard result that any such $D\mathbb{U}$ is a metric space whose completion is $\mathbb{U}$. The particular case $\mathbb{QU}$ has been well-studied, and is known as the *rational Urysohn space.*

We now extend T_{MS} to an L_{MS}-theory T_U whose countable models will be precisely those L_{MS}-structures $\mathcal{N}$ for which the completion of $\mathcal{P}_{\mathcal{N}}$ is isomorphic to $\mathbb{U}$. We will work with finite sublanguages of L_{MS}, rather than all of L_{MS}, because there is no (countable) Fraïssé limit of the class of finite models of T_{MS}; in particular, there are continuum-many non-isomorphic finite models of T_{MS}, even of size 2. On the other hand, in every finite sublanguage L of L_{MS}, there is a Fraïssé limit of the countably many (up to isomorphism) finite models of $T_{\mathrm{MS}} \cap \mathcal{L}_{\omega,\omega}(L)$.

Definition 5.11. Let L be a finite sublanguage of L_{MS}. Note that the class of finite models of $T_{\mathrm{MS}} \cap \mathcal{L}_{\omega,\omega}(L)$ is an amalgamation class. Let T_U^L be the $\mathcal{L}_{\omega,\omega}(L)$-theory of the Fraïssé limit of this class, and define

$$T_U := \bigcup \{T_U^L : \text{finite } L \subseteq L_{\mathrm{MS}}\}.$$

Proposition 5.12. *The theory T_U is consistent.*

Proof. Consider the L_{MS}-structure $\mathcal{M}_{\mathbb{QU}}$. It is a Fraïssé limit of the class of those finite models $\mathcal{N}$ of T_{MS} for which $\mathcal{P}_{\mathcal{N}}$ is a metric space with only rational distances. By Proposition 5.9, and as $\mathbb{Q}$ is dense in $\mathbb{R}$, for any finite sublanguage L of L_{MS}, the Fraïssé limit of the class of finite models of $T_{\mathrm{MS}} \cap \mathcal{L}_{\omega,\omega}(L)$ is isomorphic to $\mathcal{M}_{\mathbb{QU}}|_L$. Hence $\mathcal{M}_{\mathbb{QU}}|_L$ is a model of T_U^L. Therefore $\mathcal{M}_{\mathbb{QU}}$ is a model of T_U, and so T_U is consistent. □

Note that by the above proof, for any countable dense subset $D \subseteq \mathbb{R}_+$, the L_{MS}-structure $D\mathbb{U}$ is a model of T_U. As these are all non-isomorphic, T_U has continuum-many countable models. Also note that for any finite sublanguage L of L_{MS} and dense $D, E \subseteq \mathbb{R}_+$, the L-structures $\mathcal{M}_{D\mathbb{U}}|_L$ and $\mathcal{M}_{E\mathbb{U}}|_L$ are isomorphic (and are both Fraïssé limits as in the above proof).

Theorem 5.13. *Let $\mathcal{S} = (S, \mathbf{d}_S)$ be a countable metric space. Then $\mathcal{M}_{\mathcal{S}}$ is a model of T_U if and only if the completion of $\mathcal{S}$ is isomorphic to $\mathbb{U}$.*

Proof. First suppose that the completion of $\mathcal{S}$ is isomorphic to $\mathbb{U}$. Without loss of generality, we may assume that $\mathcal{S} \subseteq \mathbb{U}$ and that $\mathcal{S}$ is dense in $\mathbb{U}$. We will show that $\mathcal{M}_{\mathcal{S}}$ is a model of T_U.

Let L be any finite sublanguage of L_{MS}, and suppose that

$$(\forall \overline{x})(\exists y)\varphi(\overline{x}, y) \in T_U \cap \mathcal{L}_{\omega,\omega}(L).$$

Because $T_U \cap \mathcal{L}_{\omega,\omega}(L)$ has a pithy Π_2 axiomatization, it suffices to show that $(\forall \overline{x})(\exists y)\varphi(\overline{x}, y)$ holds in $\mathcal{M}_{\mathcal{S}}$.

Fix some $\overline{a} \in \mathcal{M}_{\mathcal{S}}$ where $|\overline{a}|$ is one less than the number of free variables of φ, and let q be the quantifier-free L-type of $\overline{a}$. We will show that there is a witness to $(\exists y)\varphi(\overline{a}, y)$ in $\mathcal{M}_{\mathcal{S}}$.

Because T_U implies the theory of the Fraïssé limit of the class of finite L-structures, there is some quantifier-free L-type $q'(\overline{x}, y)$ extending $q(\overline{x})$ (where $|\overline{x}| = |\overline{a}|$) that is consistent with both $\varphi(\overline{x}, y)$ and $T_{\mathrm{MS}} \cap \mathcal{L}_{\omega,\omega}(L)$.

Now, $\mathbb{U}$ is universal for separable metric spaces, and so there is some tuple $\overline{c}f \in \mathbb{U}$ such that q' holds of $\mathcal{M}_{\mathcal{C}}$ (under the corresponding order of elements), where $\mathcal{C}$ is the substructure of $\mathbb{U}$ with underlying set $\overline{c}f$. As $\mathbb{U}$ is ultrahomogeneous and q is the quantifier-free type of $\overline{a}$, there must be an automorphism σ of $\mathbb{U}$ such that $\sigma(\overline{c}) = \overline{a}$. Define $b := \sigma(f)$. Then q' holds of $\mathcal{M}_{\mathcal{B}}$ (in the corresponding order), where $\mathcal{B}$ is the substructure of $\mathbb{U}$ with underlying set $\overline{a}b$.

But no quantifier-free L-type can ever completely determine the distance between any two distinct points, as L is finite. Hence there is some $\varepsilon > 0$ such that q' also holds of $\mathcal{M}_{\mathcal{A}}$ (in the corresponding order) whenever $\mathcal{A}$ is any finite $(|\overline{a}| + 1)$-element substructure of $\mathbb{U}$ that can be put into one-to-one correspondence with $\overline{a}b$ in such a way that each element of $\mathcal{A}$ is less than ε away from the corresponding element of $\overline{a}b$ and from no other. By assumption, $\mathcal{S}$ is dense in $\mathbb{U}$, and so there is some $b' \in \mathcal{S}$ such that $d_{\mathbb{U}}(b, b') < \varepsilon$. Hence $\mathcal{M}_{\mathcal{S}} \models q'(\overline{a}, b')$, and so $\mathcal{M}_{\mathcal{S}} \models \varphi(\overline{a}, b')$, as desired.

Conversely, suppose that $\mathcal{S}$ is a countable metric space such that $\mathcal{M}_{\mathcal{S}}$ is a model of T_U. We will show that the completion $\mathcal{U}$ of $\mathcal{S}$ is isomorphic to $\mathbb{U}$.

We do this by showing that for every finite metric space $\mathcal{A}$ with underlying set $A \subseteq \mathcal{U}$ and metric space $\mathcal{B}$ extending $\mathcal{A}$ by some element b (not necessarily in $\mathcal{U}$), there is some $b' \in \mathcal{U}$ such that the metric space induced (in $\mathcal{U}$) by $A \cup \{b'\}$ is isomorphic to $\mathcal{B}$. From this it follows that if σ is an isomorphism from $\mathcal{A}$ to another submetric space $\mathcal{A}'$ of $\mathcal{U}$, then for every $c \in \mathcal{U}$, there is some $c' \in \mathcal{U}$ such that the function that extends σ by mapping c to c' is also an isomorphism of induced metric spaces. By a standard back-and-forth argument, this implies the universality and ultrahomogeneity of $\mathcal{U}$. Hence $\mathcal{U}$ is isomorphic to $\mathbb{U}$, as $\mathbb{U}$ is the unique (up to isomorphism) universal ultrahomogeneous complete separable metric space.

