On 0,1-laws and super-simplicity

Cameron Donnay Hill

Abstract

In Hill [9], it is shown (as a special case) that for an algebraically trivial Fraïssé class \( K \), if the generic theory \( T_K \) is an almost-sure theory by way of an asymptotic probability measure with independent sampling, then \( T_K \) is super-simple of \( SU \)-rank 1. Proving the converse requires a method for producing asymptotic probability measures with the appropriate properties, and this turns out to be rather difficult.

Viewing super-simplicity as just one level of an infinite hierarchy of higher amalgamation properties, Kruckman [14] uses all of these properties to produce asymptotic probability measures that yield the 0,1-law. The class of theories that meet the requirements of that construction is certainly smaller than the super-simple theories, likely much smaller. Here, building on Ackerman-Freer-Patel [2], Lovasz [16], and related work, we give a much more subtle construction of useful asymptotic probability measures using nothing more than super-simplicity cast as a single higher amalgamation property. With this construction, we show that if \( T_K \) is super-simple of \( SU \)-rank 1, then it is an almost-sure theory. This resolves a conjecture of Hill-Kruckman [10], an open question on the generic tetrahedron-free 3-hypergraph, and an (apparently) open question of Cherlin [4].

1 Introduction

1.1 Background

The first notice of the phenomenon of 0,1-laws for first-order logic over classes of finite structures lies in [6] and [7], where the 0,1-law for finite graphs is demonstrated, meaning that the almost-sure theory is complete. Naturally, there have been a number of developments in the line of demonstrating 0,1-laws for other classes of finite structures. This includes [13], for example, demonstrating the 0,1-law for the class of \( K_n \)-free graphs (fixed but arbitrary \( n \geq 3 \)) relative to uniform probability measures. A second example is in work on “sparse” classes of structures, as in [3, 15, 20]. The project of this paper is slightly different: We wish to answer the question, “What are almost-sure theories like?” To address this question, we restrict our attention to a smaller part of the 0,1-law phenomenon:

- The asymptotic probability measures in question must be compatible with the invariant-measures framework of [2]. The work of [6] and [7] falls under this rubric, but the uniform measures on triangle-free graphs and the measures associated with classes of sparse structures (e.g. edge-probability \( n^{-\alpha} \) for irrational \( \alpha > 0 \)) are excluded.

- We interest ourselves only in 0,1-laws for Fraïssé classes \( K \) such that the almost-sure theory is equal to the generic theory \( T_K \) of \( K \) (the theory of the Fraïssé limit). This excludes [13] again (since the almost-sure theory is the theory of the generic bi-partite graph), but [6] and [7] remain.
A consequence of these restrictions, proved in [8], is that if \( T_K \) is an almost-sure theory under these restrictions, then it must be algebraically trivial. Even with these restrictions, though, there seems to be a very rich theory to investigate both for its own sake and for its relevance to the more general phenomenon of pseudo-finiteness of \( \aleph_0 \)-categorical theories. A much deeper result than algebraic triviality was obtained in [9], showing that if the asymptotic probability measure in question is “reasonable” in that it has good conditional independence properties, then the generic theory – identical to the almost-sure theory – is super-simple of rank 1; even more, the Fraïssé class is “almost” a 1-dimensional asymptotic class in the sense of [17], and its generic theory is measurable:

\textbf{Theorem 1.1.1} ([9]). Let \( K \) be a combinatorial class.\(^1\) Suppose there is an asymptotic probability measure \( m \) for \( K \) with independent sampling such that \( T^m = T_K \). Then \( T_K \) is MS-measurable over the counting/algebraic dimension, and from this, it follows that \( T_K \) super-simple of SU-rank 1.

\subsection*{1.2 Contributions}

A bit of reflection on Theorem 1.1.1 reveals that a proof of its converse would yield a complete characterization of “strongly” almost-sure combinatorial theories without any reference to probabilities at all. Moreover, such a proof would, presumably, require a general method for producing asymptotic probability measures with appropriate properties for any super-simple combinatorial theory. Our main contribution in this paper is precisely a proof of the converse of Theorem 1.1.1 (as conjectured in [10]), and indeed, we present a general method of generating asymptotic probability measures of the right kind. That is, we prove:

\textbf{Theorem 4.1.1.} Let \( K \) be a combinatorial class. The following are equivalent:

1. \( T_K \) is super-simple of SU-rank 1.

2. There is an asymptotic probability measure \( m \) for \( K \) with independent sampling whose almost-sure theory \( T^m \) is equal to \( T_K \).

As an immediate consequence of Theorem 4.1.1, we find that whenever \( T_K \) is super-simple of SU-rank 1, we obtain (modulo the asymptotic probability measure) tight estimates of cardinalities of definable sets in members of \( K \). Moreover, for combinatorial theories, super-simplicity and MS-measurability are actually identical.

\textbf{Corollary 1.2.1.} Let \( K \) be a combinatorial class. If \( T_K \) is super-simple of SU-rank 1, then it is MS-measurable over the counting/algebraic dimension, and \( K \) itself is “almost” a 1-dimensional asymptotic class.

We also resolve an open question about concrete combinatorial classes. To the best of our knowledge, it was not know previously whether the theory of the generic tetrahedron-free 3-hypergraph is even pseudo-finite. As a corollary of Theorem 4.1.1, we find that this theory is not just pseudo-finite but that it is an almost-sure theory.\(^2\)

\(^1\)Combinatorial class means “algebraically trivial Fraïssé class in a countable relational language with finitely many isomorphism types of each finite cardinality.” A combinatorial theory, then, is just the generic theory of a combinatorial class.

\(^2\)The construction from [2] of invariant measures for arbitrary algebraically trivial countable structures is, perhaps, complicated enough to justify our previous ignorance of whether \( T_{G_{3,4}} \) can be recovered as an almost-sure theory.
Corollary 1.2.2. In the language $\mathcal{L}$ with one 3-ary relation symbol, and let $G_{3,4}$ be the Fraïssé class of all finite tetrahedron-free 3-hypergraphs. Then $G_{3,4}$ is a combinatorial class, and its generic theory $T_{G_{3,4}}$ is super-simple of SU-rank 1. Thus, there is an asymptotic probability measure $m$ for $G_{3,4}$ with independent sampling such that $T^m = T_{G_{3,4}}$.

This result generalizes to a very large family of combinatorial classes in finite relational languages. For a finite relational language $\mathcal{L}$, let $S = S(\mathcal{L})$ be the class of finite $\mathcal{L}$-structures $A$ such that for every $n$-ary relation symbol, (i) $R^A(\pi) \Rightarrow R^A(a_{\sigma(0)}, \ldots, a_{\sigma(n-1)})$ for each $\sigma \in \text{Sym}(n)$, and (ii) $R^A(\pi) \Rightarrow a_i \neq a_j$ whenever $i < j < n$. For each $k$, a structure $A \in S$ is $k$-irreducible if for any $a_0, \ldots, a_{k-1} \in A$, there are $R \in \text{sig}(\mathcal{L})$ and $\bar{b} \in R^A$ such that $a_i \in \bar{b}$ for each $i < k$. For $F \subset S$, we write $S_F$ to denote the class of structures $B \in S$ such that for no $A \in F$ is there an injective homomorphism $A \to B$. It is well-known that $S_F$ is a combinatorial class provided that every $A \in F$ is 2-irreducible. In unpublished work, G. Conant has shown that for $F \subset S$ such that $S_F$ is a combinatorial class, $T_{S_F}$ is super-simple of SU-rank 1 if and only if every $A \in F$ is 3-irreducible. This fact together with Theorem 4.1.1 gives us a complete characterization of almost-sure theories of classes of “societies” (in the language of [19]) obtained by excluding/forbidding irreducible structures – again, eliminating any reference to probabilities. By choosing $\mathcal{L}$ in the right way, we then use Corollary 1.2.3 resolve a question of Cherlin [4].

Corollary 1.2.3. Let $\mathcal{L}$ be a finite relational language, and let $F \subset S$ such that every $A \in F$ is 2-irreducible. The following are equivalent:

1. Every $A \in F$ is 3-irreducible.

2. $T_{S_F}$ is an almost-sure theory by way of an asymptotic probability measure with independent sampling.

1.3 Outline of the paper

In Section 2, we formally define combinatorial classes and their generic models. We also introduce a family of higher amalgamation properties (appearing also in [14] with a slightly different names), and we give the characterization of super-simplicity in terms of one of these properties. Finally, we translate types over the generic model into certain kinds of expansions of the generic model, and we show that super-simplicity is sufficient for the existence of a generic type (in a Baire category sense).

In Section 3, we recall three different sorts of measures that are relevant to model theory – Keisler measures, invariant measures, and asymptotic probability measures – and we expose how these three notions are related to each other. We recall a relatively innocuous theorem of [10] which draws an equivalence between asymptotic probability measures with independent sampling and certain kinds of Keisler measures. This result is used later in the proof of Theorem 4.1.1.

In work spread across Sections ?? and 4, we show that a “balanced” invariant measure concentrated on the generic type engenders the sort of Keisler measures that, in turn, yield asymptotic probability measures with independent sampling. That is, we show that the existence of such a balanced invariant measure is enough to prove Theorem 4.1.1, and then (in Section 4), we show how to construct balanced invariant measures from certain $a$ priori weaker invariant measures.

In Section 5, we collect some further consequences of Theorem 4.1.1. We prove a generalization of a theorem of [14] that allows one to prove pseudo-finiteness of a combinatorial theory by showing that
it is approximable by simple theories in a certain sense. We also make some conjectures connecting our work in this paper to a program for resolving Cherlin’s Question: “Is the Henson graph pseudo-finite?” Finally, we use Corollary 1.2.3 to address another question of Cherlin (Problem A of [4]): “Are there uncountably many pseudo-finite countable $\mathbb{R}_0$-categorical structures?”
2 Combinatorial classes, higher amalgamation, and expansions of generic models

In this section, we introduce the definitions and conventions that \textit{a priori} do not pertain directly to pseudo-finiteness or 0,1-laws. The most useful new result here is Theorem 2.3.4, which asserts that super-simplicity with \(SU\)-rank 1 is sufficient for the existence of a generic type.\(^3\) Later, concentrating measure on the isomorphism type of a structure associated with the generic type is a key step in the proof of Theorem 4.1.1.

2.1 Combinatorial classes

Throughout this paper, the signature \(\text{sig}(L)\) of any language \(L\) under discussion will consist of countably many relation symbols and no constant or function symbols. We often write \(R(n) \in \text{sig}(L)\) to mean that \(R\) is an \(n\)-ary relation symbol in the signature underlying \(L\).

For infinite structures, our notation follows that of [18]. So, for example, we use calligraphic upper-case letters like \(M\) to denote infinite structures with universe \(M\), and in general, our notation for such structures is more or less standard. For finite structures, we use simple upper-case letters like \(A, B, C\), and we identify finite structures with their universes unless some other set is explicitly mentioned. For a subset \(A\) of \(M\), where \(M\) is an infinite structure, \(M \upharpoonright A\) is the induced substructure of \(M\) with universe \(A\); when no confusion is likely to arise, however, we will often write \(A\) instead of \(M \upharpoonright A\). If \(B\) is a finite structure and \(X \subseteq B\), then \(B \upharpoonright X\) is likewise the induced substructure of \(B\).

