Abstract

The celebrated Erdős–Hajnal Conjecture says that in any proper hereditary class of finite graphs we are guaranteed to have a clique or anti-clique of size $n^c$, which is a much better bound than the logarithmic size that is provided by Ramsey’s Theorem in general. On the other hand, in uncountable cardinalities, the model-theoretic property of stability guarantees a uniform set much larger than the bound provided by the Erdős–Rado Theorem in general.

Even though the consequences of stability in the finite have been much studied in the literature, the countable setting seems a priori quite different, namely, in the countably infinite the notion of largeness based on cardinality alone does not reveal any structure as Ramsey’s Theorem already provides a countably infinite uniform set in general. In this paper, we show that the natural notion of largeness given by upper density reveals that these phenomena meet in the countable: a countable graph has an almost clique or anti-clique of positive upper density if and only if it has a positive upper density almost stable set. Moreover, this result also extends naturally to countable models of a universal theory in a finite relational language.

Our methods explore a connection with the notion of convergence in the theory of limits of dense combinatorial objects, introducing and studying a natural approximate version of the Erdős–Hajnal property that allows for a negligible error in the edges (in general, predicates) but requires linear-sized uniform sets in convergent sequences of models (this is much stronger than what stable regularity can provide as the error is required to go to zero). Finally, surprisingly, we completely characterize all hereditary classes of finite graphs that have this approximate Erdős–Hajnal property. The proof highlights both differences and similarities with the original conjecture.

*Institute for Advanced Study, lenacore@ias.edu. This material is based upon work supported by a grant from the Institute for Advanced Study School of Mathematics.
†University of Chicago, mem@math.uchicago.edu. Research partially supported by NSF-BSF 2051825.
1 Introduction

The celebrated Ramsey’s Theorem [Ram29] guarantees that sufficiently large structures have uniform substructures. Without any extra restrictions, the size of the guaranteed uniform substructure is typically tiny in comparison to the ambient structure: for example, to guarantee a clique or independent set of size $n$ in a graph, its size must be exponential in $n$. The famous Erdős–Hajnal Conjecture [EH89] then asks if this bound can be improved to polynomial in $n$ if we restrict the problem to any hereditary (i.e., closed under induced subgraphs) proper subclass of finite graphs. Several hereditary proper subclasses of finite graphs are known to satisfy the Erdős–Hajnal Conjecture (see [Chu14] for a survey).

In the uncountable, Ramsey type theorems also detect important differences between structures. This is best stated in the language of logic, specifically set theory and model theory: in general, the Erdős–Rado Theorem [ER56] gives a tower bound on the size of a uniform subset, however, in structures which are stable in the sense of model theory, see below, we can extract uniform subsets of essentially the same size as the model.

However, in the case of countable structures balanced between these two, all structures seem to behave in the same way since the infinite version of Ramsey’s Theorem yields a uniform set of the maximum possible cardinality of $\aleph_0$, so it is natural to ask if there is any version of uniformity that would be able to distinguish between countable structures.

We propose an answer through probability, more specifically through the language of graph limits and continuous combinatorics (see [Lov12] for the graph case and [Aus08, AC14, CR20b] for the case of universal theories on finite relational languages). When considering universal theories of graphs, we work with graphons, which are continuum-sized limits of convergent sequences of finite graphs, in which relative sizes of sets of vertices are encoded by a probability measure. We start by proving a dichotomy theorem for graphons (Theorem 3.6): a graphon contains a positive measure almost clique or a positive measure almost independent set if and only if it has an almost stable positive measure subgraphon (see Definitions 3.2 and 3.5). We also give examples to illustrate how the negative side works, preventing any positive measure uniform sets for basic instances of instability (quasirandom graphs and “recursive” half-graphs). We also generalize this theorem for arbitrary universal theories $T$ in finite relational languages (Theorem 5.11): a $T$-on (i.e., a limit of a convergent sequence of finite models of $T$) has a pos-
itive measure “uniform” sub-object if and only if it has a positive measure sub-object in which all predicate symbols are stable (see Definition 5.10).