Let $\mathcal{A}$ and $\mathcal{B}$ be as above, and suppose $A = \{a_0, \dots, a_{n-1}\}$, where $n = |A|$. Let $\mathcal{U}^*$ be any metric space extending $\mathcal{U}$ by b. and define

$$\gamma_j := \mathbf{d}_{\mathcal{U}^*}(a_j, b)$$

for $0 \leq j < n$. Let $\langle L_i \rangle_{i \in \mathbb{N}}$ be an increasing sequence of finite sublanguages of L_{MS} such that for each $i \in \mathbb{N}$, the language L_i contains enough symbols of the form d_r to imply that whenever two finite models of T_{MS}, both of diameter less than twice that of $\mathcal{B}$, satisfy the same quantifier-free L_i-type (in some order), then each pairwise distance in the first structure is within $2^{-(i+6)}$ of the corresponding distance in the second structure. For each j such that $0 \leq j \leq n-1$, let $\langle a_j^i \rangle_{i \in \mathbb{N}}$ be a Cauchy sequence in $\mathcal{S}$ that converges to a_j with

$$\mathbf{d}_{\mathcal{S}}(a_j^i, a_j^{i+1}) \leq 2^{-(i+3)}$$

for $i \in \mathbb{N}$.

Consider the inductive claim that for $h \in \mathbb{N}$ we have defined $b_0 \cdots b_h \in \mathcal{S}$ that satisfy

$$\mathbf{d}_{\mathcal{S}}(b_i, b_{i+1}) \leq 2^{-i}$$

for $i < h$, and

$$\left|\mathbf{d}_{\mathcal{S}}(a^i_j, b_i) - \gamma_j\right| \leq 2^{-(i+2)},$$

for $0 \leq j \leq n-1$ and $i \leq h$.

If this claim holds for all $h \in \mathbb{N}$, then $\langle b_i \rangle_{i \in \mathbb{N}}$ is a Cauchy sequence in $\mathcal{S}$, which therefore must converge to an element $b' \in \mathcal{U}$. Furthermore, $\mathbf{d}_{\mathcal{U}}(a_j, b') = \gamma_j$ for $0 \leq j \leq n-1$, and so the metric space induced by $A \cup \{b'\}$ is isomorphic to $\mathcal{B}$, as desired.

We now show the inductive claim for $h+1$. Because

$$\left|\mathbf{d}_{\mathcal{U}^*}(a^h_j, b) - \gamma_j\right| \leq 2^{-(h+2)}$$

for $0 \leq j \leq n-1$, and since $\mathcal{M}_{\mathcal{S}}|_{L_{h+1}}$ is the Fraïssé limit of the finite models of $T_{\mathrm{MS}} \cap \mathcal{L}_{\omega,\omega}(L_{h+1})$, we can find a $b_{h+1} \in \mathcal{S}$ satisfying

$$\left|\mathbf{d}_{\mathcal{S}}(a^h_j, b_{h+1}) - \gamma_j\right| \leq 2^{-(h-1)}$$

for $0 \leq j \leq n-1$. We may further assume that $\mathbf{d}_{\mathcal{S}}(b_h, b_{h+1}) \leq 2^{-h}$, as there is a finite metric space containing such a b_{h+1} that extends the one induced by $a^h_0, \ldots, a^h_{n-1}, b_h$. Now, for $0 \leq j \leq n-1$, we have $\mathbf{d}_{\mathcal{S}}(a^h_j, a^{h+1}_j) \leq 2^{-(h+3)}$, and so $\mathbf{d}_{\mathcal{S}}(a^h_j, a_j) \leq 2^{-(h+1)}$; hence

$$\left|\mathbf{d}_{\mathcal{S}}(a^{h+1}_j, b_{h+1}) - \gamma_j\right| \leq 2^{-(h+2)},$$

and so b_{h+1} satisfies the inductive claim. □

Although T_U is not itself $\aleph_0$-categorical, as shown by the examples $D\mathbb{U}$, it is approximately $\aleph_0$-categorical. Let $\alpha\colon \mathbb{N} \to \mathbb{Q}_{\geq 0}$ be a bijection, and for each $i \in \mathbb{N}$ define the finite sublanguage of L_{MS} to be

$$L_i := \{d_{\alpha(j)} : 0 \leq j \leq i\}.$$

Proposition 5.14. *The theory T_U is approximately $\aleph_0$-categorical with witnessing sequence $\langle L_i \rangle_{i \in \mathbb{N}}$.*

Proof. For every $i \in \mathbb{N}$, the restriction $T_U \cap \mathcal{L}_{\omega,\omega}(L_i)$ is the theory of the Fraïssé limit of all finite models of $T_U \cap \mathcal{L}_{\omega,\omega}(L_i)$, hence $\aleph_0$-categorical. □

Proposition 5.15. *The theory T_U and witnessing sequence $\langle L_i \rangle_{i \in \mathbb{N}}$ satisfy the assumptions of Theorem 5.3. Hence there is an S_∞-invariant probability measure m_U on $\mathrm{Str}_{L_{\mathrm{MS}}}$ that is concentrated on the class of models of T_U and that assigns probability 0 to each isomorphism class.*

Proof. For each $i \in \mathbb{N}$, the countable model of $T_U \cap \mathcal{L}_{\omega,\omega}(L_i)$ is isomorphic to $\mathcal{M}_{\mathbb{QU}}|_{L_i}$. Its age has the strong amalgamation property, because the age of $\mathcal{M}_{\mathbb{QU}}$ has the strong amalgamation property.

For each $i \in \mathbb{N}$, let Q_i be the set of quantifier-free L_i-types that are consistent with $T_U \cap \mathcal{L}_{\omega,\omega}(L_i)$. We will show that $\langle Q_i \rangle_{i \in \mathbb{N}}$ has splitting of order 2. Let $j \in \mathbb{N}$ and

$q \in Q_j$. We show that there is some $j' > j$ such that each quantifier-free L_j-type with two free variables has a splitting in the language $L_{j'}$.

Let k be the number of free variables of q. There is an iterated duplicate q' of q having $2k$-many free variables, and there is some finite metric space $\mathcal{S}$ whose positive distances are distinct and such that q' holds of $\mathcal{M}_{\mathcal{S}}$ (under some ordering of the elements of $\mathcal{M}_{\mathcal{S}}$). Let $j' > j$ be such that

$$\{\alpha(i) \,:\, 0 \leq i \leq j'\}$$

partitions $\mathbb{Q}$ so that each part contains at most one positive distance occurring in $\mathcal{S}$. Let $q^{\natural}$ be the quantifier-free $L_{j'}$-type of $\mathcal{M}_{\mathcal{S}}$. Then $q^{\natural}$ is a splitting of q of order 2. □

As with the Kaleidoscope random graphs above, the measure m_U cannot be obtained via the methods in [5]. This is because almost every sample from m_U has nontrivial definable closure, as we now show. Let $\mathcal{N}$ be a structure sampled from m_U, and consider its corresponding metric space $\mathcal{P}_{\mathcal{N}} = (\mathbb{N}, \mathbf{d}_{\mathcal{N}})$. Then with probability 1, for $(i, j), (i', j') \in \mathbb{N}^2$ satisfying $i < j$ and $i' < j'$, we have

$$\mathbf{d}_{\mathcal{N}}(i, j) \neq \mathbf{d}_{\mathcal{N}}(i', j')$$

whenever $(i, j) \neq (i', j')$.

Also m_U does not arise from the standard examples of the form $D\mathbb{U}$, as for any two independent samples $\mathcal{N}_0, \mathcal{N}_1$ from m_U, the sets of real distances

$$\{\mathbf{d}_{\mathcal{N}_w}(i, j) \,:\, i, j \in \mathbb{N} \text{ and } i \neq j\}$$

for $w \in \{0, 1\}$ are almost surely disjoint (and so any two independent samples from m_U are almost surely non-isomorphic — as we already knew). As a consequence, a sample $\mathcal{N}$ is almost surely such that $\mathcal{P}_{\mathcal{N}}$ is not isometric to $D\mathbb{U}$ for any countable dense set $D \subseteq \mathbb{R}_+$.

6. *G*-orbits admitting *G*-invariant probability measures

In this section we characterize, for certain Polish groups G, those transitive Borel G-spaces that admit G-invariant measures. In particular, we do so for all countable Polish groups and for countable products of symmetric groups on a countable (finite or infinite) set. Throughout this section, let $(G, \cdot)$ be a Polish group.

6.1. S_∞-actions

For a countable first-order language L, recall that Str_L is the space of L-structures with underlying set $\mathbb{N}$, with $\circledast_L \colon S_\infty \times \mathrm{Str}_L \to \mathrm{Str}_L$ the logic action of S_∞ on Str_L by permutation of the underlying set.