A structure \(M\) is algebraically trivial if \(\text{acl}_M(A) = A\) for all \(A \subseteq M\). When \(M\) is \(\aleph_0\)-categorical (by convention, in a countable language), we often tacitly identify a type \(p(x)\) over a finite set \(C \subseteq \text{fin} M\) with a formula that isolates it; in any case, the solution set \(p(M)\) is a definable set. We write \(S^*_1(M)\) to denote the set of complete types \(p(x)\) over \(M\) such that \(p(x) \models x \neq a\) for each \(a \in M\).

We coin the term “combinatorial class” mainly as a convenience — it allows us to avoid saying “algebraically trivial Fraïssé class” over and over again.

\textbf{Definition 2.1.1} (Combinatorial classes). Let \(K\) be a class of finite \(L\)-structures. Then \(K\) is a \textit{combinatorial class} if it has the following properties:

- \(K\) is closed under isomorphism, and for every \(n\), the set \(\{A \in K : |A| = n\} / \cong\) is finite.
- (Heredity property - HP) For any \(B \in K\) and any induced substructure \(A \leq B, A \in K\) as well.
- (Disjoint joint-embedding property - disjoint-JEP) For any \(A_1, A_2 \in K\), there are \(B \in K\) and embeddings \(f_i : A_i \to B (i = 1, 2)\) such that \(f_1 A_1 \cap f_2 A_2 = \emptyset\).
- (Disjoint amalgamation property - disjoint-AP) For all \(A, B_1, B_2 \in K\) and embeddings \(f_i : A \to B_i (i = 1, 2)\), there are \(C \in K\) and embeddings \(g_i : B_i \to C\) such that \(g_1 \circ f_1 = g_2 \circ f_2\) and \(g_1 B_1 \cap g_2 B_2 = g_1 f_1 A = g_2 f_2 A\).

Thus, a combinatorial class is just a slightly special sort of Fraïssé class.

\(^3\)The terminology “generic type” is appropriate here because it is generic in the sense of Baire category. It’s just a little unfortunate that “generic type” is also used heavily in the model theory of groups.
The “fundamental” theorem of combinatorial classes is the following (whose proof can be found in [11] and several other places):

**Theorem 2.1.2.** If $K$ is a combinatorial class, then there is a countably infinite structure $M$, a generic model of $K$, with the following properties:

- *(K-closedness)* For every $X \subset_{fin} M$, $M|X \in K$.
- *(K-universality)* For every $A \in K$, there is an embedding $A \to M$.
- *(Ultraphomengeity)* For any $A, B \in K$ with $A \leq B$ and any embedding $f_0 : A \to M$, there is an embedding $f : B \to M$ such that $f_0 \subseteq f$.
- $M$ is algebraically trivial.

Any countable structure with the first three of these properties is isomorphic to $M$, so $M$ is usually called the generic model of $K$ (also called the Fraïssé limit of $K$), and the generic theory $T_K = Th(M)$ is well-defined in terms of $K$. This theory $T_K$ is $\aleph_0$-categorical, eliminates quantifiers, and every one of its models is algebraically trivial.

**Remark 2.1.3.** Suppose $T$ is an algebraically trivial $\aleph_0$-categorical theory in a countable language. Then there is a combinatorial class $K$ such that $T$ and $T_K$ are inter-definable – take $K$ to be the age of the expansion of the countable model $M \models T$ obtained by adding relation symbols to name every $0$-definable set (i.e. by morleyizing $T$). Thus, if we wish to study pseudo-finiteness of algebraically trivial $\aleph_0$-categorical theories, we lose nothing in restricting our attention to combinatorial classes and their generic theories.

**From now on...**

Throughout the remainder of this paper, $K$ is a combinatorial class with language $\mathcal{L}$, and $M$ is (a fixed copy of) its generic model, with universe $M$.

### 2.2 Higher amalgamation

We introduce a hierarchy of higher amalgamation properties for classes of finite. With a slightly different notation, these appeared in [14], and very similar, if not identical, notions appear in [21] and in the literature on abstract elementary classes.

**Definition 2.2.1.** Let $2 \leq n < \omega$. As usual, it seems, we write $\mathcal{P}^-(n)$ for the set of proper subsets of $n = \{0, 1, ..., n - 1\}$. An $n$-disjoint amalgamation problem in $K$ ($n$-DAP problem) is a family $(A_s)_{s \in \mathcal{P}^-(n)}$ of structures in $K$ such that for all $s, t \subseteq n$:

- $s \subseteq t \Rightarrow A_s \leq A_t$;
- $A_s \cap A_t = A_{s \cap t}$ as sets.

A solution of $(A_s)_{s \in \mathcal{P}^-(n)}$ in $K$ is a structure $B \in K$ such that $A_s \leq B$ for each $s \subseteq n$. We allow $A_{\varnothing} = \varnothing$, but this is not required. We say that $K$ has $n$-DAP if every $n$-DAP problem in $K$ has a solution in $K$.

We note that any class can have $n$-DAP or not for any $n$. In fact, “$K$ is a combinatorial class” is equivalent to “$K$ has HP and 2-DAP.”
It is not difficult to verify that the class of all finite graphs has \( n \)-DAP for every \( n \), and in general, \( S(\mathcal{L}) \) has \( n \)-DAP for every \( n \), whenever \( \mathcal{L} \) is a finite relational language. On the other hand, the class \( \text{H} \) of finite triangle-free graphs does not have 3-DAP, and the class \( \text{G}_{3,4} \) has 3-DAP but does not have 4-DAP. “\( n \)-DAP for every \( n \)” makes it easy to produce asymptotic probability measures that engender \( 0,1 \)-laws (see Definition 3.1.1), but it is much less obvious that 3-DAP alone is sufficient. However, 3-DAP does have an immediate equivalent in model-theoretic terminology, as expressed in Proposition 2.2.3 below.

**Observation 2.2.2.** Suppose \( T_K \) is super-simple of \( SU \)-rank 1. Then:

1. \( T_K \) has trivial forking dependence (in the real sorts): \( A \downarrow C B \Leftrightarrow A \cap B \subseteq C \).
2. Every finite subset \( C \subset \text{fin} M \) is an amalgamation base for non-forking independence (in real sorts).

**Proposition 2.2.3.** The following are equivalent:

1. \( T_K \) is super-simple of \( SU \)-rank 1.
2. \( K \) has 3-DAP.

**Proof.** For 1\( \Rightarrow \)2: Let \( (A_s)_{s \in \mathcal{P}-(3)} \) be a 3-DAP problem in \( K \). We may rename the structures involved as follows

- \( A_{\emptyset} = C \);
- \( A_{\{1\}} = B_1, A_{\{2\}} = B_2, \) and \( A_{\{3\}} = C \bar{\tau} \), where \( \tau \cap B_1 B_2 = \emptyset \);
- \( A_{\{1,3\}} = B_1 \bar{\tau}, A_{\{2,3\}} = B_2 \bar{\tau}, \) and \( A_{\{1,2\}} = B_1 B_2 \).

Since \( K \) has 2-DAP, we may assume that \( C, B_1, B_2 \) are substructures of \( M \) and that there are \( \bar{a}_1, \bar{a}_2 \) in \( M \) such that \( B_1 \bar{a}_1 \cong_{B_1} B_1 \tau \bar{\tau} \) and \( B_2 \bar{a}_2 \cong_{B_2} B_2 \bar{\tau} \). By item 1 (since \( \bar{a}_1 \cap B_1 = \emptyset \subseteq C \)), we see that \( B_1 \cap B_2 = C, \bar{a}_1 \equiv_C \bar{a}_2, B_1 \downarrow C B_2, \bar{a}_1 \downarrow C B_1, \) and \( \bar{a}_2 \downarrow C B_2 \). Also by item 1, there is an \( \bar{a} \) in \( M \) such that \( \bar{a} \equiv_{B_1} \bar{a}_1, \bar{a} \equiv_{B_2} \bar{a}_2, \) and \( \bar{a} \downarrow C B_1 B_2 \). Setting \( D = B_1 B_2 \bar{a} \), we have a solution of \( (A_s) \) in \( K \).

For 2\( \Rightarrow \)1 (from [14]): One easily verifies that \( \downarrow \) has all the properties of an independence relation required except possibly independent-amalgamation (where \( A \downarrow C B \) means \( A \cap B \subseteq C \)). We will show that if \( K \) has 3-DAP, then \( \downarrow \) has independent-amalgamation – in which case, it must be non-forking independence in a simple theory, yielding item 1. So, let \( C, B_1, B_2 \subset \text{fin} M, \bar{a}_1, \bar{a}_2 \in M^n \) such that \( B_1 \cap B_2 = C \) (so \( B_1 \downarrow C B_2 \)), \( \bar{a}_1 \equiv_C \bar{a}_2, \bar{a}_1 \downarrow C B_1, \) and \( \bar{a}_2 \downarrow C B_2 \). We pick a new tuple of elements \( \bar{\tau} \) and make structures in \( K \) as follows:

- \( A_{\emptyset} = C \)
- \( A_{\{1\}} = B_1, A_{\{2\}} = B_2, \) and \( A_{\{3\}} = C \bar{\tau} \cong_C C \bar{a}_1 \cong_C C \bar{a}_2 \).
- \( A_{\{1,3\}} = B_1 \bar{\tau} \cong_{B_1} B_1 \bar{a}_1, A_{\{2,3\}} = B_2 \bar{\tau} \cong_{B_2} B_2 \bar{a}_2, \) and \( A_{\{1,2\}} = B_1 B_2 \)

We have constructed a 3-DAP problem \( (A_s)_{s \in \mathcal{P}-(3)} \) in \( K \), so since \( K \) has 3-DAP, there is a solution for it, \( D \in K \). It follows that \( \text{qftp}(\bar{a}_1/B_1) \cup \text{qftp}(\bar{a}_2/B_2) \cup \"\bar{x}_1 \cap \bar{x}_2 = \emptyset\" \) is consistent, and any realization of this type suffices. \( \Box \)
2.3 Special expansions

In this subsection, we will see that 3-DAP is sufficient for the existence of a generic type over $\mathcal{M}$. To prove existence and to use generic types later on, it is natural recast types $q(x) \in S_1^*(\mathcal{M})$ as expansions $\mathcal{M}_q$ of $\mathcal{M}$. For the proof of Theorem 4.1.1, this point of view also provides us with a useful isomorphism-type on which to concentrate measure in the sense of [2]. The existence of generic types given 3-DAP is proved in Theorem 2.3.4 below, and that is the one important fact verified in this subsection.

**Definition 2.3.1.** We fix a type $p(x_0, \ldots, x_{k-1}) \in S(T_K)$ such that $p(\bar{x}) \models \bigwedge_{i<j} x_i \neq x_j$. We write $S_p^*(\mathcal{M})$ for the set of complete types $q(\bar{x})$ over $\mathcal{M}$ extending $p$ such that $\bar{a} \models q \Rightarrow \bar{a} \cap \mathcal{M} = \emptyset$.