Since graphons and theons are limits of convergent sequences, the dichotomy can be pushed down to convergent sequences (Theorems 4.3) as: a convergent sequence of finite graphs \((H_n)_{n \in \mathbb{N}}\) has non-negligible (i.e., linear-sized) sets \(U_n \subseteq V(H_n)\) such that the edge density in \((H_n|_{U_n})_{n \in \mathbb{N}}\) either converges to 1 or to 0, (i.e., it is an almost clique or an almost independent set) if and only if there are non-negligible sets \(U'_n \subseteq V(H_n)\) that make the edge relation almost stable in the induced sequence \((H_n|_{U'_n})_{n \in \mathbb{N}}\). A similar theorem is also obtained for arbitrary universal theories (Theorem 6.3).

For a countable graph (or model) \(G\), this dichotomy (Theorems 4.7 and 6.4) takes the following form: there exists a set \(U \subseteq \mathbb{N}_+\) and an increasing sequence \((n_\ell)_{\ell \in \mathbb{N}}\) of positive integers along which \((G|_{U \cap [n_\ell]})_{\ell \in \mathbb{N}}\) is almost “uniform” and \(U \cap [n_\ell]\) is non-negligible in \([n_\ell]\) if and only if there exists a set \(U' \subseteq \mathbb{N}\) and an increasing sequence \((n'_\ell)_{\ell \in \mathbb{N}}\) of positive integers along which \((G|_{U' \cap [n'_\ell]})_{\ell \in \mathbb{N}}\) is almost stable and \(U' \cap [n'_\ell]\) is non-negligible in \([n'_\ell]\).

From a general point of view, the theorems above on convergent sequences may be understood as characterizing existence of a linear-sized subset which is an almost clique or almost empty graph. This characterization is in terms of almost-absence of a certain finite structure (but notice that we remain agnostic about some of the edges; this is in some ways very satisfying as it coincides with the major structural property mentioned above, stability, in model theory). Focusing on the case of graphs, by analogy to the usual Erdős–Hajnal Conjecture, this begs the question: does there exist a family of finite graphs (with no agnostic edges) whose absence precisely characterizes the existence of a linear-sized almost clique or almost anti-clique in a convergent sequence?

In the last, substantial section of the paper, Section 8, we prove that the answer is yes and provide a complete characterization (Theorem 8.10) of this family as all induced subgraphs of some recursive blow-up of the 4-cycle (see Definition 8.1 and Figure 6). This doubly unexpected characterization (its existence and the nature of the forbidden family were both a surprise) sheds a quite different light on the proofs above.

Using the characterization above, we can show (Theorem 8.11) that this approximate Erdős–Hajnal property (AEHP) implies the usual Erdős–Hajnal property (EHP). Namely, if a hereditary class of graphs \(\mathcal{G}\) is such that every convergent sequence of graphs in \(\mathcal{G}\) has a linear-sized almost clique or almost anti-clique, then there exists a constant \(c_\mathcal{G} > 0\) such that every graph of \(\mathcal{G}\) on
$n$ vertices has a clique or anti-clique of size $n^{c_2}$. Our proof is somewhat indirect and crucially relies on the characterization of AEHP and in the concluding section, we ask if a more direct construction is possible.

These points are further explored in the text once the details of the proofs are available for discussion and commentary.