Also recall that for any formula $\varphi \in \mathcal{L}_{\omega_1,\omega}(L)$ and any $\ell_1, \ldots, \ell_n \in \mathbb{N}$, we have defined the collection of models

$$[\![\varphi(\ell_1, \ldots, \ell_n)]\!] := \big\{\mathcal{M} \in \mathrm{Str}_L \,:\, \mathcal{M} \models \varphi(\ell_1, \ldots, \ell_n)\big\}.$$

The following is an equivalent formulation of the main result of [5].

Theorem 6.1 ([5]). *Let $(X, \circ)$ be a transitive Borel S_∞-space, and suppose that $\iota\colon X \to \mathrm{Str}_L$ is a Borel embedding, where L is some countable language. Note that the image of ι is the S_∞-space*

$$(\{\mathcal{M} \in \mathrm{Str}_L : \mathcal{M} \cong \mathcal{M}^*\}, \circledast_L)$$

consisting of the orbit in Str_L of some countably infinite L-structure $\mathcal{M}^*$ under the action of $\circledast_L$. Then X admits an S_∞-invariant probability measure if and only if $\mathcal{M}^*$ has trivial definable closure.

The following well-known result will be useful in our classification of transitive Borel S_∞-spaces admitting S_∞-invariant probability measures.

Theorem 6.2 ([19, Theorem 2.7.3]). *Let L be a countable language having relation symbols of arbitrarily high arity. Then $(\mathrm{Str}_L, \circledast_L)$ is a universal Borel S_∞-space.*

Note that by Theorem 6.2, for any transitive Borel S_∞-space $(X, \circ)$, we can always find an embedding $X \to \mathrm{Str}_L$, where L is as in Theorem 6.2. Hence Theorem 6.1 provides a complete characterization of those transitive Borel S_∞-spaces admitting S_∞-invariant probability measures. The main result of this section, Theorem 6.11, is a generalization of Theorem 6.1 to the case of invariance under certain products of symmetric groups.

6.2. Countable G-spaces

We now characterize, for countable groups G, those transitive Borel G-spaces admitting G-invariant probability measures.

Lemma 6.3. *Let $(X, \circ)$ be a finite Borel G-space. Then $(X, \circ)$ admits a G-invariant probability measure.*

Proof. The counting measure ρ_X, given by $\rho_X(A) = |A|/|X|$, is G-invariant. □

Corollary 6.4. *Suppose G is finite. Then every transitive Borel G-space admits an invariant probability measure.*

Proof. Because G is finite, every transitive Borel G-space is also finite. By Lemma 6.3, every such G-space admits a G-invariant probability measure. □

Lemma 6.5. *Let $(X, \circ)$ be a countably infinite transitive Borel G-space. Then $(X, \circ)$ does not admit a G-invariant probability measure.*

Proof. Suppose μ_X is a G-invariant probability measure on $(X, \circ)$. By the transitivity of X, for all $x, y \in X$ we must have $\mu_X(\{x\}) = \mu_X(\{y\})$. Let $\alpha := \mu_X(\{x\})$. As X is countable and μ_X is countably additive, we have

$$1 = \mu_X(X) = \sum_{x \in X} \mu_X(\{x\}) = \sum_{x \in X} \alpha.$$

But this is impossible as X is infinite, and so for any non-zero α the right-hand side is infinite. □

Corollary 6.6. *Suppose G is countable. Then a transitive Borel G-space X admits a G-invariant probability measure if and only if X is finite.*

Proof. As G is countable and X is transitive, X must be countable. The conclusion then follows from Lemmas 6.3 and 6.5. □

6.3. Products of symmetric groups

We now consider those groups G that are a countable product of symmetric groups on countable sets. For such G, we will characterize those transitive Borel G-spaces that admit a G-invariant probability measure, using the following standard result from descriptive set theory.

Recall the definition of $(\mathrm{Str}^{\mathcal{M}_0}_{L_0,L}, \circledast^{\mathcal{M}_0}_L)$ from §2.5.3.

Theorem 6.7 ([19, Theorem 2.7.4]). *Let L be a countable language and let L_0 be a sublanguage of L such that $L \setminus L_0$ contains relations of arbitrarily high arity. Let $\mathcal{M}_0 \in \mathrm{Str}_{L_0}$. Then $\mathrm{Aut}(\mathcal{M}_0)$ is a closed subgroup of S_∞, and $(\mathrm{Str}^{\mathcal{M}_0}_{L_0,L}, \circledast^{\mathcal{M}_0}_L)$ is a universal $\mathrm{Aut}(\mathcal{M}_0)$-space.*

Note that the $\mathrm{Aut}(\mathcal{M}_0)$-orbit of any structure $\mathcal{M}^* \in \mathrm{Str}^{\mathcal{M}_0}_{L_0,L}$ is of the form

$$\mathrm{Orb}_{L_0}(\mathcal{M}^*) := \big\{\mathcal{M} \in \mathrm{Str}^{\mathcal{M}_0}_{L_0,L} \,:\, \mathcal{M} \cong \mathcal{M}^*\big\}.$$

We will be interested in the case when L_0 is a unary language, i.e., consists entirely of unary relations.

For completeness, and to fix notation for later, we now recall basic facts about the relationship between universal G-spaces and structures in a given language, when G is the product of symmetric groups. For the remainder of the section, let $\ell_0, \ell_1, \dots, \ell_\infty$ be finite or countably infinite, define

$$G_\infty := S_\infty^{\ell_\infty} \quad \text{and}$$
$$G_{\mathrm{fin}} := \prod_{n\in\mathbb{N}} S_n^{\ell_n},$$

and let $G := G_\infty \times G_{\mathrm{fin}}$.

Define the countable language

$$L_G := \{U_i^\infty \,:\, 1 \le i \le \ell_\infty\} \,\cup\, \bigcup_{n\in\mathbb{N}} \{U_i^n \,:\, 1 \le i \le \ell_n\} \,\cup\, \{V_\infty, V_{\mathrm{fin}}\},$$

consisting of unary relation symbols. Consider the theory $T_G \subseteq \mathcal{L}_{\omega_1,\omega}(L_G)$ defined by the axioms

- $(\forall x)\neg\big(U_i^\infty(x) \wedge U_j^\infty(x)\big)$ whenever $1 \le i < j \le \ell_\infty$,
- $(\forall x)\neg\big(U_i^n(x) \wedge U_j^m(x)\big)$ for all $n, m \in \mathbb{N}$ and i, j such that $1 \le i \le \ell_n$ and $1 \le j \le \ell_m$ for which $(i, n) \ne (j, m)$,
- $(\forall x)\big(V_{\mathrm{fin}}(x) \leftrightarrow \bigvee_{n\in\mathbb{N}} \bigvee_{1\le i\le \ell_n} U_i^n(x)\big)$,
- $(\forall x)\big(V_\infty(x) \leftrightarrow \bigvee_{1\le i\le \ell_\infty} U_i^\infty(x)\big)$,
- $(\forall x)\big(V_{\mathrm{fin}}(x) \leftrightarrow \neg V_\infty(x)\big)$,
- for all i such that $1 \le i \le \ell_\infty$, the set $\{x \,:\, U_i^\infty(x)\}$ is infinite, and
- for all $n \in \mathbb{N}$ and i such that $1 \le i \le \ell_n$, we have $|\{x \,:\, U_i^n(x)\}| = n$.

These axioms are consistent; in particular, they can be realized by any L_G-structure partitioned by the U-relations for which each U^∞ relation is infinite, each U^n relation has size n, the relation V_∞ is the union of all U^∞-relations, and V_{fin} is the union of all U^n relations.

Fix some $\mathcal{A}_G \in \mathrm{Str}_{L_G}$ that is a model of T_G. For each U-relation, write $\widetilde{U} := U^{\mathcal{A}_G} = \{x \in A : \mathcal{A}_G \models U(x)\}$, and similarly for each V-relation. Let $P(\widetilde{U})$ be the collection of permutations of $\widetilde{U}$.