Let $\tau_0^p$ be the signature consisting of relation symbols $R_p(n)$ for each type $p(\bar{x}, \bar{y}) \in S_n+k(T_K)$ such that $p(\bar{x}, \bar{y}) \models p(\bar{x}) \land "\bar{x} \cap \bar{y} = \emptyset" \land \bigwedge_{i<j} y_i \neq y_j$ $(n > 0)$. Then let $\mathcal{L}'$ be the language with $\text{sig}(\mathcal{L}') = \text{sig}(\mathcal{L}) \cup \tau_0^p$. For $q(\bar{x}) \in S_p^*(\mathcal{M})$, let $\mathcal{M}_q$ be the $\mathcal{L}'$-expansion of $\mathcal{M}$ such that $\bar{a} \in R_{\mathcal{M}_q}^p \Leftrightarrow p(\bar{x}, \bar{a}) \subset q(\bar{x})$ whenever $\bar{a} \in M^n$, $p(\bar{x}, \bar{y}) \in S_{n+k}(T_K)$ such that $p(\bar{x}, \bar{y}) \models p(\bar{x}) \land "\bar{x} \cap \bar{y} = \emptyset" \land \bigwedge_{i<j} y_i \neq y_j$ $(n > 0)$.

We define $\mathcal{K}^p$ to be the isomorphism-closure of

$$\{ \mathcal{M}_q | C : C \subset M, q \in S_p^*(\mathcal{M}) \}$$

which has HP but need not have AP or JEP in general.

Let’s say that a type $q(x) \in S_1^*(\mathcal{M})$ is $p$-generic if $\mathcal{K}^p$ is a combinatorial class and $\mathcal{M}_q$ is (a copy of) its generic model.

**Observation 2.3.2.** Fix $p(\bar{x}) \in S_k(T_K)$ such that $p(\bar{x}) \models \bigwedge_{i<j} x_i \neq x_j$. Let $A \in \mathcal{K}^p$ — say $|A| = n$ — and let $\bar{a}$ be an enumeration of $A$. Let $p(\bar{x}, \bar{y}) \in S_{n+k}(T_K)$ such that $A \models R_p(\bar{x})$, and let $A^- = A|\mathcal{L}$.

Then, the pair $(A^-, p(\bar{x}, \bar{a}))$ completely determines $A$ in the following sense: Suppose $B^- \in \mathcal{K}$ and $f : A^- \rightarrow B^-$ is an isomorphism of $\mathcal{L}$-structures; if $B \in \mathcal{K}^p$ is an expansion of $B^-$ such that $B \models R_p(f(\bar{a}))$, then $f$ is also an $\mathcal{L}'$-isomorphism $A \rightarrow B$. We also note that $p(\bar{x}, \bar{y})$ need not be a complete type — a quantifier-free-complete type would do.

This observation suggests a convenient notation for passing from structures $A \in \mathcal{K}$ to expansions $A' \in \mathcal{K}^p$. Namely, suppose $A \in \mathcal{K}$, $\bar{a}$ is an enumeration of $A$, and $p = p(\bar{x}, \bar{a})$ is a quantifier-free-complete type over $A$ such that $p(\bar{x}, \bar{a}) \models p(\bar{x}) \land "\bar{x} \cap \bar{y} = \emptyset"$. Then, $A+p$ denotes the unique structure $A' \in \mathcal{K}^p$ described in the previous paragraph.

**Lemma 2.3.3. If $\mathcal{K}$ has 3-DAP, then $\mathcal{K}^p$ is a combinatorial class for every $p$.**

**Proof.** Let $C, B_1, B_2 \in \mathcal{K}^p$ with $C \leq B_1 \leq B_2$, and $B_1 \cap B_2 = C$. We must show that there is some $D \in \mathcal{K}^+$ such that $B_1 \leq D$ and $B_2 \leq D$. Let $\mathcal{C}^- = C|\mathcal{L}'$, $B_1^- = B_1|\mathcal{L}'$, and $B_2^- = B_2|\mathcal{L}'$.

We allow $C$ to be the empty structure, and without loss of generality, we assume that $B_1^-, B_2^- < \mathcal{M}$.

Let $\bar{c}, \bar{b}_1, \bar{b}_2$ be enumerations of $C$, $B_1 \setminus C$, and $B_2 \setminus C$, respectively. Then there are quantifier-free-complete types $q(\bar{x}, \bar{y}), q_1(\bar{x}, \bar{y}, z_1), q_2(\bar{x}, \bar{y}, z_2)$ such that $C \models R_q(\bar{c}), B_1 \models R_{q_1}(\bar{c}, \bar{b}_1), B_2 \models R_{q_2}(\bar{c}, \bar{b}_2)$, and these completely determine $C, B_1, B_2$ over $\mathcal{C}^-, B_1^-, B_2^-$ in the sense of Observation 2.3.2. That is, $C = \mathcal{C}^- + q(\bar{x}, \bar{c})$, and so on. Clearly, $q_i(\bar{x}, \bar{y}, z_i) \models q(\bar{x}, \bar{y})$ for both $i = 1, 2$. Now, we define a 3-DAP problem in $\mathcal{K}$:

- $A_{\varnothing} = \mathcal{C}^-$;
• $A_{\{1\}} = B_1^-$ and $A_{\{2\}} = B_2^-$;

• $A_{\{3\}} = \tau C^-$, where $\tau$ is a $k$-tuple of distinct new elements outside of $B_1B_2$, and we require that $A_{\{3\}} \models q(\tau, \bar{\tau})$;

• $A_{\{1,2\}} = \mathcal{M}|(B_1B_2)$;

• for both $i = 1, 2$, $A_{\{i,3\}}$ is the $\mathcal{L}$-structure with universe $\bar{\tau}B_i^-$ extending $B_i^-$ and such that $A_{\{i,3\}} \models q(\tau, \bar{\tau}, \bar{b}_i)$.

Since $K$ has 3-DAP, we recover a solution $E^-$ of $(A_s)_{s \in \mathcal{P}^-}(3)$; we may assume that $A_s \subseteq E^- < M$ for each $s \subseteq 3$. Let $p(x, \bar{y}, \bar{z}_1, \bar{z}_2)$ be the quantifier-free-complete type such that $\mathcal{M} \models p(\bar{\tau}, \bar{\tau}, \bar{b}_1\bar{b}_2)$. Finally, we define $D$ to be the $\mathcal{L}^+_{\text{e}}$-structure such that $D|\mathcal{L} = \mathcal{M}|(B_1B_2)$ and $D \models R_p(\bar{\tau}, \bar{b}_1, \bar{b}_2)$. (As above, these data completely determine a unique $\mathcal{L}^+_{\text{e}}$-structure expanding $\mathcal{M}|(B_1B_2)$.) Since $p(\bar{\tau}, \bar{y}, \bar{z}_1, \bar{z}_2) \models q_i(x, \bar{y}, \bar{z}_1)$ for both $i = 1, 2$, we have $B_1 \leq D$ and $B_2 \leq D \in K^p$, and this completes the proof. \hfill \qed

**Theorem 2.3.4.** $K$ has 3-DAP if and only if there is a $p$-generic type over $\mathcal{M}$ for every $p \in S^*(T_K)$.

**Proof.** For “only if,” let $p(x_0, ..., x_{k-1})$ such that $p(\bar{\tau}) \equiv \bigwedge_{i \neq j} x_i \neq x_j$. Since $K^p$ is a combinatorial class (by Lemma 2.3.3), it has a generic model $M^p$, which we may assume is an expansion of $M$ — i.e., $M^p|\mathcal{L} = M$. For each $s \subset \text{fin} M$, we fix an enumeration $\bar{\tau}_s$ of $s$ and a type $p(\bar{\tau}, \bar{\tau}_s) \in S_{|s|+k}(T_K)$ such that $M^p \models R_{p_s}(\bar{\tau}_s)$ and so that the data $(M|s, p_s(\bar{\tau}, \bar{\tau}_s))$ determines $M^p|s$ in the sense of Observation 2.3.2. We observe that $p_t(\bar{\tau}, \bar{\tau}_t) \models p_s(x, \bar{\tau}_s)$ whenever $s \subseteq t \subset \text{fin} M$, and from this it follows that $p(x) = \bigcup\{p_s(\bar{\tau}, \bar{\tau}_s) : s \subset \text{fin} M\}$ is consistent. Since $\pi|C$ isolates a complete type over $C$ for every $C \subset \text{fin} M$, it’s also clear that $\pi$ extends uniquely to a complete type $q$ over $M$. Since for each $s \subset \text{fin} M$, $(M|s, p_s(\bar{\tau}, \bar{\tau}_s))$ determines $M^p|s$, we have $M^p|s < M_q$ for each $s \subset \text{fin} M$; hence $M_q = M^p$.

For “if”, let $(A_s)_{s \in \mathcal{P}^-}(3)$ be a 3-DAP problem in $K$. Then, let $\bar{\tau} = (e_0, ..., e_{k-1})$ enumerate $A_{\{0\}} \setminus A_{\{2\}}$, and set $p(\bar{\tau}) = \text{qftp}^{A(0)}(\bar{\tau})$. Moreover:

• Let $C = A_{\{2\}}$ and $q(\bar{\tau}, \bar{\epsilon}) = \text{qftp}^{A(0)}(\bar{\tau}/C)$.

• Let $B_1 = A_{\{1\}}$ and $B_2 = A_{\{2\}}$, and for $i = 1, 2$, set $q_i(\bar{\tau}, \bar{\epsilon}_i) = \text{qftp}^{A(0)}(\bar{\tau}/B_i)$, where $\bar{\epsilon}_i$ enumerates $B_i \setminus C$.

We may assume that $C < \mathcal{M}$. Consequently, there are embeddings $f_i : B_i \to \mathcal{M}$ such that $f_i|C = id_C$ ($i = 1, 2$) such that $f_1B_1 \cap f_2B_2 = C$. Now, let $C^+ = C + q$, $B_1^+ = B_1 + q_1$, and $B_2^+ = B_2 + q_2$. Since $K^p$ is a combinatorial class (has 2-DAP) and arguing as in the previous paragraph to convert a model $M' \models T_{K^p}$ into a type, there is a type $q \in S^+_k(M)$ extending $p$ such that $B_1^+ < M_q$ and $B_2^+ < M_q$. Since $B_1B_2$ is finite, there is a realization $\bar{\tau}'$ of $q|B_1B_2$ in $M$. Up to exchanging $\bar{\tau}$ for $\bar{\tau}'$, the structure $\mathcal{M}|B_1B_2\bar{\tau}'$ solves the original 3-DAP problem $(A_s)_{s \in \mathcal{P}^-}(3)$. \hfill \qed

**Observation 2.3.5.** Let $p(x_0, ..., x_{k-1}) \in S_k(T_K)$ and $p_0(x_0, ..., x_{m-1}) \in S_m(T_K)$, where $m \leq k$ and $p \models p_0$. Also, let $\sigma \in \text{Sym}(k)$, and let $p_\sigma$ be the type such that $p_\sigma(x_{\sigma(0)}, ..., x_{\sigma(k-1)}) = p(\bar{\tau})$.

1. $M^p|\mathcal{L}^{p_0} \cong M^{p_0}$ via an automorphism of $M$.

2. After applying an automorphism of $M$, $M^p$ and $M^{p_0}$ are inter-definable.
3 Definitions and some background: Keisler measures, invariant measures, and asymptotic probability measures

There are several different sorts of probability measures that have some relevance to model theory. Here, we make use of three of these, and we introduce their definitions here. The connections between them were investigated in some detail in [10], and we will make use of some of those results here.