We conclude the introduction with some comments on the model theoretic notion of stability, which has been mentioned several times above. We emphasize that it is not necessary to be familiar with stability either to read or to appreciate the present paper, but let us explain this remark. It was very interesting to us to discover in the course of writing this paper that stability appears in a characteristic sense in some of our main results, despite investigating a priori unrelated questions. Stability is not only a key concept in Shelah’s classification theory [She90] but has had recent applications in finite combinatorics via the stable regularity lemma and stable Ramsey’s Theorem [MS14] (see also [AFP18, MS21]) and learning theory via Littlestone dimension [ALMM19, BLM20]. Stability has several equivalent definitions in the language of model theory. The one we will use is combinatorial and says essentially that no relation forms a half-graph with respect to any partition of the variables (see Definition 5.10 below), which may be thought of as ensuring a certain kind of symmetry for relations. For model theorists, we note that our framework only requires working with stability of specific formulas, not of all formulas, and we will note in the text where ideas from stable regularity will play a role in guiding certain proofs. This said, because of the new context, the present proofs require working by hand with the combinatorial definitions and building up everything necessary from scratch. To the extent that stability appears, it is as a characterization: its appearance is fully justified, so to speak, by the proofs in either direction. Moreover, the interaction of convergence and stability in the present context seems to point to something interesting and new about stability’s effect, which was certainly not explored in the usual infinitary context. As a result, not only is the paper self-contained in this respect, but also readers encountering stability for the first time in the present context may bring new understanding by taking these results as a starting point.

The paper is organized as follows. In Section 2 we establish some basic notation that will be used throughout the text. The remainder of the paper can be thematically divided into three main parts.

- In the first part, Sections 3 and 4, we study the interaction of large
almost cliques or almost anti-cliques with large almost stable sets. More specifically, in Section 3, we prove the stability dichotomy theorem for graphons, Theorem 3.6. In Section 4, we prove the finite and countable versions of the stability dichotomy theorem for graphs, Theorems 4.3 and 4.7, respectively. Sections 3 and 4 also contain a gentle introduction to the concepts of the theory of graphons needed for the results.

• In the second part, Sections 5 and 6, we extend the stability results of the first part to universal theories in finite relational languages: in Section 5 we prove the stability dichotomy theorem for arbitrary universal theories, Theorem 5.11 and in Section 6 we prove its finite and countable versions, Theorems 6.3 and 6.4, respectively. These sections also contain a gentle introduction to the concepts of the theory of theons needed for the results.

• In the third part, we revisit the characterization of the first part, in the pure language of hereditary classes. More specifically, in Section 7 we define and prove basic properties of the approximate Erdős–Hajnal property (AEHP), an analogue of the usual Erdős–Hajnal property requiring linear-sized almost uniform sets in convergent sequences of models (rather than polynomial-sized uniform sets in a single model). In Section 8 we characterize universal theories of graphs with AEHP as precisely the theories that forbid some induced subgraph of some recursive blow-up of the 4-cycle (Theorem 8.10) and show that the approximate Erdős–Hajnal property implies the usual Erdős–Hajnal property for graphs (Theorem 8.11).

Although the parts can probably be read in any order (provided the readers are willing to assume the referenced results from previous sections), reading in the given order may provide the most insight into the interconnected aspects of the emerging theory.

We conclude in Section 9 with remarks and open questions.

2 Some notation

Throughout the text, we will use the notation \( \mathbb{N} \overset{\text{def}}{=} \{0, 1, \ldots\} \) for the non-negative integers and \( \mathbb{N}_+ \overset{\text{def}}{=} \mathbb{N} \setminus \{0\} \) for the positive integers. We also let \( [n] \overset{\text{def}}{=} \{1, \ldots, n\} \) and \( (n)_m \overset{\text{def}}{=} n(n-1) \cdots (n-m+1) \). The usage of the
arrow $\hookrightarrow$ for a function will always presume the function to be injective. We let $2^V \overset{\text{def}}{=} \{ A \subseteq V \}$ be the set of all the subsets of $V$, let $(V^\ell) \overset{\text{def}}{=} \{ A \subseteq V \mid |A| = \ell \}$. We also let $r(V)$ be the set of all finite non-empty subsets of $V$ and $r(V, \ell) \overset{\text{def}}{=} \{ A \in r(V) \mid |A| \leq \ell \}$ be the set of all non-empty subsets of $V$ of size at most $\ell$. Given a function $\alpha: V \to W$, we may use the notation $\alpha_v$ for $\alpha(v)$ when convenient.