Lemma 6.8. *The group G is isomorphic to the automorphism group of $\mathcal{A}_G$.*

Proof. A permutation of $\mathbb{N}$ induces an automorphism of $\mathcal{A}_G$ if and only if it preserves each U-relation. Hence $\mathrm{Aut}(\mathcal{A}_G)$ is isomorphic to

$$\prod_{1\le i\le \ell_\infty} P(\widetilde{U_i^\infty}) \times \prod_{n\in\mathbb{N}} \prod_{1\le i\le \ell_n} P(\widetilde{U_i^n}).$$

However, as each $P(\widetilde{U_i^\infty})$ is isomorphic to S_∞, and each $P(\widetilde{U_i^n})$ is isomorphic to S_n, we have that $\mathrm{Aut}(\mathcal{A}_G) \cong G$. ☐

Lemma 6.9. *Let L be a countable unary language and $\mathcal{M}$ be a countably infinite L-structure. Then* $\mathrm{Aut}(\mathcal{M})$ *is isomorphic to a product of symmetric groups.*

Proof. For $x, y \in \mathcal{M}$, define $x \sim y$ to hold when x and y have the same quantifier-free L-type. Let E be the collection of $\sim$-equivalence classes. As L is unary, the automorphisms of $\mathcal{M}$ are precisely those permutations of the underlying set of $\mathcal{M}$ that preserve $\sim$. Hence $\mathrm{Aut}(\mathcal{M}) \cong \prod_{Y\in E} S_{|Y|}$. ☐

Note that Lemmas 6.8 and 6.9 imply the standard fact that the countable products of symmetric groups on countable (finite or infinite) sets are precisely those groups isomorphic to automorphisms of structures in countable unary languages.

6.4. Non-existence of invariant probability measures

Recall that $G = \mathrm{Aut}(\mathcal{A}_G)$ by Lemma 6.8. For the rest of the section, fix a countable relational language L that extends L_G.

We now classify those orbits in $\mathrm{Str}^{\mathcal{A}_G}_{L_G,L}$ that admit an $\mathrm{Aut}(\mathcal{A}_G)$-invariant probability measure. Then in particular, if $L \setminus L_G$ has relations of arbitrarily high arity, then $\mathrm{Str}^{\mathcal{A}_G}_{L_G,L}$ will be a universal G-space, and so we will obtain a classification of those transitive G-spaces that admit G-invariant probability measures.

Notice that in any structure $\mathcal{M} \in \mathrm{Str}^{\mathcal{A}_G}_{L_G,L}$, the algebraic closure of the empty set contains $V^{\mathcal{A}_G}_{\text{fin}}$, which is non-empty precisely when G is not a countable power of S_∞. Hence, when $V^{\mathcal{A}_G}_{\text{fin}}$ is non-empty, $\mathcal{M}$ does not have trivial definable closure. To deal with this issue, we define the following notion.

Definition 6.10. An L-structure $\mathcal{M} \in \mathrm{Str}^{\mathcal{A}_G}_{L_G,L}$ has **almost-trivial definable closure** if and only if for every tuple $\overline{a} \in \mathcal{M}$, we have

$$\mathrm{dcl}(\overline{a} \cup V^{\mathcal{A}_G}_{\text{fin}}) = \overline{a} \cup V^{\mathcal{A}_G}_{\text{fin}}.$$

Note that the analogous notion of almost-trivial algebraic closure coincides with almost-trivial definable closure, similarly to the way that trivial definable closure and trivial algebraic closure coincide. Using this notion, we can now state our main classification.

Theorem 6.11. *Let $\mathcal{M} \in \mathrm{Str}^{\mathcal{A}_G}_{L_G,L}$. Then $\mathrm{Orb}_{L_G}(\mathcal{M})$ admits a G-invariant probability measure if and only if $\mathcal{M}$ has almost-trivial definable closure.*

We will prove Theorem 6.11 in two steps. We prove the forward direction in Proposition 6.12. This argument is very similar to an analogous result in [5], but we include it here for completeness. In Proposition 6.14, we prove the reverse direction.

Proposition 6.12. *Let $\mathcal{M} \in \mathrm{Str}^{\mathcal{A}_G}_{L_G,L}$, and suppose that $\mathrm{Orb}_{L_G}(\mathcal{M})$ admits a G-invariant probability measure. Then $\mathcal{M}$ has almost-trivial definable closure.*

Proof. Let μ be a G-invariant probability measure on $\mathrm{Orb}_{L_G}(\mathcal{M})$, and suppose that there is a finite tuple $\overline{a} \in \mathcal{M}$ such that

$$b \in \mathrm{dcl}(\overline{a} \cup V^{\mathcal{A}_G}_{\mathrm{fin}}) \setminus (\overline{a} \cup V^{\mathcal{A}_G}_{\mathrm{fin}}).$$

Let $p(\overline{x}y)$ be a formula that generates a (principal) $\mathcal{L}_{\omega_1,\omega}(L)$-type of $\overline{a}b$, i.e., a formula of $\mathcal{L}_{\omega_1,\omega}(L)$ with free variables $\overline{x}y$ such that for any $\mathcal{L}_{\omega_1,\omega}(L)$-formula ψ whose free variables are among $\overline{x}y$, either

$$\models (\forall \overline{x})(\forall y)\big(p(\overline{x}y) \rightarrow \psi(\overline{x}y)\big) \qquad \text{or} \qquad \models (\forall \overline{x})(\forall y)\big(p(\overline{x}y) \rightarrow \neg\psi(\overline{x}y)\big).$$

Because $\mathcal{M} \models (\exists \overline{x}y)\, p(\overline{x}y)$, the measure μ is concentrated on $[\![(\exists \overline{x}y)\, p(\overline{x}y)]\!]_{\mathcal{A}_G}$. By the countable additivity of μ, there is some $\overline{m} \in \mathbb{N}$ such that $\mu\big([\![(\exists y)\, p(\overline{m}y)]\!]_{\mathcal{A}_G}\big) > 0$.

Now, $b \notin V^{\mathcal{A}_G}_{\mathrm{fin}}$, and so $b \in V^{\mathcal{A}_G}_{\infty}$. Hence we must have $\mathcal{M} \models U^{\infty}_k(b)$ for some k such that $1 \leq k \leq \ell_\infty$. Let

$$F := \{n^* \in \mathbb{N} : \mathcal{A}_G \models U^{\infty}_k(n^*) \text{ and } n^* \notin \overline{m}\}.$$

As $b \notin \overline{a}$, note that $[\![(\exists y)\, p(\overline{m}y)]\!]_{\mathcal{A}_G} = \bigcup_{n \in F} [\![p(\overline{m}n)]\!]_{\mathcal{A}_G}$. Because $b \in \mathrm{dcl}(\overline{a} \cup V^{\mathcal{A}_G}_{\mathrm{fin}}) \setminus (\overline{a} \cup V^{\mathcal{A}_G}_{\mathrm{fin}})$, for any distinct $n_0, n_1 \in F$ we have $[\![p(\overline{m}n_0)]\!]_{\mathcal{A}_G} \cap [\![p(\overline{m}n_1)]\!]_{\mathcal{A}_G} = \emptyset$, and so $[\![(\exists y)\, p(\overline{m}y)]\!]_{\mathcal{A}_G} = \sum_{n^* \in F} [\![p(\overline{m}n^*)]\!]_{\mathcal{A}_G}$.

By countable additivity, there is some $n \in F$ such that $\alpha := \mu\big([\![p(\overline{m}n)]\!]_{\mathcal{A}_G}\big) > 0$. Further, by the definition of F, for every $n^* \in F$ there is some $g \in G$ such that $g(\overline{m}n) = \overline{m}n^*$ and g fixes $V^{\mathcal{A}_G}_{\mathrm{fin}}$. As μ is G-invariant, for all $n^* \in F$ we have $\mu\big([\![p(\overline{m}n^*)]\!]_{\mathcal{A}_G}\big) = \mu\big([\![p(\overline{m}n)]\!]_{\mathcal{A}_G}\big)$, and so $[\![(\exists y)\, p(\overline{m}y)]\!]_{\mathcal{A}_G} = \sum_{n^* \in F} \alpha$. This is a contradiction, as $\alpha > 0$ and F is infinite. ☐

This concludes the forward direction of Theorem 6.11.