3.1 Asymptotic probability measures

**Definition 3.1.1.** For each $s \subset \text{fin } M$, $K[s]$ is the set of structures $A \in K$ with universe exactly $s$. An asymptotic probability measure for $K$ (APM for $K$) is a family $m = (m_s)_{s \subset \text{fin } M}$ where:

- for each $s \subset \text{fin } M$, $m_s : K[s] \to [0, 1]$ is a probability mass function;
- for all $s \subseteq t \subset \text{fin } M$ and $A \in K[s]$, $m_s(A) = \sum_{B \in K[t]} m_t(B)$;
- for all $s \subset \text{fin } M$, $\sigma \in \text{Sym}(M)$, and $A \in K[s]$, $m_{\sigma s}(\sigma A) = m_s(A)$.

For a sentence $\varphi$ of $L$ and $s \subset \text{fin } M$, we define $P_{m_s}[\varphi] = \sum \{m_s(A) : A \models \varphi\}$. The almost-sure theory of $m$ is $T_m = \{\varphi \in \text{Sent}(\mathcal{L}) : \lim_{|s| \to \infty} P_{m_s}[\varphi] = 1\}$.

An APM $m = (m_s)_s$ for $K$ is said to have independent sampling if for all $s_0 \subset s$, $C \in K_{s_0}$, pairwise distinct $i_0, ..., i_{n-1} \in s \setminus s_0$, and $A_t \in K_{s_0 \cup \{i_t\}}$ extending $C$ ($t < n$), we have

$$P_{m_0}[A_0 \wedge ... \wedge A_{n-1} | C] = \prod_{t<n} P_{m_0}[A_t \mid C]$$

where $A_0 \wedge \cdots \wedge A_{n-1}$ stands for the event $\{B \in K_s : \bigwedge_{t<n} A_t \leq B\}$, and so forth.

**Theorem 3.1.2** (cite:Kruckman). Suppose $K$ has $n$-DAP for all $n$. Then there is an APM $m$ for $K$ with independent sampling such that $T_m = T_K$.

3.2 Probabilistic MS-measures

Here, we begin by introducing the original formulation of MS-measurability, which is appropriate for arbitrary infinite structures.

**Definition 3.2.1.** Let $\mathcal{N}$ be an infinite structure. Consider a pair of functions $\delta : \text{Def}(\mathcal{N}) \to N$ and $\mu : \text{Def}(\mathcal{N}) \to [0, \infty)$; we say that $\mathcal{N}$ is MS-measurable via $(\delta, \mu)$ if the following requirements are met:

- If $X$ is finite, then $\delta(X) = 0$ and $\mu(X) = |X|$.
- If $X$ is infinite, then $\delta(X) > 0$ and $\mu(X) > 0$.
- For each formula $\varphi(\overline{x}, \overline{y})$ in the language of $\mathcal{N}$, the set $H_\varphi = \{(\delta(\varphi(\mathcal{M}, \overline{b})), \mu(\varphi(\mathcal{M}, \overline{b}))) : \overline{b} \in M[\overline{y}]\}$

10
is finite, and for each pair \((d_0, m_0) \in H_\varphi\), the set
\[
\{ \bar{b} \in M^{[\overline{y}]} : (\delta(\varphi(M, \bar{b})), \mu(\varphi(M, \bar{b}))) = (d_0, m_0) \}
\]
is 0-definable.

- (Fubini) Let \(X, Y\) be definable sets, and let \(f : X \to Y\) be a surjective definable function. Suppose that for some \((d_0, m_0) \in \mathbb{N} \times [0, \infty)\), \(\delta(f^{-1}(\bar{b})) = d_0\) and \(\mu(f^{-1}(\bar{b})) = m_0\) for all \(\bar{b} \in Y\). Then
  \[
  \delta(X) = d_0 + \delta(Y),
  \mu(X) = m_0 \cdot \mu(Y).
  \]

When \(\mathcal{N}\) is MS-measurable via \((\delta, \mu)\), the function \(h : \text{Def}(M) \to \mathbb{N} \times [0, \infty) : X \mapsto (\delta(X), \mu(X))\) is called the measuring function, while \(\delta\) alone is called a dimension function.

A fundamental theorem of MS-measurable structures is the following, already remarked upon above. Of course, it is essential to our proof of Theorem ?? (on the super-simplicity of generic theories of 0,1-classes).

**Theorem 3.2.2** (Corollary 3.6 of [5]). Suppose \(\mathcal{N}\) is MS-measurable via \((\delta, \mu)\). Then for every \(X \in \text{Def}(\mathcal{N})\), \(D(X) \leq \delta(X)\). It follows that \(\text{Th}(\mathcal{N})\) is super-simple of finite SU-rank.

In our setting, the definable functions are severely restricted by the algebraic-triviality assumption, and the same assumption provides an especially natural candidate for the dimension function \(\delta\) – the “algebraic” dimension \(d\) of Definition 3.2.3. Subsequently, we may restrict our attention just to functions \(m\) (MS-measures) that form measuring functions when combined with this \(d\), and we can exploit the dearth of definable functions to formulate a version of MS-measurability that is particularly easy to verify, as in Definition 3.2.4.

**Definition 3.2.3.** For \(C \subset \text{fin} M\) and \(\overline{a} \in M^n\) \((0 < n < \omega)\), we define
\[
I(\overline{a}/C) = \{ i < n : a_i \notin C \},
E(\overline{a}/C) = \{ (i, j) \in I(\overline{a}/C)^2 : a_i = a_j \}, \text{ an equivalence relation}
\]
\[
d(\overline{a}/C) = \left| I(\overline{a}/C)/E(\overline{a}/C) \right|.
\]
We observe that
\[
d(\overline{a}\overline{b}/C) = d(\overline{a}/C) + d(\overline{b}/C)\overline{a}
\]
whenever \(C \subset \text{fin} M\) and \(\overline{a}, \overline{b} \in M^{<\omega}\). For a definable set \(X \subseteq M^n\), we define
\[
d(X) = \max \{ d(\overline{a}/C) : \overline{a} \in X, X \text{ is over } C \}
\]

**Definition 3.2.4.** We write \(S[\mathcal{M}]\) for the set of all types \(tp(\overline{a}/C)\), where \(C \subset \text{fin} M\) and \(\overline{a} \in M^{<\omega}\) is a non-empty. An **MS-measure on \(\mathcal{M}\)** is a map \(m : S[\mathcal{M}] \to [0, \infty)\) satisfying the following:

**MS-1.** \(m\) is \(\text{Aut}(\mathcal{M})\)-invariant

**MS-2.** If \(\overline{a} \subseteq C\), then \(m(\overline{a}/C) = 1\).
MS-3. If \( \overline{a} \not\in C \), then \( m(\overline{a}/C) > 0 \).

MS-4. For \( C_0 \subseteq C \subseteq_{\text{fin}} M \) and \( \overline{a} \in M^{<\omega} \),

\[
m(\overline{a}/C_0) = \sum_{i<k} m(\overline{a}_i/C)
\]

where \( \overline{a}_0, \ldots, \overline{a}_{k-1} \) is any set of \( \equiv_C \)-representatives of \( \{ \overline{a}' : \overline{a}' \equiv_{C_0} \overline{a} \land d(\overline{a}'/C) = d(\overline{a}/C_0) \} \).

MS-5. Fubini:

(a) Let \( C \subseteq_{\text{fin}} M \), \( 0 < n < \omega \), and \( \overline{a} \in M^n \), and let \( \sigma \in \text{Sym}(n) \). Then \( m(\overline{a}/C) = m(\overline{a}_\sigma/C) \)

where \( \overline{a}_\sigma = (a_{\sigma(0)}, \ldots, a_{\sigma(n-1)}) \).

(b) Let \( C \subseteq_{\text{fin}} M \) and \( \overline{a}, \overline{b} \in M^{<\omega} \); then \( m(\overline{a}/C)m(\overline{b}/C\overline{a}) \).

An MS-measure \( m : S[M] \to [0, \infty) \) is called probabilistic if the following holds:

Let \( C \subseteq_{\text{fin}} M \) and \( 0 < n < \omega \), and let \( \pi(\overline{a}) \) be a quantifier-free-complete \( n \)-type over \( C \) in the language of equality such that \( \pi \models "\overline{a} \not\in C." \) Then

\[
\sum_{p \in S_\pi(C)} m(p) = 1
\]

where \( S_\pi(C) = \{ p \in S_n(C) : \pi \subseteq p \} \).

We write \( \text{Meas}_{MS}(M) \) for the set of all probabilistic MS-measures on \( M \).

We have suggested strongly that the presence of MS-measures amounts to MS-measurability, so of course, we need to verify this claim. We remark that there is no need to restrict to probabilistic MS-measures at this point in the discussion, but it is not altogether clear that there are any non-probabilistic MS-measures.

**Proposition 3.2.5.** If \( m : S[M] \to [0, \infty) \) is an MS-measure on \( M \), then \( M \) is MS-measurable via \( (d,m^*) \), where for a definable set \( X \), we define \( S_X(C) = \{ \text{tp}(\overline{a}/C) : \overline{a} \in X, X \text{ is over } C \} \) and

\[
m^*(X) = \sup \left\{ \sum_{p \in S_X(C)} m(p) : X \text{ is over } C \right\}.
\]

**Proposition 3.2.6** (Theorem ## of CITE:Hill-Kruckman). APMs with independent sampling and probabilistic MS-measures correspond in the following way.

1. Let \( a \) be a positive APM for \( K \) with independent sampling. Define \( m_a : S^*[M] \to [0, 1] \) by setting

\[
m_a(\overline{a}/C) = \frac{a(\overline{a}/C)}{a(M(C))}
\]

for all \( C \subseteq_{\text{fin}} M \) and non-repeating \( \overline{a} \in M^{<\omega} \) disjoint from \( C \). Then \( m_a \) extends uniquely to a probabilistic MS-measure.

2. Let \( m : S[M] \to [0, 1] \) be a probabilistic MS-measure. Define \( a^m \) by setting

\[
a^m(M|A) = m(a_0)m(a_0/a_1) \cdots m(a_n/a_{<n})
\]

for \( A = \{ a_0, \ldots, a_n \} \subseteq_{\text{fin}} M \), and closing to enforce invariance. Then, this definition is well-defined on finite structures (i.e. the ordering of \( a_0, \ldots, a_n \) does not matter), and \( a^m \) is a positive APM with independent sampling.
3.3 Invariant measures

One of the results that animates our proof of Theorem 4.1.1 is the following. We apply Theorem 3.3.1 only for the special case of $\aleph_0$-categorical theories, so we will not explicate all of its moving parts in general.

**Theorem 3.3.1** ([2]). For a countable relational language, let $\text{Str}_{\mathcal{L}}$ be the space of all $\mathcal{L}$-structures $\mathcal{N}$ with universe $\omega$. For $\mathcal{N}_0 \in \text{Str}_{\mathcal{L}}$, the following are equivalent:

1. $\mathcal{N}_0$ has trivial group-theoretic algebraic closure.

2. There is a $\text{Sym}(\omega)$-invariant Borel probability measure $\mu$ on $\text{Str}_{\mathcal{L}}$ concentrated on the isomorphism-type of $\mathcal{N}_0$ (i.e. $\mu(\mathcal{N}_0/\cong) = 1$).

We now express the technology of Theorem 3.3.1 for the special case that interests us. (In particular, since we are working with $\aleph_0$-categorical theories, logical and group-theoretic algebraic closure operators are identical, and there is no need to concern ourselves with $\mathcal{L}_{\omega_1,\omega}$.)