Given an injection $\alpha: U \hookrightarrow V$ and a set $X$, we let $\alpha^*: X^V \to X^U$ be the contra-variantly defined “projection” given by $\alpha^*(x)_u = x_{\alpha(u)}$. We will be using these “projections” both in the situation where we are interested in the coordinates of some point $x \in X^V$ that are indexed by elements of $\text{im}(\alpha)$ and since $\alpha: U \hookrightarrow V$ induces an injection $\alpha: r(U) \hookrightarrow r(V)$ (denoted by abuse with the same letter), this in turn gives the projection $\alpha^*: X^{r(V)} \to X^{r(U)}$ that allows us to inspect coordinates of $x \in X^{r(V)}$ that are indexed by non-empty subsets of $\text{im}(\alpha) \subseteq V$.

We will be frequently abusing notation by identifying $[n]$ with $n$, e.g., we will use $r(n, \ell)$ as a shorthand for $r([n], \ell)$. Random variables will always be typed in \textbf{math bold face}. We denote by $S_V$ the group of bijections $V \to V$ so that $S_n$ is the group of permutations on $n$ elements.

We denote the complete graph on $n$ vertices by $K_n$ and the empty graph on $n$ vertices by $\overline{K}_n$. We use the terms “anti-clique”, “empty graph” and “independent set” interchangeably.

\section{Almost cliques or anti-cliques in graphons}

In this section, we state and prove the stability dichotomy theorem for graphons, Theorem 3.6. Along the way, we give a gentle introduction to the concepts of graphon theory that we will be using (we refer the reader to [Lov12] for a more thorough introduction to the theory). Let us also remark that some of the techniques of limit theory can be traced back to way before the development of graphons at least as far as [DF81].

Given finite graphs $G$ and $H$, let $T_{\text{ind}}(G, H)$ be the set of all graph embeddings of $G$ in $H$ (i.e., injective functions $f: V(G) \hookrightarrow V(H)$ that preserve edges and non-edges) and let

$$t_{\text{ind}}(G, H) \overset{\text{def}}{=} \frac{|T_{\text{ind}}(G, H)|}{(|H|)|G|}$$


be the normalized number of embeddings of $G$ and $H$; this is sometimes called the labeled (induced) density of $G$ in $H$. The “labeled” here is to differentiate from the (induced) density of $G$ in $H$, which is the normalized number of induced subgraphs of $H$ that are isomorphic to $G$ given by

$$p(G, H) \overset{\text{def}}{=} \frac{|\{ U \subseteq V(H) \mid G|_U \cong H\}|}{|H| |G|} = \frac{|G|!}{|\text{Aut}(G)|} \cdot t_{\text{ind}}(G, H),$$

where $\text{Aut}(G)$ is the group of automorphisms of $G$. We denote by $\rho$ the edge graph, so that $p(\rho, H) = t_{\text{ind}}(\rho, H)$ denotes the edge density of $H$.

A sequence $(H_n)_{n \in \mathbb{N}}$ of finite graphs is called convergent if it is increasing in the sense that for every $n \in \mathbb{N}, |H_n| < |H_{n+1}|$ and if for every finite graph $G$, the limit $\lim_{n \to \infty} p(G, H_n)$ exists. An alternative way of seeing convergence is that each graph $H$ defines a point $p(-, H) \in [0, 1]^\mathcal{M}$, where $\mathcal{M}$ is the set of all finite graphs up to isomorphism and convergence of an increasing sequence $(H_n)_{n \in \mathbb{N}}$ is simply convergence of the corresponding sequence $(p(-, H_n))_{n \in \mathbb{N}}$ with respect to the product topology of $[0, 1]^\mathcal{M}$. Since $\mathcal{M}$ is countable, $[0, 1]^\mathcal{M}$ is metrizable and since it is compact, it follows that any increasing sequence of finite graphs has a convergent subsequence.