6.5. Constructing the invariant probability measure

The reverse direction of Theorem 6.11 will use the construction in Section 4 analogously to the way in which the main construction in [5] is used to classify those transitive S_∞-spaces admitting S_∞-invariant probability measures.

Lemma 6.13. *Let $\mathcal{M} \in \mathrm{Str}^{\mathcal{A}_G}_{L_G,L}$, and suppose that μ is a G_∞-invariant probability measure on $\mathrm{Orb}_{L_G}(\mathcal{M})$. Then there is a G-invariant probability measure μ_{fin} on $\mathrm{Orb}_{L_G}(\mathcal{M})$.*

Proof. First note that, for each $n \in \mathbb{N}$ and $1 \leq i \leq \ell_n$, there is a unique order-preserving bijection

$$\iota^n_i \colon \widetilde{U^n_i} \rightarrow \{1, \ldots, n\}.$$

Recall that these relations $\widetilde{U_i^n}$, along with $\widetilde{U_i^\infty}$, partition $\mathcal{A}_G$. Define the maps

$$\alpha\colon \mathbb{N} \to \mathbb{N} \text{ and}$$
$$\beta\colon \mathbb{N} \to \mathbb{N} \cup \{\infty\}$$

to be such that for all $n \in \mathbb{N}$,

$$\mathcal{A}_G \models U^{\beta(n)}_{\alpha(n)}(n).$$

For every finite subset $Y \subseteq \mathbb{N}$, let

$$Y^* := \bigcup_{y\in Y} \widetilde{U^{\beta(y)}_{\alpha(y)}}.$$

Further, define the finite group

$$G_Y := \prod_{a,b\in\mathbb{N}} \big\{S_b : (\exists y \in Y)\,(\alpha(y) = a \text{ and } \beta(y) = b)\big\}.$$

In other words, G_Y contains the product of $\big|\{\alpha(y) : y \in Y \text{ and } \beta(y) = b\}\big|$-many copies of S_b.

There is a natural action of G_Y on Y^* that fixes $\widetilde{V_\infty}$ pointwise, and uses the $\alpha(y)$-th copy of $S_{\beta(y)}$ to permute $\widetilde{U^{\beta(y)}_{\alpha(y)}}$.

We will define μ_{fin} via a sampling procedure. Begin by sampling an element $\mathcal{N}^* \in \mathrm{Orb}_{L_G}(\mathcal{M})$ according to μ. Next, for each unary relation U_i^n where $n \in \mathbb{N}$ and $1 \le i \le \ell_n$, independently select an element σ_i^n of S_n, uniformly at random. Finally, let μ_{fin} be the distribution of the structure $\mathcal{N} \in \mathrm{Str}^{\mathcal{A}_G}_{L_G,L}$ defined as follows. For every relation symbol $R \in L$ and every $h_1, \ldots, h_j \in \mathbb{N}$, where j is the arity of R, let

$$\mathcal{N} \models R(h_1, \ldots, h_j) \quad \text{iff} \quad \mathcal{N}^* \models R(h_1^*, \ldots, h_j^*),$$

where for $1 \le p \le j$, when $h_p^* \in \widetilde{U_i^n}$ for some $n \in \mathbb{N}$ and i such that $1 \le i \le \ell_n$, we have

$$\big((\iota_i^n)^{-1}\sigma_i^n\iota_i^n\big)\big(h_p^*\big) = h_p,$$

and when $h_p^* \in \widetilde{V_\infty}$, we have $h_p^* = h_p$. Now, $\mathcal{N}$ is almost surely isomorphic to $\mathcal{N}^*$ via the isomorphism that is the identity on $\widetilde{V_\infty}$ and is $(\iota_i^n)^{-1}\sigma_i^n\iota_i^n$ on each $\widetilde{U_i^n}$. Thus μ_{fin} is is a measure on $\mathrm{Orb}_{L_G}(\mathcal{M})$, as claimed.

We now show that the probability measure μ_{fin} is G_{fin}-invariant. Because, in the definition of μ_{fin}, each finite permutation σ_i^n was selected uniformly independently from S_n, we have

$$\mu_{\text{fin}}\big([\![R(h_1, \ldots, h_j)]\!]_{\mathcal{A}_G}\big) = \frac{1}{\big|G_{\{h_1,\ldots,h_j\}}\big|} \sum_{g\in G_{\{h_1,\ldots,h_j\}}} \mu\Big([\![R\big(g(h_1), \ldots, g(h_j)\big)]\!]_{\mathcal{A}_G}\Big),$$

where each $g \in G_{\{h_1,\ldots,h_j\}}$ acts on each h_p (for $1 \le p \le j$) as described above.

Note, however, that for all $g^* \in G_{\text{fin}}$, there is some $g \in G_{\{h_1,\ldots,h_j\}}$ such that the actions of g and g^* agree on $\{h_1, \dots, h_j\}$. Hence

$$\mu_{\mathrm{fin}}\Big(\big[\!\big[R(g^*(h_1), \dots, g^*(h_j))\big]\!\big]_{\mathcal{A}_G}\Big) = \mu_{\mathrm{fin}}\big([\![R(h_1, \dots, h_j)]\!]_{\mathcal{A}_G}\big),$$

and so μ_{fin} is G_{fin}-invariant.

Recall that μ is G_∞-invariant. We now show that μ_{fin} is also G_∞-invariant, so that μ_{fin} is invariant under $G = G_\infty \times G_{\mathrm{fin}}$, as desired. Let $f \in G_\infty$, let $R \in L$ be a relation symbol, and let j be the arity of R. We now show that, for all $h_1, \dots, h_j \in \mathbb{N}$,

$$\begin{aligned}
&\mu_{\mathrm{fin}}\Big(\big[\!\big[R\big(f(h_1), \dots, f(h_j)\big)\big]\!\big]_{\mathcal{A}_G}\Big) \\
&= \frac{1}{\big|G_{\{f(h_1),\dots,f(h_j)\}}\big|} \sum_{g \in G_{\{f(h_1),\dots,f(h_j)\}}} \mu\Big(\big[\!\big[R\big(g(f(h_1)), \dots, g(f(h_j))\big)\big]\!\big]_{\mathcal{A}_G}\Big) \\
&= \frac{1}{\big|G_{\{h_1,\dots,h_j\}}\big|} \sum_{g \in G_{\{h_1,\dots,h_j\}}} \mu\Big(\big[\!\big[R\big(g(h_1), \dots, g(h_j)\big)\big]\!\big]_{\mathcal{A}_G}\Big) \\
&= \mu_{\mathrm{fin}}\big([\![R(h_1, \dots, h_j)]\!]_{\mathcal{A}_G}\big),
\end{aligned}$$

where each $g \in G_{\{h_1,\dots,h_j\}}$ again acts on each $g(h_p)$ and h_p (for $1 \le p \le j$) as described above. The first and third equalities are as before. Note that f is the identity on $\widetilde{V_{\mathrm{fin}}}$ and so $G_{\{f(h_1),\dots,f(h_j)\}} = G_{\{h_1,\dots,h_j\}}$; the second equality follows from this and our assumption that μ is G_∞-invariant. Therefore μ_{fin} is G_∞-invariant, hence G-invariant. ☐

Proposition 6.14. *Let $\mathcal{M} \in \mathrm{Str}^{\mathcal{A}_G}_{L_G,L}$, and suppose that $\mathcal{M}$ has almost-trivial definable closure. Then $\mathrm{Orb}_{L_G}(\mathcal{M})$ has a G-invariant probability measure.*

Proof. There are two cases. Suppose $\widetilde{V_\infty}$ is empty. In this case, G_∞ is the trivial group, and so every measure on $\mathrm{Orb}_{L_G}(\mathcal{M})$ is G_∞-invariant.

Otherwise, $\widetilde{V_\infty}$ is non-empty. Hence $U_1^\infty \in L_G$, and so $\widetilde{U_1^\infty}$ is a countably infinite set. Therefore $\widetilde{V_\infty}$ is countably infinite, and so there is a bijection $\tau\colon \widetilde{V_\infty} \to \mathbb{N}$. Let $\mathcal{M}^\tau \in \mathrm{Str}_{\mathscr{C}_0,L}$ be such that for any quantifier-free L-type q,

$$\mathcal{M}^\tau \models q(h_1, \dots, h_j) \quad \text{iff} \quad \mathcal{M} \models q\big(\tau^{-1}(h_1), \dots, \tau^{-1}(h_j)\big),$$

where j is the number of free variables of q.