**Definition 3.3.2.** $\mathcal{X}$ is the space of all $\mathcal{L}$-structures $\mathcal{M}'$ with universe $M$ such that $\text{age}(\mathcal{M}') \subseteq K$, and $K[M]$ is the set of all $A \in K$ whose universes are subsets of $M$. As usual, if $A \in K[M]$, then $[A] = \{\mathcal{M}' \in \mathcal{X} : A < M', \}$, and $\mathcal{X}$ is the Stone space with $\{[A] : A \in K[M]\}$ as a sub-base of clopen sets. $\mathcal{B}$ is the boolean algebra of subsets of $\mathcal{X}$ so that $\mathcal{X} = S(\mathcal{B})$ up to homeomorphism.) Obviously, $\text{Sym}(M)$ acts on $\mathcal{X}$ by homeomorphisms and compatibly on $\mathcal{B}$ by automorphisms.

**An invariant measure on $\mathcal{X}$** is a Borel probability measure $\mu : \text{Bor}(\mathcal{X}) \to [0, 1]$ such that $\mu(\sigma X) = \mu(X)$ whenever $X \in \text{Bor}(\mathcal{X})$ and $\sigma \in \text{Sym}(M)$. For $\mathcal{M}' \in \mathcal{X}$, we identify $\mathcal{M}/\cong$, the isomorphism type of $\mathcal{M}'$, with its $\text{Sym}(M)$-orbit, and we say that $\mu$ is concentrated on $\mathcal{M}'$ if $\mu(\mathcal{M}'/\cong) = 1$.

**Remark 3.3.3.** By the Hahn-Kolmogorov Theorem again, invariant measures and APMs are essentially the same thing. Clearly, if $\mu : \text{Bor}(\mathcal{X}) \to [0, 1]$ is an invariant measure, then $m = (m_s)_s$ is an APM for $K$, where $m_s(A) = \mu([A])$ for all $s \subset \text{fin} M$ and $A \in K_s$. Conversely, suppose $m$ is an APM for $K$. Then we define a “pre-measure” $\mu_0 : \mathcal{B} \to [0, 1]$ by setting $\mu_0([A]) = m_s(A)$ whenever $s \subset \text{fin} M$ and $A \in K_s$, and extending to enforce finite-additivity in the natural way. Then we recover a $\sigma$-additive measure $\mu : \text{Bor}(\mathcal{X}) \to [0, 1]$ extending $\mu_0$.

In the proof of Theorem 4.1.1, we will need to address invariant measures on some other spaces besides $\mathcal{X}$. We introduce those spaces (and the meaning of “invariance” for each) in the next to definitions.

**Definition 3.3.4.** Fix $p(\bar{x}) \in S_k(T_K)$ such that $p(\bar{x}) \models \bigwedge_{i<j} x_i \neq x_j$. $\mathcal{X}^p$ is the set of all $\mathcal{L}^p$-structures $\mathcal{N}$ with universe $M$ such that $\text{Age}(\mathcal{N}) \subseteq K^p$. $K^p[M]$ is the set of all $A \in K^p$ whose universes are subsets of $M$. As usual, if $A \in K^p[M]$, then $[A] = \{\mathcal{N} \in \mathcal{X}^p : A < \mathcal{N}\}$, and $\mathcal{X}^p$ is the Stone space with $\{[A] : A \in K^p[M]\}$ as a sub-base of clopen sets. $\mathcal{B}^p$ is the boolean algebra associated with $\mathcal{X}^p$ – i.e. the boolean algebra generated by $\{[A] : A \in K^p[M]\}$. Invariant measures on $\mathcal{X}^p$ are defined as expected.

We note that for $C \in K[M]$, the set $\{\mathcal{N} \in \mathcal{X}^p : C \leq \mathcal{N} \downarrow \mathcal{L}\}$ is also a clopen set; we denote this set by $[C]$ and just hope that this will not cause any confusion.

---

4$\mu_0$ is a pre-measure" means that if $X_0, \ldots, X_n, \ldots$ are pairwise disjoint sets in $\mathcal{B}$ such that $\bigcup_n X_n$ is also in $\mathcal{B}$, then $\kappa(\bigcup_n X_n) = \sum_n \kappa(X_n)$. \]

---

13
Not all invariant measures are created equal. In particular, the ergodic invariant measures have a particularly close relationship with the asymptotics of finite structures, and we will make essential use of this fact later on. For now, we just introduce the definition of “ergodic” in this context and some of its equivalents. The theorem expressing the equivalence is given here, but with the readers indulgence, we defer discussion of convergent sequences of finite structures to Subsection 4.2, when we actually use this formulation.

We remark that “dissociated” invariant measures constitute a weaker variant of APMs with independent sampling. In general, this appears to be a strict weakening, but we will see (Theorem 4.1.3) that given 3-DAP all dissociated invariant measures concentrated on $T_K$ have independent sampling.

Definition 3.3.5. Let $\mu : \text{Bor}(X) \rightarrow [0,1]$ be an invariant measure.

- We say that $\mu$ is ergodic if for every Borel set $X \subseteq X$, if $\mu(X \triangle gX) = 0$ for all $g \in \text{Sym}(M)$, then either $\mu(X) = 0$ or $\mu(X) = 1$.

- We say that $\mu$ is dissociated if for any finite structures $A, B \in K[M]$, if $A \cap B = \emptyset$ as sets, then $\mu([A] \cap [B]) = \mu([A])\mu([B])$.

Theorem 3.3.6 (see [12], [1] and/or [16]). Let $\mu : \text{Bor}(X) \rightarrow [0,1]$ be an invariant measure. The following are equivalent:

1. $\mu$ is ergodic.
2. $\mu$ is dissociated.
3. $\mu$ is an extreme point of $\mathcal{C}$, where $\mathcal{C}$ is the compact convex set of all invariant measures living inside the Banach space $\mathcal{B}$ of finite signed measures $\text{Bor}(X) \rightarrow \mathbb{R}$ under the total variation norm.
4. $\mu$ is induced by a convergent sequence of finite structures.

3.4 Measuring systems

Definition 3.4.1. Consider a family $\underline{\mu} = (\mu^p : p \in S^*(T_K))$ as follows:

MSys-1. For each $p$, $\mu^p : \text{Bor}(X^p) \rightarrow [0,1]$ is an invariant measure concentrated on $T_{Kp}$.

MSys-2. Let $p(x_0, ..., x_{k-1}) \in S^*(T_K)$, and let $\sigma \in \text{Sym}(k)$; also, let $p_\sigma$ be the type such that $p_\sigma(x_{\sigma(0)}, ..., x_{\sigma(k-1)}) = p(\pi)$. Let $C \subseteq \text{fin} M, q(\pi) \in S_k^*(C)$ an extension of $p$, and $q_\sigma(\pi)$ such that $q_\sigma(x_{\sigma(0)}, ..., x_{\sigma(k-1)}) = q(\pi)$. Then

$$\mu^p([C+q]) = \mu^{p_\sigma}([C+q_\sigma]).$$

MSys-3. For all $p$, all $A \in K^p$ and $A^- = A \upharpoonright \mathcal{L}$, and all $A^- \leq B \leq C$,

$$\mu^p([A] \upharpoonright [B]) = \mu^p([A] \upharpoonright [C]).$$

MSys-4. Let $C \subseteq \text{fin} M$, and let $\tilde{a}, \tilde{b} \in M^{<\omega}$ be non-repeating and both disjoint from $C$ and from each other. Let

$$p = tp^M(\tilde{a}, \tilde{b}), \quad p_a = tp^M(\tilde{a}), \quad p_b = tp^M(\tilde{b}),$$
\[ q(\bar{x}, \bar{y}) = \text{tp}^M(\bar{a}, \bar{b}/C), \quad q_b(\bar{y}) = \text{tp}^M(\bar{b}/C), \quad q_{a|b}(\bar{x}) = \text{tp}^M(\bar{a}/\bar{b}C). \]

Then identifying \( C \) and \( C\bar{b} \) with \( \mathcal{M}|C \) and \( \mathcal{M}|C\bar{b} \), respectively,

\[
\mu^P([C+q] \mid [C]) = \mu^P_b([C+q_b] \mid [C]) \cdot \mu^P_a([C\bar{b}+q_{a|b}] \mid [C\bar{b}])
\]

Then \( \mu \) is called a measuring system for \( K \).

**Proposition 3.4.2.** Suppose \( \mu \) is a measuring system for \( K \). Define \( m^\mu : S^*[M] \to [0, 1] \) by setting

\[
m^\mu(\bar{a}/C) = \mu^P([C+q] \mid [C])
\]

whenever \( C \subseteq \text{fin} M, \bar{a} \in M^{<\omega} \) is non-repeating and disjoint from \( C \), \( p(\bar{x}) = \text{tp}^M(\bar{a}) \), and \( q(\bar{x}) = \text{tp}^M(\bar{a}/C) \). Then \( m^\mu \) extends uniquely to a probabilistic MS-measure.

**Proof.** We must first extend \( m^\mu \) to a map \( S[M] \to [0, 1] \). If \( \bar{a} \subseteq C \), then we set \( m^\mu(\bar{a}/C) = 1 \) by fiat. If \( \bar{a} \not\subseteq C \), then:

We choose a complete subset \( J \subseteq I(\bar{a}/C) \) of \( E(\bar{a}/C) \)-representatives, and we set \( \bar{a}' = (a_{j_0}, \ldots, a_{j_{k-1}}) \), where \( J = \{j_0 < \cdots < j_{k-1}\} \). Then \( m^\mu(\bar{a}/C) = m^\mu(\bar{a}'/C) \).

Then MS-1, MS-2, and MS-3 for \( m^\mu \) are clear. Further, MS-4 follows immediately from the fact that each \( \mu^P \) is a probability measure. MS-5, part (a), follows from MSys-2, and MS-5, part (b), follows from MSys-3 and MSys-4. The fact that \( m^\mu \) is probabilistic also follows from the fact that each \( \mu^P \) is a probability measure. \( \square \)
4 Construction of measuring systems

4.1 Proof of the main theorem (4.1.1)

In this and the following two subsections, we complete the proof of the main theorem of this paper, which we now restate for the reader’s convenience:

**Theorem 4.1.1.** Let $K$ be a combinatorial class. The following are equivalent:

1. $T_K$ is super-simple of $SU$-rank 1.
2. There is an asymptotic probability measure $a$ for $K$ with independent sampling whose almost-sure theory $T^a$ is equal to $T_K$.

As shown in Proposition 3.4.2, the existence of a measuring system $\mu$ for $K$ would yield a probabilistic $MS$-measure, which in turn would yield $1 \Rightarrow 2$ of Theorem 4.1.1. Thus, we need to devise a mechanism to produce such measuring systems, and the next theorem does just this.

**Theorem 4.1.2.** Assume $K$ has 3-DAP. Suppose that $\mu : \text{Bor}(X) \to [0,1]$ is some invariant measure concentrated on $T_K$. If $\mu$ is ergodic, then it induces a measuring system for $K$.

We now present the proof of Theorem 4.1.1 using Theorem 4.1.2. The proof of the latter requires a bit more understanding of ergodic invariant measures, which occupies us in the following two subsections.

**Proof of Theorem 4.1.1 (using Theorem 4.1.2).** $2 \Rightarrow 1$ is precisely Theorem 1.1.1, so we just need to prove $1 \Rightarrow 2$. So, assume $T_K$ is super-simple of $SU$-rank 1. To apply Theorem 4.1.2, we really just need to verify that there are ergodic invariant measures concentrated on $T_K$.