The main point of graphon theory is that convergent sequences can be encoded by a geometric limit object in which limits of labeled densities can be naturally computed. Formally, given an atomless standard probability space $\Omega = (X, \mathcal{A}, \mu)$, a graphon over $\Omega$ is a function $W: X \times X \to [0, 1]$ that is symmetric and is measurable with respect to the completion of the product $\sigma$-algebra $\mathcal{A} \otimes \mathcal{A}$ with respect to the product measure $\mu \otimes \mu$. Typically, we take $\Omega$ as $[0, 1]$ equipped with the Lebesgue measure $\lambda$ over Borel sets, in which case we say “graphon over $[0, 1]$” or simply “graphon” (which is simply a symmetric Lebesgue measurable function $[0, 1]^2 \to [0, 1]$). The intuition is that a graphon over $\Omega = (X, \mathcal{A}, \mu)$ is a graph with vertex set $X$ in which edges can have fractional values and $W(x, y)$ should be interpreted as the “probability” that $x$ and $y$ are adjacent. With this intuition in mind, the labeled (induced) density of a graph $G$ in a graphon $W$ is naturally defined as

$$t_{\text{ind}}(G, W) \overset{\text{def}}{=} \int_{X^V(G)} \prod_{\{v, w\} \in E(G)} W(x_v, x_w) \prod_{\{v, w\} \in E(\bar{G})} (1 - W(x_v, x_w)) \, d\mu(x),$$

where $\text{Aut}(G)$ is the group of automorphisms of $G$. We denote by $\rho$ the edge graph, so that $p(\rho, H) = t_{\text{ind}}(\rho, H)$ denotes the edge density of $H$.
where $E(G) \overset{\text{def}}{=} \{(v, w) \in \binom{V(G)}{2} \mid G \models E(v, w)\}$ is the edge set of $G$ and $\overline{G}$ is the complement graph of $G$. We also define $p(G, W) \overset{\text{def}}{=} |G|! \cdot t_{\text{ind}}(G, W)/|\text{Aut}(G)|$ in analogy with (1) and we let $\phi_W \overset{\text{def}}{=} p(-, W)$. We say that $W$ is a limit of a convergent sequence $(H_n)_{n \in \mathbb{N}}$ if $\lim_{n \to \infty} p(G, H_n) = \phi_W(G)$ for every finite graph $G$.

The following theorem, sometimes referred to as Existence Theorem for graphons, is the main theorem of graphon theory.

**Theorem 3.1** (Lovász–Szegedy [LS06]). Let $\Omega$ be an atomless standard probability space. If $(H_n)_{n \in \mathbb{N}}$ is a convergent sequence of graphs, then there exists a graphon $W$ over $\Omega$ that is a limit of $(H_n)_{n \in \mathbb{N}}$. Conversely, every graphon is a limit of a convergent sequence of graphs.

It is trivial that two convergent sequences can converge to the same limit graphon as only the tail behavior of the convergent sequences matters and changes to $o(|H_n|^2)$ edges do not affect densities. On the other side, more than one graphon can represent the limit of the same convergent sequence. For example, any graphon $W : [0, 1]^2 \to [0, 1]$ over $[0, 1]$ represents the same limit as $W'$ given by $W'(x, y) = W(2x \mod 1, 2y \mod 1)$ (see Figure 1 for an example). The Uniqueness Theorem for graphons [BCL10] (see also [Lov12, Theorem 13.10]) characterizes when two graphons represent the same limit using measure-preserving functions.

For our theorems, we will be considering subgraphons, which are a limit world generalization of the notion of induced subgraph, but only by non-negligible sets (i.e., sets of linear size).

**Definition 3.2.** Given a graphon $W$ over $\Omega$, a (positive measure) subgraphon $W'$ of $W$ is a graphon over a space $\Omega'$ such that there exist a sequence $(H_n)_{n \in \mathbb{N}}$ converging to $W$ and sets $U_n \subseteq V(H_n)$ such that $\lim_{n \to \infty} |U_n|/|H_n| > 0$ and $(H_n|_{U_n})_{n \in \mathbb{N}}$ converges to $W'$. When we want to be more specific, we say that $W'$ is a subgraphon of $W$ of measure $c \in (0, 1]$, if the condition above holds with $\lim_{n \to \infty} |U_n|/|H_n| = c$.