Fix some countable admissible set A containing the Scott sentence σ of $\mathcal{M}^\tau$ (equivalently, of $\mathcal{M}$). Let the L_A-theory Σ_A be the definitional expansion (as in Lemma 2.1) of A. Let $T_A := \Sigma_A \cup \{\sigma_A\}$, where $\sigma_A \in L_A$ is a pithy Π_2 sentence such that

$$\Sigma_A \models \sigma_A \leftrightarrow \sigma.$$

For each $i \in \mathbb{N}$ define the language $L_i := L_A$ and theory $T_i := T_A$, and let Q_i be any enumeration of all quantifier-free L_A-types over A (of which there are only countably many).

Let $\mathcal{M}^\tau_A$ be the unique expansion of $\mathcal{M}^\tau$ to a model of Σ_A. We will now show that there is an $S_\infty^{C_0}$-invariant probability measure on $\mathrm{Str}_{\mathscr{C}_0,L_A}$ that is concentrated on the class of models of T_A. We will do so by showing that $\langle Q_i \rangle_{i\in\mathbb{N}}$ satisfies conditions (W), (D), (E) and (C) of our main construction, and so Proposition 4.3 applies. Now, (W), (E), and (C) follow immediately as each Q_i enumerates all quantifier-free types consistent with $T_i = T_A$.

Suppose we do not have condition (D), i.e., duplication of quantifier-free types. Then there is some $i \in \mathbb{N}$, some non-redundant non-constant quantifier-free type $q \in Q_i$, and some tuple $\overline{a} \in \mathcal{M}_A^\tau$ such that there is a unique $b \in V_\infty^{\mathcal{M}_A^\tau}$ (as q is non-constant) for which

$$\mathcal{M}_A^\tau \models q(\overline{a}, b).$$

In particular, if $g \in \mathrm{Aut}(\mathcal{M}_A^\tau)$ fixes $\overline{a} \cup V_\infty^{\mathcal{M}_A^\tau}$ pointwise, then $g(b) = b$, and so $\mathcal{M}_A^\tau$ does not have almost-trivial definable closure (since b is disjoint from $\overline{a}$ as q is non-redundant). This violates our assumption of almost-trivial definable closure for $\mathcal{M}$, as $\mathcal{M}$ is isomorphic to $\mathcal{M}_A^\tau$. Hence condition (D) holds, and so by Proposition 4.3 there is an invariant measure m_∞° on $\mathrm{Str}_{\mathscr{C}_0, L_A}$ that is concentrated on the class of models of T_A, i.e., the isomorphism class of $\mathcal{M}_A^\tau$.

Now let μ be the probability measure on $\mathrm{Str}_{L_G,L}^{\mathcal{A}_G}$ satisfying, for any relation symbol $R \in L$,

$$\mu\big(\llbracket R(h_1, \ldots, h_j)\rrbracket_{\mathcal{A}_G}\big) = \mu_\infty^\circ\big(\llbracket R(\tau(h_1), \ldots, \tau(h_j)\rrbracket_{\mathscr{C}_0}\big),$$

where j is the arity of R. The measure μ is concentrated on $\mathrm{Orb}_{L_G}(\mathcal{M})$, as m_∞° is concentrated on the isomorphism class of $\mathcal{M}_A^\tau$. Hence the restriction μ' of μ to $\mathrm{Orb}_{L_G}(\mathcal{M})$ is a probability measure. Furthermore, μ is G_∞-invariant because m_∞° is $S_\infty^{C_0}$-invariant.

By Lemma 6.13 applied to $\mathcal{M}$ and μ', there is a G-invariant probability measure on $\mathrm{Orb}_{L_G}(\mathcal{M})$. $\square$

This concludes the reverse direction of Theorem 6.11.

7. Concluding remarks

In this paper we have provided conditions under which the class of models of a theory admits an invariant measure that is not concentrated on any single isomorphism class. But much remains to be explored. In particular, there are natural constructions of invariant measures that do not arise by the techniques that we have described, but which would be interesting to capture through general constructions.

7.1. Other invariant measures

The best-known invariant measures concentrated on the Rado graph are the distributions of the countably infinite Erdős-Rényi random graphs $\mathbb{G}(\mathbb{N}, p)$ for $0 < p < 1$, in which edges are chosen independently using weight p coins. These are not produced by our constructions. In particular, when considered as arising from dense graph limits, these limits all have positive entropy (as defined in, e.g., [34, §D.2]), while any of our invariant measures concentrated on graphs corresponds to a dense graph limit that has zero entropy; equivalently, our measures arise from graphons that are $\{0, 1\}$-valued a.e., or "random-free" (see [34, §10]).

7.1.1. Kaleidoscope theories

A similar phenomenon occurs with the following natural construction of an invariant measure concentrated on the class of models of the Kaleidoscope theory built from certain ages. Consider an age A in a language L, both satisfying the hypotheses of

Proposition 5.7, and let $n \in \mathbb{N}$ be such that A has at least two non-equal elements of size n on the same underlying set.

Since A is a strong amalgamation class, there is some invariant measure μ concentrated on the (isomorphism class of the) Fraïssé limit of A, as proved in [5]. We now describe an invariant measure, constructed using μ, that is concentrated on the class of models of the Kaleidoscope theory T_∞ built from the age A.

Namely, consider the distribution μ_∞ of the following random construction. Let $\mathcal{X}$ be a random structure in Str_{L_∞} such that for each $i \in \mathbb{N}$, $\mathcal{X}|_{L^i}$ is an L^i-structure consisting of an independent sample from μ. Observe that this procedure almost surely produces a model of T_∞, and so μ_∞ is an invariant measure concentrated on the class of models of T_∞.

For any n-tuple $\overline{a} \in \mathbb{N}$ and any distinct $i, j \in \mathbb{N}$, the random quantifier-free L^i-type of $\overline{a}$ induced by sampling from μ_∞ is independent from the random quantifier-free L^j-type of $\overline{a}$. Hence the set of structures realizing any given quantifier-free L_∞-type in n variables has measure 0, and so μ_∞ assigns measure 0 to any single isomorphism class. Furthermore, for ages consisting of graphs, when μ is not random-free, one can show that the resulting invariant measure is not captured via our constructions above.

For example, consider the case of the Kaleidoscope random graphs, where μ is the distribution of the Erdős-Rényi graph $\mathbb{G}(\mathbb{N}, 1/2)$, in which edges are determined by independent flips of a fair coin. Then μ_∞ is an invariant measure determined by independently flipping a fair coin to determine the presence of a c-colored edge for each pair of vertices, for each of countably many colors c. The measure μ_∞ is concentrated on the class of Kaleidoscope random graphs and assigns measure 0 to each isomorphism class, but does not arise via our methods.

7.1.2. Urysohn space

Likewise, there is another natural invariant measure on $\mathrm{Str}_{L_{\mathrm{MS}}}$ concentrated on the class of countable L_{MS}-structures $\mathcal{N}$ that are models of T_U (i.e., such that the completion of $\mathcal{P}_\mathcal{N}$ is $\mathbb{U}$), but which assigns measure 0 to each isomorphism class.

Namely, for any countable dense set $D \subseteq \mathbb{R}_+$, recall that $D\mathbb{U}$ is the metric space induced by the Fraïssé limit of all finite metric spaces (considered as L_{MS}-structures) whose set of non-zero distances is contained in D. Note that for any such D, the L_{MS}-structure $\mathcal{M}_{D\mathbb{U}}$ has trivial definable closure (unlike the L_{MS}-structure corresponding to a typical sample of the invariant measure m_U that we constructed in Proposition 5.15). Hence, as proved in [5], there is an invariant measure m_D on $\mathrm{Str}_{L_{\mathrm{MS}}}$, concentrated on the isomorphism class of $\mathcal{M}_{D\mathbb{U}}$.