**Claim.** There is an ergodic invariant measure $\mu : \text{Bor}(X) \to [0,1]$ concentrated on $T_K$ (i.e. on $M/\cong$).

**Proof of claim.** We use the notations, $B$ and $C$, of Theorem 3.3.6. Further, let $C_1$ denote the set of all invariant measures that are concentrated on $M/\cong$. Once again, $C_1$ is a compact convex subset of $B$, so it is equal to the closure of the set of convex combinations of its own extreme points. Moreover, $C_1$ is non-empty because $T_K$ is algebraically trivial (using Theorem 3.3.1, which comes from [2]). Let $\mu$ be an extreme point of $C_1$; we claim that $\mu$ is also an extreme point of $C$, hence ergodic by Theorem 3.3.6.

Let $Z$ be the complement of $M/\cong$ in $X$. Towards a contradiction, suppose there are invariant measures $\nu_1, \nu_2 : \text{Bor}(X) \to [0,1]$ and a number $0 < \alpha < 1$ such that $\mu = \alpha \nu_1 + (1-\alpha) \nu_2$. Since $\mu(Z) = 0$, $\alpha > 0$, and $1-\alpha > 0$, we find that $\nu_1(Z) = \nu_2(Z) = 0$, so $\nu_1(M/\cong) = \nu_2(M/\cong) = 1$, and so $\nu_1, \nu_2 \in C_1$. This shows that $\mu$ is not an extreme point of $C_1$ – a contradiction.

Now, let $\mu : \text{Bor}(X) \to [0,1]$ be an ergodic invariant measure concentrated on $T_K$, and let $q \in S^*_s(M)$ be generic. By Theorem 4.1.2, we recover a measuring system $\mu = (\mu^p : p \in S^*(T_K))$ for $K$. By Proposition 3.4.2, we then recover from $\mu$ a probabilistic $MS$-measure $m = m^\mu$. Finally, by Proposition 3.2.6, we recover $a = a^m$, a positive APM with independent sampling, and by Theorem $\#\#\#$, $T^a = T_K$.

We note that this proof of Theorem 4.1.1 actually yields a somewhat stronger result, partially characterizing all ergodic invariant measures concentrated on the generic model of $K$. 

16
Theorem 4.1.3. Assume $T_K$ is super-simple of SU-rank $1$ (i.e. $K$ has 3-DAP). Then there are ergodic invariant measures concentrated on $T_K$, and for any APM $a$ for $K$, if $\nu^a$ is ergodic and concentrated on $T_K$, then $a$ has independent sampling and $T^a = T_K$.

4.2 Ergodicity and convergent sequences of finite structures

In Subsection 3.3, we presented Theorem 3.3.6 characterizing ergodic invariant measures. We deferred discussion of convergent sequences of finite structures, but we remedy this omission now. As we shall see, the essential fact about ergodic invariant measures is that they are recoverable from a process of sampling from finite structures.

Definition 4.2.1 (Convergent sequences of structures). For sets finite sets $s$, $s'$, we define $\text{Inj}(s, s')$ to be the set of injections $s \to s'$ and $\text{inj}(s, s') = |\text{Inj}(s, s')|$. If $A \in K_s$ and $B \in K_{s'}$, we define $t(A, B) := \frac{\text{emb}(A, B)}{\text{inj}(s, s')}$, where $\text{emb}(A, B) = |\text{Emb}(A, B)|$ and $\text{Emb}(A, B)$ is the set of embeddings $A \to B$.

A sequence of structures $\sigma = (B_n)_n$ is convergent if for any $s \subseteq \text{fin} M$ and every $A \in K_s$, the sequence $(t(A, B_n))_n$ converges. One easily verifies that if $\sigma$ is convergent, then it induces an APM $a^\sigma$ by setting $a^\sigma(A) = \lim_{n \to \infty} t(A, B_n)$ for every $s \subseteq \text{fin} M$ and $A \in K_s$. The APM $a^\sigma$ then induces an invariant measure $\nu^\sigma : \text{Bor}(\mathcal{X}) \to [0, 1]$. It is also easy to verify that $\nu^\sigma$ is dissociated, hence ergodic by Theorem 3.3.6.

Suppose $s$ is a finite set, $A$ is finite structure, and $u : s \to A$ is an injection; then define $u^{-1}A$ to be the unique structure with universe $s$ such that $u$ is an embedding $u^{-1}A \to A$. For a quantifier-free sentence $\varphi$ with parameters from $s$, we sometimes write $A \models_u \varphi$ to mean that $u^{-1}A \models \varphi$.

At last, we are in position to prove Theorem 4.1.2, completing the proof of Theorem 4.1.1.

4.3 Proof of the Theorem 4.1.2

In all of this subsection, we assume that $T_K$ is super-simple of SU-rank $1$. This fact is used to ensure that for each $p \in S^*(T_K)$, $K_p$ is a combinatorial class, so it is actually possible to concentrate probability measures on the models of $T_{K_p}$.

Let $\mu : \text{Bor}(\mathcal{X}) \to [0, 1]$ be an invariant measure concentrated on $T_K$. We assume that $\mu$ is ergodic, so it is induced by a convergent sequence of structures $\sigma = (B_n)_n$ i.e. $\mu = \nu^\sigma$. We also set $a = a^\sigma$.

For $p(x_0, \ldots, x_{k-1}) \in S^*(T_K)$, we fix distinct new elements $a_0, \ldots, a_{k-1}$ outside of $M$, and for any $s \subseteq \text{fin} M$, we define $s^p = s \cup \{a_0, \ldots, a_{k-1}\}$. If $A \in K_s$ and $p(x_0, \ldots, x_{k-1}) \in S^*(A)$ is an extension of $p$, then $A^*p$ is the unique structure with universe $s^p$ such that $(A^*p)|s = A$ and $A^*p \models p(a_0, \ldots, a_{k-1})$.

Definition 4.3.1. Let $p(x_0, \ldots, x_{k-1}) \in S^*(T_K)$. Let $s_0 \subseteq s \subseteq s' \subseteq \text{fin} M$, $A \in K_s$, and $A_0 = A|s_0$, and let $p \in S^*(A_0)$ be an extension of $p$. Let $C \in K_s$. We then define

$$\xi_a(p, A; C; s', u) = \frac{\left| \{ \bar{v} \in (s' \setminus s)^{k} : C \models_u p(\bar{v}) \} \right|}{|s'|^k}$$

for each $u \in \text{Inj}(s', C)$ such that $u|s \in \text{Emb}(A, C)$. 

17
Lemma 4.3.3. Let \( k < k_1 < \cdots < k_n < \cdots \) be a sequence of positive integers, and for each \( n \), suppose \( Z_{n,0}, \ldots, Z_{n,k_n-1} \) is an i.i.d. sequence of Bernoulli random variables. Suppose \( \lim_{n \to \infty} \mathbb{P}(Z_{n,0} = 1) = p \). Then for every \( \varepsilon > 0 \),

\[
\lim_{k \to \infty} \mathbb{P} \left( \left| \frac{1}{k} \sum_{i<k} Z_i - p \right| \geq \varepsilon \right) = 0.
\]

It is easy to generalize the classical statement slightly as follows: Let \( k_0 < k_1 < \cdots < k_n < \cdots \) be a sequence of positive integers, and for each \( n \), suppose \( Z_{n,0}, \ldots, Z_{n,k_n-1} \) is an i.i.d. sequence of Bernoulli random variables. Suppose \( \lim_{n \to \infty} \mathbb{P}(Z_{n,0} = 1) = p \). Then for every \( \varepsilon > 0 \),

\[
\lim_{n \to \infty} \mathbb{P} \left( \left| \frac{1}{k_n} \sum_{i<k_n} Z_{n,i} - p \right| \geq \varepsilon \right) = 0.
\]

Lemma 4.3.3. There is a function \( F_a : \omega \times \omega \times (0,1) \to \omega \) such that for any \( k, m < \omega \) and \( \delta > 0 \), for any \( s_0 \subseteq s \subseteq s' \subseteq \text{fin} M, A \in K_s A_0 = A \upharpoonright s_0 \), and \( p(\overline{x}) \in S^*(A_0) \) extending \( p(x_0, \ldots, x_{k-1}) \in S^*(T_K) \), if \( |s| \leq m \) and \( |s'| \geq F_a(k, m, \delta) \), then

\[
\lim_{N \to \infty} \mathbb{P} \left( \left| \xi_a(p, A; B_N; s', u) - \frac{a_{s_0}(A_0 \wedge p(\overline{\gamma}, \overline{\varphi}))}{a_{s_0}(A_0)} \right| \geq \delta \right| u \upharpoonright s \in \text{Emb}(A, B_N) \right) = 0.
\]

Proof. It is enough to prove that for fixed \( \delta > 0 \), \( s_0 \subseteq s \subseteq \text{fin} M, A \in K_s A_0 = A \upharpoonright s_0 \), and \( p(\overline{x}) \in S^*(A_0) \) extending \( p(x_0, \ldots, x_{k-1}) \in S^*(T_K) \), the following holds: For every \( \varepsilon > 0 \), there is some \( n_{\varepsilon} < \omega \) such that if \( |s'| \geq n_{\varepsilon} \), then

\[
\lim_{N \to \infty} \mathbb{P} \left( \left| \xi_a(p, A; B_N; s', u) - \frac{a_{s_0}(A_0 \wedge p(\overline{\gamma}, \overline{\varphi}))}{a_{s_0}(A_0)} \right| \geq \delta \right| u \upharpoonright s \in \text{Emb}(A, B_N) \right) \leq \varepsilon.
\]

Let \( \overline{e}_0, \overline{e}_1, \ldots, \overline{e}_n, \ldots \) be an enumeration of the non-repeating \( k \)-tuples from \( M \setminus s \), and for each \( n \), let \( w_n = s \cup \bigcup_{i<n} \overline{e}_i \). For \( n, N < \omega \), we define random variables

\[
X_{N;n:0}, \ldots, X_{N;n:n-1} : \text{Inj}(w_n, B_N) \to \{0, 1\}
\]

and

\[
Y_{N;n:0}, \ldots, Y_{N;n:n-1} : w_n B_N \to \{0, 1\}
\]

as follows:

\[
X_{N;n;i}(u) = \begin{cases} 
1 & \text{if } B_N \models_u p(\overline{e}_i) \\
0 & \text{otherwise}
\end{cases}
\]

\[
Y_{N;n;i}(u) = \begin{cases} 
1 & \text{if } B_N \models_u p(\overline{e}_i) \\
0 & \text{otherwise}
\end{cases}
\]

Abusing notation, we write \( A \) to stand for the event \( "u \upharpoonright s \in \text{Emb}(A, B_N)" \) and \( A_0 \) to stand for the event \( "u \upharpoonright s \in \text{Emb}(A_0, B_N)" \) in all circumstances. Clearly,

\[
\frac{a_{s_0}(A_0 \wedge p(\overline{\gamma}, \overline{\varphi}))}{m_{s_0}(A_0)} = \lim_{n \to \infty} \left( \lim_{N \to \infty} \mathbb{P}(X_{N;n:0} = 1 | A_0) \right)
\]

18
so to prove the lemma, it suffices to show that for every $\varepsilon > 0$, there is some $n_\varepsilon$ such that if $n \geq n_\varepsilon$, then
\[
\lim_{N \to \infty} \mathbb{P} \left( \left| \frac{1}{n} \sum_{i<n} X_{N:n;i} - \frac{1}{n} \sum_{i<n} Y_{N:n;i} \right| \geq \varepsilon \mid A \right) \leq \varepsilon.
\]
The proof of this fact amounts to a few straightforward observations and an application of the Law of Large Numbers for a sequence of Bernoulli random variables.