Now let $\widetilde{D}$ be a random subset of $\mathbb{R}_+$ chosen via a countably infinite set of independent samples from any non-degenerate atomless probability measure on $\mathbb{R}_+$. Then with probability 1, the set $\widetilde{D}$ is infinite, dense, and for any given $r \in \mathbb{R}_+$ does not contain r. Finally, consider the random measure $m_{\widetilde{D}}$. Its distribution is also an invariant measure on $\mathrm{Str}_{L_{\mathrm{MS}}}$ concentrated on the class of countable L_{MS}-structures $\mathcal{N}$ such that the completion of the corresponding metric space $\mathcal{P}_\mathcal{N}$ is isometric to $\mathbb{U}$, but which assigns measure 0 to each isomorphism class. However, this invariant measure is different from the measure m_U that we constructed in Proposition 5.15, as a typical sample from it has trivial definable closure, whereas a typical sample from m_U does not.

We now discuss a more elaborate case of invariant measures that can also be described explicitly but which do not arise from our construction. This set of examples, along with the explicit Kaleidoscope and Urysohn constructions described above, motivate the search for further general conditions that lead to invariant measures.

7.1.3. Continuous transformations

The previous example involved no relationship between the various copies L^j of the original language. We now consider a more complex example, in which interactions within a sequence of languages allow us to describe "transformations" from one structure to another. Although the invariant measure in this example will assign measure 0 to every isomorphism class, it is not clear how it could arise from the methods of this paper.

Let L be a countable relational language. Consider the larger language L_{tr}, which consists of the disjoint union of countably infinitely many copies L^t of L indexed by $t \in \mathbb{Q} \cap [0, 1]$. For each relation symbol $R \in L$, write R^t for the corresponding symbol indexed by $t \in \mathbb{Q} \cap [0, 1]$. One can think of the L_{tr}-structure as describing a "time-evolution" starting with a structure which occurs in the first sublanguage L^0, and ending at another structure which occurs in the last sublanguage L^1, progressing through structures in intermediate sublanguages.

Definition 7.1. Let $\mathcal{M}_0$ be an L^0-structure and $\mathcal{M}_1$ an L^1-structure. We call an L_{tr}-structure $\mathcal{M}$ a **transformation** of $\mathcal{M}_0$ into $\mathcal{M}_1$ when

$$\mathcal{M}|_{L^0} = \mathcal{M}_0 \quad \text{and} \quad \mathcal{M}|_{L^1} = \mathcal{M}_1,$$

and for all relation symbols $R \in L$, where n is the arity of R, and all $s, t \in \mathbb{Q}$ such that $0 \le s < t \le 1$,

$$\mathcal{M} \models (\forall x_1, \dots, x_n)\big(R^s(x_1, \dots, x_n) \to R^t(x_1, \dots, x_n)\big).$$

We now define a notion, called a *nesting*, that will ensure coherence between structures in languages with intermediate indices, as "time" progresses.

Definition 7.2. Suppose A_0 is an age in the language L^0 and A_1 is an age in the language L^1. We define a **nesting of** A_0 **in** A_1 to be an age A in the language $L^0 \cup L^1$ that satisfies the following properties:

- A is a strong amalgamation class.
- For every $\mathcal{K} \in A$ and every relation R in L,

  $$\mathcal{K} \models (\forall x_1, \dots, x_n)\big(R^0(x_1, \dots x_n) \to R^1(x_1, \dots x_n)\big),$$

  where n is the arity of R.
- If $\mathcal{N}$ is a Fraïssé limit of A, then $\mathcal{N}|_{L^0}$ is a Fraïssé limit of A_0 and $\mathcal{N}|_{L^1}$ is a Fraïssé limit of A_1.

For example, consider the age consisting of all those ways that a finite graph can be overlaid on a finite triangle-free graph (using a different edge relation) such that whenever there is an edge in the latter there is a corresponding edge in the former. This is a nesting of the collection of finite triangle-free graphs in the collection of finite graphs. The Fraïssé limit of the joint age consists of a copy of the Rado graph overlaid on a copy of the Henson triangle-free graph (using different edge relations) such that whenever a pair of vertices has an edge in the latter, it has one in the former.

Given a nesting A of A_0 in A_1 as in Definition 7.2, we will now describe a *random* L_{tr}-structure $\mathcal{M}$ that is a.s. a transformation of $\mathcal{M}|_{L^0}$ into $\mathcal{M}|_{L^1}$, and for which

$\mathcal{M}|_{L^0\cup L^1}$ is a Fraïssé limit of A, almost surely. Furthermore, the distribution of $\mathcal{M}$ will be invariant under arbitrary permutations of the underlying set.

Because A has the strong amalgamation property, there is some probability measure μ on $\mathrm{Str}_{L^0\cup L^1}$, invariant under S_∞, that is concentrated on the isomorphism class of the Fraïssé limit of A. Our procedure starts by first sampling μ to obtain a random structure $\mathcal{N} \in \mathrm{Str}_{L^0\cup L^1}$.

Conditioned on $\mathcal{N}$, for every relation symbol $R \in L$ and every $j_1, \ldots, j_n \in \mathbb{N}$, where n is the arity of R, choose $r_{R,j_1,\ldots,j_n} \in \mathbb{R}$ as follows. If

$$\mathcal{N} \models \neg R^0(j_1, \ldots, j_n) \wedge R^1(j_1, \ldots, j_n),$$

then independently choose a real number $r_{R,j_1,\ldots,j_n} \in (0, 1)$ uniformly at random; if

$$\mathcal{N} \models \neg R^0(j_1, \ldots, j_n) \wedge \neg R^1(j_1, \ldots, j_n),$$

then let $r_{R,j_1,\ldots,j_n} := 2$, so that $R^s(j_1, \ldots, j_n)$ will not hold for any s; otherwise let $r_{R,j_1,\ldots,j_n} := 0$. Define $\mathcal{M}$ to be the L_{tr}-structure such that for all $s \in \mathbb{Q} \cap [0, 1]$,

$$\mathcal{M} \models R^s(j_1, \ldots, j_n)$$

if and only if $s \geq r_{R,j_1,\ldots,j_n}$, for all $R \in L$ and every $j_1, \ldots, j_n \in \mathbb{N}$, where n is the arity of R.

The real $r_{R,j_1,\ldots,j_n}$ can be thought of as the point in time at which $R(j_1, \ldots, j_n)$ "appears", in that it flips from not holding (in sublanguages L^s for $s < r_{R,j_1,\ldots,j_n}$) to holding (in sublanguages L^s for $s \geq r_{R,j_1,\ldots,j_n}$). Each $\mathcal{M}|_{L^s}$ then provides a "snapshot" of the structure over time as it transitions from $\mathcal{M}|_{L^0}$ to $\mathcal{M}|_{L^1}$, whereby the relations hold of more and more tuples. In particular, for any tuple and relation (of the same arity), the set of "times" for which the relation holds of the tuple is upwards-closed.

Note that whenever there are such points $r_{R,j_1,\ldots,j_n}$ other than 0 and 2, i.e., when there is some tuple of which a relation holds in $\mathcal{M}|_{L^1}$ but not in $\mathcal{M}|_{L^0}$, then any two independent samples from the distribution of $\mathcal{M}$ are a.s. non-isomorphic, as their respective sets of transition points are a.s. distinct. Hence, under this hypothesis, the distribution of $\mathcal{M}$ is an invariant measure that assigns measure 0 to every isomorphism class of L_{tr}-structures.

7.2. Open questions

In this paper, we have given conditions on a first-order theory that ensure the existence of an invariant measure concentrated on the class of its models but on no single isomorphism class; but a complete characterization has yet to be determined. It would be interesting also to characterize the structure of these invariant measures.

Another question is to find conditions under which one can formulate similar results for appropriate models of more sparse structures. Various notions of sparse graphs and intermediate classes have recently been studied extensively (see, e.g., [35] and [36]); for a presentation of graph limits for bounded-degree graphs, see [7].

One may also ask whether one can obtain measures concentrated on the class of models of the theory of continuous transformations described in §7.1.3, and still not on any single isomorphism class, in a "random-free" way, i.e., by sampling from a (two-valued) continuum-sized structure, as in our main construction.

Acknowledgments

This research was facilitated by participation in the workshop on Graph and Hypergraph Limits at the American Institute of Mathematics (Palo Alto, CA), the second workshop on Graph Limits, Homomorphisms and Structures at Hraniční Zámeček (Czech Republic), the conference on Graphs and Analysis at the Institute for Advanced Study (Princeton, NJ), the Arbeitsgemeinschaft on Limits of Structures at the Mathematisches Forschungsinstitut Oberwolfach (Germany), and the Workshop on Homogeneous Structures of the Hausdorff Trimester Program on Universality and Homogeneity at the Hausdorff Research Institute for Mathematics (Bonn, Germany).

Work on this publication by CF was made possible through the support of NSF grant DMS-0901020, ARO grant W911NF-13-1-0212, and 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 or the U.S. Government.

Work by JN has been partially supported by the Project LL-1201 ERCCZ CORES and by CE-ITI P202/12/G061 of the GAČR.

The authors would like to thank Daniel Roy and the referee for helpful comments, Alex Kruckman for detailed suggestions on an earlier draft, and Robert Israel for the statement and proof of Lemma 3.8.

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

[2] A. M. Vershik, A random metric space is a Urysohn space, Dokl. Akad. Nauk 387 (6) (2002) 733–736.

[3] A. M. Vershik, Random metric spaces and universality, Uspekhi Mat. Nauk 59 (2(356)) (2004) 65–104. doi:10.1070/RM2004v059n02ABEH000718.
URL http://dx.doi.org/10.1070/RM2004v059n02ABEH000718

[4] F. Petrov, A. Vershik, Uncountable graphs and invariant measures on the set of universal countable graphs, Random Structures & Algorithms 37 (3) (2010) 389–406. doi:10.1002/rsa.20309.
URL http://dx.doi.org/10.1002/rsa.20309

[5] N. Ackerman, C. Freer, R. Patel, Invariant measures concentrated on countable structures, ArXiv e-print 1206.4011.

[6] L. Lovász, B. Szegedy, Limits of dense graph sequences, J. Combin. Theory Ser. B 96 (6) (2006) 933–957. doi:10.1016/j.jctb.2006.05.002.
URL http://dx.doi.org/10.1016/j.jctb.2006.05.002

[7] L. Lovász, Large networks and graph limits, Vol. 60 of American Mathematical Society Colloquium Publications, American Mathematical Society, Providence, RI, 2012.

[8] D. J. Aldous, Representations for partially exchangeable arrays of random variables, J. Multivariate Anal. 11 (4) (1981) 581–598.

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

[10] O. Kallenberg, Symmetries on random arrays and set-indexed processes, J. Theoret. Probab. 5 (4) (1992) 727–765. doi:10.1007/BF01058727.
URL http://dx.doi.org/10.1007/BF01058727

[11] A. M. Vershik, Classification of measurable functions of several arguments, and invariantly distributed random matrices, Funktsional. Anal. i Prilozhen. 36 (2) (2002) 12–27, 95. doi:10.1023/A:1015662321953.
URL http://dx.doi.org/10.1023/A:1015662321953

[12] P. Diaconis, S. Janson, Graph limits and exchangeable random graphs, Rend. Mat. Appl. (7) 28 (1) (2008) 33–61.

[13] T. Austin, On exchangeable random variables and the statistics of large graphs and hypergraphs, Probab. Surv. 5 (2008) 80–145. doi:10.1214/08-PS124.
URL http://dx.doi.org/10.1214/08-PS124

[14] J. T. Baldwin, M. C. Laskowski, S. Shelah, Constructing Borel models in the continuum, In preparation (2015).

[15] S. Shelah, Classification theory and the number of nonisomorphic models, 2nd Edition, Vol. 92 of Studies in Logic and the Foundations of Mathematics, North-Holland Publishing Co., Amsterdam, 1990.

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

[17] D. Marker, Model theory, Vol. 217 of Graduate Texts in Mathematics, Springer-Verlag, New York, 2002.

[18] W. Hodges, Model theory, Vol. 42 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1993. doi:10.1017/CBO9780511551574.

[19] H. Becker, A. S. Kechris, The descriptive set theory of Polish group actions, Vol. 232 of London Mathematical Society Lecture Note Series, Cambridge University Press, Cambridge, 1996. doi:10.1017/CBO9780511735264.
URL http://dx.doi.org/10.1017/CBO9780511735264

[20] A. H. Lachlan, R. E. Woodrow, Countable ultrahomogeneous undirected graphs, Trans. Amer. Math. Soc. 262 (1) (1980) 51–94. doi:10.2307/1999974.
URL http://dx.doi.org/10.2307/1999974

[21] P. Hell, J. Nešetřil, Graphs and homomorphisms, Vol. 28 of Oxford Lecture Series in Mathematics and its Applications, Oxford University Press, Oxford, 2004. doi:10.1093/acprof:oso/9780198528173.001.0001.
URL http://dx.doi.org/10.1093/acprof:oso/9780198528173.001.0001

[22] O. Kallenberg, Probabilistic symmetries and invariance principles, Probability and its Applications, Springer-Verlag, New York, 2005.

[23] L. Lovász, B. Szegedy, Random graphons and a weak Positivstellensatz for graphs, J. Graph Theory 70 (2) (2012) 214–225. doi:10.1002/jgt.20611.
URL http://dx.doi.org/10.1002/jgt.20611

[24] R. Fraïssé, Sur certaines relations qui généralisent l'ordre des nombres rationnels, C. R. Acad. Sci. Paris 237 (1953) 540–542.

[25] M. Hušek, Urysohn universal space, its development and Hausdorff's approach, Topology Appl. 155 (14) (2008) 1493–1501. doi:10.1016/j.topol.2008.03.020.

URL http://dx.doi.org/10.1016/j.topol.2008.03.020
[26] P. S. Urysohn, Sur les espace métrique universel, Bull. Sci. Math. 51 (1927) 43–64, 74–90.
[27] M. Katětov, On universal metric spaces, in: General topology and its relations to modern analysis and algebra, VI (Prague, 1986), Vol. 16 of Res. Exp. Math., Heldermann, Berlin, 1988, pp. 323–330.
[28] J. Hubička, J. Nešetřil, A finite presentation of the rational Urysohn space, Topology Appl. 155 (14) (2008) 1483–1492. doi:10.1016/j.topol.2007.06.022.
URL http://dx.doi.org/10.1016/j.topol.2007.06.022
[29] P. J. Cameron, A. M. Vershik, Some isometry groups of the Urysohn space, Ann. Pure Appl. Logic 143 (1-3) (2006) 70–78. doi:10.1016/j.apal.2005.08.001.
URL http://dx.doi.org/10.1016/j.apal.2005.08.001
[30] C. Ealy, I. Goldbring, Thorn-Forking in continuous logic, J. Symbolic Logic 77 (1) (2012) 63–93. doi:10.2178/jsl/1327068692.
URL http://dx.doi.org/10.2178/jsl/1327068692
[31] J. Melleray, Some geometric and dynamical properties of the Urysohn space, Topology Appl. 155 (14) (2008) 1531–1560. doi:10.1016/j.topol.2007.04.029.
URL http://dx.doi.org/10.1016/j.topol.2007.04.029
[32] V. G. Pestov, A theorem of Hrushovski-Solecki-Vershik applied to uniform and coarse embeddings of the Urysohn metric space, Topology Appl. 155 (14) (2008) 1561–1575. doi:10.1016/j.topol.2008.03.002.
URL http://dx.doi.org/10.1016/j.topol.2008.03.002
[33] A. Usvyatsov, Generic separable metric structures, Topology Appl. 155 (14) (2008) 1607–1617. doi:10.1016/j.topol.2007.08.023.
URL http://dx.doi.org/10.1016/j.topol.2007.08.023
[34] S. Janson, Graphons, cut norm and distance, couplings and rearrangements, Vol. 4 of New York Journal of Mathematics. NYJM Monographs, State University of New York University at Albany, Albany, NY, 2013.
[35] J. Nešetřil, P. Ossona de Mendez, Sparsity, Vol. 28 of Algorithms and Combinatorics, Springer, Heidelberg, 2012. doi:10.1007/978-3-642-27875-4.
URL http://dx.doi.org/10.1007/978-3-642-27875-4
[36] J. Nešetřil, P. Ossona de Mendez, A unified approach to structural limits (with application to the study of limits of graphs with bounded tree-depth), ArXiv e-print 1303.6471.