**Observation 1:** For each $n$, we have
\[
\lim_{N \to \infty} \mathbb{P} \left( \left| \frac{1}{n} \sum_{i<n} X_{N:n;i} - \frac{1}{n} \sum_{i<n} Y_{N:n;i} \right| \geq \varepsilon \mid A \right) = 0.
\]
and similarly with $A_0$ in place of $A$.

**Observation 2:** For every $\varepsilon > 0$, there is an $n_\varepsilon$ such that if $n \geq n_\varepsilon$, then
\[
\lim_{N \to \infty} \mathbb{P} \left( \left| \frac{1}{n} \sum_{i<n} X_{N:n;i} - \frac{1}{n} \sum_{i<n} Y_{N:n;i} \right| \geq \varepsilon \mid A \right) = 0.
\]

**Observation 3:** For all $n < \omega$, $\lim_{N \to \infty} \mathbb{P}(X_{N;n;0} = 1 \mid A_0) - \mathbb{P}(Y_{N;n;0} = 1 \mid A_0) = 0$.

**Observation 4:** For all $n < \omega$, $\lim_{N \to \infty} \mathbb{P}(Y_{N;n;0} = 1 \mid A) - \mathbb{P}(Y_{N;n;0} = 1 \mid A_0) = 0$.

**Observation 5:** For all $N, n$, the conditioned random variables $Y_{N;n;0}|A, ..., Y_{N;n;0}|A$ are i.i.d. Bernoulli random variables.

Observation 1 is enforced by the fact that a uniformly random function $w_n \to B_N$ is injective when $N$ (hence $|B_N|$) grows while $n$ is held constant, and Observations 2 and 3 follow immediately from Observation 1. Observations 4 and 5 can be verified similarly to Observation 1.

Now, by Observation 5, we may apply the Law of Large Numbers, finding that for every $\varepsilon > 0$, there is an $n_\varepsilon$ such that
\[
n \geq n_\varepsilon \Rightarrow \lim_{N \to \infty} \mathbb{P} \left( \left| \frac{1}{n} \sum_{i<n} Y_{N;n;i} - \mathbb{P}(Y_{N;n;0} = 1 \mid A) \right| \geq \varepsilon \mid A \right) = 0.
\]
Given $\varepsilon > 0$, we apply Observations 2 through 4 and the triangle inequality to find $n'_\varepsilon$ such that if $n \geq n'_\varepsilon$, then
\[
\lim_{N \to \infty} \mathbb{P} \left( \left| \frac{1}{n} \sum_{i<n} X_{N:n;i} - \mathbb{P}(X_{N;n;0} = 1 \mid A) \right| \geq \delta \mid A \right) \leq \varepsilon.
\]
This completes the proof of the lemma.

For a structure $C \in K$ and $s \subseteq M$, consider the following random process for generating structures $A \in K_s^P$ (where $p \in S^*(T_K)$):

1. Draw $v \in \text{Inj}(s^P, C)$ uniformly at random; then put $u = v|s$, $A^- = u^{-1}C$, $B = v^{-1}C$, and $q(x) = \text{qftp}^B(\pi/A^-)$; finally, let $A = A^- + q$. 

19
Next, we define $t^p(-, C) : K^p[M] \to [0, 1]$ by taking $t^p(A, C)$ to be the probability of generating $A$ according to the above process.

**Lemma 4.3.4.** For every $s \subset \mathfrak{m} M$, for each $A \in K^p_s$, the limit $\lim_{N \to \infty} t^p(A, B_N)$ converges.

**Proof.** Immediate when we apply Lemma 4.3.3 with $s_0 = s$. \hfill \Box

Now, for each $p \in S^*(T_K)$, we may define an APM $a^p$ for $K^p$, defining $a^p_s : K^p_s \to [0, 1]$ ($s \subset \mathfrak{m} M$) by setting $a^p_s(A) = \lim_{N \to \infty} t^p(A, B_N)$ for each $A \in K^p_s$. We define $\mu^p : \text{Bor}(X^p) \to [0, 1]$ to be the unique $\sigma$-additive measure induced by $a^p$, viewed as a pre-measure. Thus, we have defined a family $\mu = (\mu^p : p \in S^*(T_K))$, and just remains to show that $\mu$ is a measuring system for $K$.

**Lemma 4.3.5.** For each $p \in S^*(T_K)$, $\mu^p$ is concentrated on $T_{K^p}$.

**Proof.** Assume $p = p(x_0, \ldots, x_{k-1})$. It is sufficient to show that for any extension axiom $\psi$ of $K^p$, $\mu^p(\{N \in X^p : N \models \psi\}) = 1$.

Let $\varphi^p = \forall \bar{\tau} (\theta^p_0(\bar{\tau}) \to \exists y \theta^p(\bar{\tau}, y))$ be an extension axiom of $K^p$. So, there are non-repeating $\bar{\tau} \in M_{\bar{\tau}}$ and $b \in M \setminus \bar{\tau}$ such that $\theta^p_0$ isolates $qftp^p(\bar{\tau})$ and $\theta^p$ isolates $qftp^p(\bar{\tau}, b)$ modulo the universal sub-theory of $T_{K^p}$. Let $A^p = M|\bar{\tau}$ and $B^p = M|\bar{\tau}b$, and let $\theta_0(\bar{\tau}), \theta(\bar{\tau}, y)$ be quantifier-free formulas that isolate $qftp^p(\bar{\tau})$, $qftp^p(\bar{\tau}, b)$ modulo the universal sub-theory of $T_K$, respectively. Thus, $\varphi = \forall \bar{\tau} (\theta_0(\bar{\tau}) \to \exists y \theta(\bar{\tau}, y))$ is an extension axiom of $T_K$. Also, let $q_0(\bar{\tau}), q(\bar{\tau}, b)$ be the types over $A = B^p|\mathcal{L}$ and $B = B^p|\mathcal{L}$, respectively, such that $A^p = A + q_0$ and $B^p = B + q$. Since $\mu$ is concentrated on $T_K$, we have $\mu(\{M^p \in X^p : M^p \models \varphi\}) = 1$.

Let $s \subset \mathfrak{m} M$ be the universe of $A$. Let $e_0, \ldots, e_{k-1} \in M \setminus s$ and $s' = s \cup \bar{\tau}$. Then, since $\mu(\{M^p : M^p \models \varphi\}) = 1$, there is a function $f : (0, 1) \to \omega$ such that for any $\varepsilon > 0$, for any $r \subset \mathfrak{m} M$, if $s' \subseteq r$ and $|r| \geq f(\varepsilon)$, then

$$\lim_{N \to \infty} \text{inj}(r, B_N)^{-1} \cdot \left\{ u \in \text{Inj}(r, B_N) : \begin{array}{l} B_N \models q_0(\bar{\tau}), \\ \bigwedge_{c \in r} B_N \not\models_u q(\bar{\tau}, c) \end{array} \right\} \leq \varepsilon.$$ 

Exchanging $\bar{\tau}$ for $\bar{\tau}$, we see that for any $\varepsilon > 0$, for any $r \subset \mathfrak{m} M$, if $s \subseteq r$ and $|r| \geq f(\varepsilon)$, then

$$\lim_{N \to \infty} \text{inj}(r^p, B_N)^{-1} \cdot \left\{ u \in \text{Inj}(r^p, B_N) : \begin{array}{l} B_N \models p_0(\bar{\tau}), \\ \bigwedge_{c \in r} B_N \not\models_u p(\bar{\tau}, c) \end{array} \right\} \leq \varepsilon.$$ 

From this, it follows that $\mu(\{N \in X^p : N \models \varphi^p\}) = 1$ as required. \hfill \Box

**Proposition 4.3.6.** $\mu$ is a measuring system for $K$.

**Proof.** There are four items to verify in the definition of a measuring system (Definition 3.4.1):

MSys-1: By Lemmas 4.3.4 and 4.3.5, for each $p \in S^*(T_K)$, $\mu^p$ is an invariant measure concentrated on $T_{K^p}$.

MSys-2: Let $p(x_0, \ldots, x_{k-1}) \in S^*(T_K)$, and let $\sigma \in \text{Sym}(k)$; also, let $p_{\sigma}$ be the type such that $p_{\sigma}(x_{\sigma(0)}, \ldots, x_{\sigma(k-1)}) = p(\bar{\tau})$. Let $s \subset \mathfrak{m} M$ and $C = M|s$, $q(\bar{\tau}) \in S^*_k(C)$ an extension of $p$, and $q_{\sigma}(\bar{\tau})$ such that $q_{\sigma}(x_{\sigma(0)}, \ldots, x_{\sigma(k-1)}) = q(\bar{\tau})$.

It is clear from the construction of the $a^p$ and $a^{p_{\sigma}}$ (whence of $\mu^p$ and $\mu^{p_{\sigma}}$) that

$$\mu^p([\mathcal{C} + q]) = a^p_\mathcal{C}([\mathcal{C} + q]) = a^{p_{\sigma}}_\mathcal{C}([\mathcal{C} + q_{\sigma}) = \mu^{p_{\sigma}}([\mathcal{C} + q_{\sigma}]).$$

20
MSys-3: Let \( p(x_0, ..., x_{k-1}) \in S^*(T_K), A \in K^p \) and \( A^- = A|L' \), and \( A^- \leq C_1 \leq C_2 \in K[M] \). Let \( s_0 \subseteq s_1 \subseteq s_2 \) be the universes of \( A, C_1, C_2 \), respectively. Also, let \( m = |s_2| \) and \( p(\overline{x}, \overline{a}) \in S_k(A) \) such that \( A = A^- + p \). By Lemma 4.3.3, we have

\[
\lim_{N \to \infty} P\left( \left| \xi_a(p, C_1; B_N; s', u) - \xi_a(p, C_2; B_N; s', u) \right| \geq \delta \mid u|s_2 \in \text{Emb}(C_2, B_N) \right) = 0.
\]

whenever \( |s'| \geq F_a(k, m, \delta) \) (any \( \delta > 0 \)). Then, by definition of \( \rho^p \) and \( a^p \), it follows that

\[
\mu^p([A] \mid [C_1]) = \frac{a^p_{s_1}(A)}{a^p_{s_2}(A)} = \frac{a^p_{s_2}(A)}{a^p_{s_2}(C_2)} = \mu^p([A] \mid [C_2]).
\]

MSys-4: Let \( s \subset \mathcal{M} \), and let \( \overline{a}, \overline{b} \in M^{<\omega} \) be non-repeating and both disjoint from \( C = M|s \) and from each other. Let

\[
p = tp^M(\overline{a}, \overline{b}), \quad p_0 = tp^M(\overline{a}), \quad p_b = tp^M(\overline{b}),
\]

\[q(\overline{x}, \overline{y}) = tp^M(\overline{a}, \overline{b}/C), \quad q_b(\overline{y}) = tp^M(\overline{b}/C), \quad q_0(x, \overline{y}) = tp^M(\overline{a}/\overline{b}C).
\]

Also, let \( s' = s \cup \overline{b} \) and \( \overline{C} = M|s' \). Then, by conditioning in finite probability spaces, we have

\[
\frac{a^p_b(C+q)}{\sum_{C'|L=C} a^p_b(C')} = \frac{a^p_b(C+q_b)}{\sum_{C'|L=C} a^p_b(C')} \cdot \frac{a^p_{s_2}(\overline{C}+q_{0|b})}{a^p_{s_2}(\overline{C}+q_{0|b})} = \frac{a^p_{s_2}(\overline{C}+q_{0|b})}{a^p_{s_2}(\overline{C}+q_{0|b})}
\]

It follows that

\[
\mu^p([C+q] \mid [C']) = \mu^p([C+q_b] \mid [C']) \cdot \mu^p([\overline{C}+q_{0|b}] \mid [\overline{C}]).
\]

\( \Box \)
5 Further questions and results on pseudo-finiteness

5.1 Approximation and the finite sub-model property

In [14], Kruckman uses a certain notion of filtration (or approximation) of a Fraïssé class by sub-
classes to demonstrate that the theories $T^*_{\text{feq}}$ and $T^*_{\text{cpz}}$ are pseudo-finite (in fact, that they have the
finite sub-model property). Those sub-classes, of course, must induce theories that have the finite
sub-model property, and to ensure this, Kruckman requires that (up to an expansion-by-definitions)
these sub-classes have $n$-DAP for all $n$. Because of this requirement, the framework of [14] excludes
the theory $T_{G_{3,4}}$ of the generic tetrahedron-free 3-hypergraph (itself or as an approximation). Using
Theorem 4.1.1, we now prove a generalization, Theorem 5.1.3, of main general theorem of [14].

First, of course, we must formalize the notion of approximation in play.

Definition 5.1.1. Let $\mathcal{M}$ be the generic model of $K$. Let

$$
M_0 \leq M_1 \leq \cdots \leq M_r \leq \cdots
$$

be a chain of (non-elementary) substructures of $\mathcal{M}$ such that $\bigcup_r M_r = \mathcal{M}$ and for each $r$, $T_r :=
Th(M_r)$ is an algebraically trivial $\aleph_0$-categorical theory. For each $r$, let $K^r$ be the isomorphism-
closure of the set of finite induced substructures of $M_r$. In general, $K^r$ need not be a Fraïssé class,
but replacing $K^r$ with the age $\tilde{K}^r$ of the morleyization of $M_r$, we find that $\tilde{K}^r$ is a
combinatorial class. Under these circumstances and provided that $T_K = \bigcap_r T_r$, the family $(K^r)_r$ is now called a filtration of $K$.

By extension, the family $(M_r)_r$ is a filtration of $\mathcal{M}$, and the family $(T_r)_r$ is a filtration of $T_K$.

We will say that $T_K$ is a simply-approximable theory if there is a filtration $(K^r)_r$ of $T_K$ such that
for each $r$, $T_{K^r}$ is super-simple of $SU$-rank 1.

Fact 5.1.2. If $T_K$ is an almost-sure theory, then it has the finite sub-model property.

Theorem 5.1.3. If $T_K$ is simply-approximable, then $T_K$ has finite sub-model property.

Proof. Let $(K^r)_r$, be a filtration of $K$ such that each $T_{K^r}$ is super-simple of $SU$-rank 1. Let
$(M_r)_r$ be the corresponding filtration of the generic model $\mathcal{M}$. By Theorem 4.1.1 and Fact 5.1.2,
for each $r$, $T_{K^r}$ has the finite sub-model property.

Now, let $A_0 \subseteq \text{fin } M$ and $\varphi \in T_K = \bigcap_r T_{K^r}$ be given. We choose $r$ such that $A_0 \subseteq M_r$. Since
$T_{K^r}$ has the finite sub-model property we recover $A \subseteq \text{fin } M_r$ containing $A_0$ such that $M_r \models \varphi$, and since $\varphi \in \mathcal{L}$, it follows that $M_r \models \varphi$. ☐

Theorem 5.1.3 (and the difficulty of concocting pseudo-finite combinatorial theories with some
version of that theorem) suggests a slightly outrageous conjecture, which would amount to a more
or less complete characterization of combinatorial theories with the finite sub-model property.

Conjecture 5.1.4. If $T_K$ has the finite sub-model property, then $T_K$ is simply-approximable.

Given Conjecture 5.1.4 and the central place of simplicity in the pseudo-finiteness phenomenon
that it suggests, we are tempted to approach another question posed by Cherlin:

Conjecture 5.1.5. Let $H$ be the combinatorial class of finite triangle-free graphs, so that $T_H$ is the
theory of the Henson graph. It’s very well-known that $T_H$ is not simple – we conjecture that $T_H$ is
not even simply-approximable.
5.2 Uncountably many pseudo-finite countable $\aleph_0$-categorical structures

In Problem A of [4], Cherlin asks whether, for $\aleph_0$-categorical theories, the pseudo-finiteness phenomenon is very rare or very common. In the following theorem, we resolve this question.

**Theorem 5.2.1.** In a particular finite relational language $\mathcal{L}$, there are $2^{\aleph_0}$ pairwise inequivalent pseudo-finite $\aleph_0$-categorical theories. More precisely, there are $2^{\aleph_0}$ pairwise inequivalent almost-sure combinatorial theories.

It is convenient to narrow the class of theories in question. Hence, we restrict attention to classes of finite structures that very much like regular hypergraphs.

**Definition 5.2.2.** Let $\mathcal{L}$ be a finite relational language. For each $n$-ary relation symbol $R$ of $\mathcal{L}$, let $\varphi_R$ be the sentence,

$$\forall \overline{x} \left( R(\overline{x}) \rightarrow \bigwedge_{i<j} x_i \neq x_j \right) \land \forall \overline{x} \left( R(\overline{x}) \rightarrow \bigwedge_{g \in \text{Sym}(n)} R(x_{g(0)}, \ldots, x_{g(n-1)}) \right).$$

In the terminology of [19], an $\mathcal{L}$-structure is a **society** just in case it is a model of $\{ \varphi_R : R \in \text{sig}(\mathcal{L}) \}$.

We write $S(\mathcal{L})$ for the class of all finite $\mathcal{L}$-societies (which is clearly a combinatorial class); when $\mathcal{L}$ is clear from context, we will write $S$ in place of $S(\mathcal{L})$.

For $k \geq 2$, a finite $\mathcal{L}$-society $A$ is called $k$-irreducible if $|A| \geq 2$ and for any $a_0, \ldots, a_{k-1} \in A$, there are $R \in \text{sig}(\mathcal{L})$ and $\overline{b} \in R^A$ such that $a_i \in \overline{b}$ for each $i < k$. It’s not difficult to see that if $A$ is $(k+1)$-irreducible, then it is $k$-irreducible as well.

For $\mathcal{F} \subset S$, we write $S_{\mathcal{F}}$ to denote the class of structures $B \in S$ such that for no $A \in \mathcal{F}$ is there an injective homomorphism $A \rightarrow B$. A **base** is a set $\mathcal{F} \subset S$ such that (i) every $A \in \mathcal{F}$ is 2-irreducible, and (ii) for all $A, B \in \mathcal{F}$, if $A \preceq B$, then $A \cong B$. (Here, $A \preceq B$ means that there is an injective homomorphism $A \rightarrow B$.)

Under somewhat different terminology and notation, Propositions 5.2.3 and 5.2.4 and Corollary 5.2.5 below are due to [?], where the existence of the Henson graph and similar structures was first demonstrated.

**Proposition 5.2.3.** Fix a finite relational language $\mathcal{L}$, and let $\mathcal{F} \subset S$. If every $A \in \mathcal{F}$ is 2-irreducible, then $S_{\mathcal{F}}$ is a combinatorial class.

**Proposition 5.2.4.** Fix a finite relational language $\mathcal{L}$, and suppose $\mathcal{F}_1, \mathcal{F}_2 \subset S = S(\mathcal{L})$ are bases. The following are equivalent:

1. $\mathcal{F}_1 \subseteq \mathcal{F}_2$ up to isomorphism.
2. $S_{\mathcal{F}_1} \supseteq S_{\mathcal{F}_2}$.
3. There is an embedding $\mathcal{M}_2 \rightarrow \mathcal{M}_1$, where $\mathcal{M}_1, \mathcal{M}_2$ are the generic models of $S_{\mathcal{F}_1}, S_{\mathcal{F}_2}$, respectively.

**Corollary 5.2.5.** Fix a finite relational language $\mathcal{L}$, and suppose $\mathcal{F}_1, \mathcal{F}_2 \subset S = S(\mathcal{L})$ are bases. Then $T_{S_{\mathcal{F}_1}} = T_{S_{\mathcal{F}_2}}$ if and only if $\mathcal{F}_1 = \mathcal{F}_2$ up to isomorphism.

The following theorem of [?] will allow us to generate a plethora of combinatorial classes with super-simple generic theories.
Theorem 5.2.6. Fix a finite relational language $\mathcal{L}$, and let $\mathcal{F} \subset \mathcal{S}$. Assume every $A \in \mathcal{F}$ is 2-irreducible. Then $T_{\mathcal{S},\mathcal{F}}$ is super-simple of $SU$-rank 1 if and only if every $A \in \mathcal{F}$ is 3-irreducible.

Let $\mathcal{L}$ be the relational language show signature consists of the binary relation symbols $R, S_0, S_1$ and 3-ary relation symbols $Q_0, Q_1, H$. Let $\mathcal{V} = \{w \in \{0, 1\}^{\leq \omega} : |w| \geq 1\}$. For $w \in \mathcal{V}$, we define an $\mathcal{L}$-society $A_w$ as follows:

- As a set, $A_w$ consists of $2n + 2$ distinct elements $a_0, ..., a_{n-1}, b_0, ..., b_{n-1}$, and $c, c'$.
- $Q^0_{A_w} = \{\{a_0, b_0, c\}\}$, $Q^1_{A_w} = \{\{a_{n-1}, b_{n-1}, c'\}\}$, and $H^A_w = \binom{A_w}{3}$, the set of all 3-element subsets of $A_w$.
- $R^A_w = \{\{a_i, a_{i+1}\}, \{b_i, b_{i+1}\} : i < n - 1\}$, $S^A_{0w} = \{\{a_i, b_i\} : w_i = 0\}$, and $S^A_{1w} = \{\{a_i, b_i\} : w_i = 1\}$.

Observation 5.2.7. Because of how the symbol $H$ is interpreted, $A_w$ is 3-irreducible for every $w \in \mathcal{V}$.

Lemma 5.2.8. For every subset $W \subseteq \mathcal{V}$, $\mathcal{F}(W) = \{A_w : w \in W\}$ is a base.

Proof of Theorem 5.2.1. For each $W \subseteq \mathcal{V}$, let $T_W := T_{\mathcal{S},\mathcal{F}(W)}$. For distinct $W_1, W_2 \subseteq \mathcal{V}$, $T_{W_1} \neq T_{W_2}$ by Lemma 5.2.8 and Corollary 5.2.5, so $\{T_W : W \subseteq \mathcal{V}\}$ has cardinality $2^{\aleph_0}$. By Theorem 5.2.6 and Observation 5.2.7, each $T_W$ is super-simple of $SU$-rank 1, hence an almost-sure theory by Theorem 4.1.1. Thus, $\{T_W : W \subseteq \mathcal{V}\}$ is a set of continuum many distinct pseudo-finite $\aleph_0$-categorical theories.

\[\square\]
References


[10] Cameron Donnay Hill and Alex Kruckman. Measures, measures, and measures. (in preparation), 201X.