Naïvely, one might think that each subgraphon of a graphon $W$ could be represented as $W|_{U \times U}$ for some positive measure set $U \subseteq X$ over the space $\Omega' \overset{\text{def}}{=} (U, \mathcal{A}|_U, \mu|_U)$, where $\mathcal{A}|_U \overset{\text{def}}{=} \{A \cap U \mid A \in \mathcal{A}\}$ and $\mu|_U(A) \overset{\text{def}}{=} \mu(A)/\mu(U)$. There are two problems with this naïve definition. The first is only technical: $\Omega'$ is not necessarily a standard probability space, but this can be addressed by conditioning the measure rather than restricting the space.
Figure 1: Different representations $W_k(x, y) \overset{\text{def}}{=} W(kx \mod 1, ky \mod 1)$ of the graphon $W$ given by

$$W(x, y) = \begin{cases} 1, & \text{if } \lfloor x/2 \rfloor \neq \lfloor y/2 \rfloor \text{ and } (\lfloor x/2 \rfloor < \lfloor y/2 \rfloor \leftrightarrow x \mod (1/2) < y \mod (1/2)), \\ 0, & \text{otherwise.} \end{cases}$$

The graphon $W$ is called the half-graphon.

by using the space $\Omega_U = (X, A, \mu_U)$, where $\mu_U(A) = \mu(A \cap U)/\mu(U)$. The second is more serious: not every limit of a sequence of the aforementioned form $(H_n|_{U_n})_{n \in \mathbb{N}}$ is necessarily encoded this way. However, the next lemma says that this description is not too far from correct, we only need to “rescale” the underlying measure by a weight function.

**Lemma 3.3.** Let $W$ be a graphon over $\Omega = (X, A, \mu)$, let $W'$ be another graphon and let $c > 0$. The following are equivalent.

i. There exist a convergent sequence of graphs $(H_n)_{n \in \mathbb{N}}$ converging to $W$ and sets $U_n \subseteq V(H_n)$ with $\lim_{n \to \infty} |U_n|/|H_n| = c$ such that $(H_n|_{U_n})_{n \in \mathbb{N}}$ converges to $W'$, that is, $W'$ is a subgraphon of $W$ of measure $c$.

ii. There exists a measurable function $f : X \to [0, 1]$ with $\int_X f \ d\mu = c$ such that $\phi_{W'} = \phi_{W_f}$, where $W_f$ is the graphon over the the space
\[ \Omega_f \overset{\text{def}}{=} (X, A, \mu_f) \text{ defined by} \]
\[ \mu_f(A) \overset{\text{def}}{=} \frac{\int_A f(x) \, d\mu(x)}{c}, \]
\[ W_f(x, y) \overset{\text{def}}{=} W(x, y). \]

We defer the proof of this lemma as it is a particular case of the more general Lemma 5.8.

As we mentioned before, one way of interpreting a graphon \( W \) is as a measurable “graph” over \( \Omega \), except that \( W(x, y) \) is the “probability” that \( x \) and \( y \) are adjacent. Under this interpretation, \( \{0, 1\} \)-valued graphons are simply measurable graphs and we can reinterpret the labeled density formula (2) as follows. The set of labeled (induced) copies of a finite graph \( G \) in a graphon \( W \) over \( \Omega = (X, A, \mu) \) is the set

\[ T_{\text{ind}}(G, W) \overset{\text{def}}{=} \left\{(x, y) \in X^{V(G)} \times [0, 1)^{V(G)} \right\} \]

\[ \forall \{v, w\} \in \binom{V(G)}{2}, (y_{\{v, w\}} < W(x_v, x_w) \leftrightarrow \{v, w\} \in E(G)) \right\}. \]

Under this definition, we have \( t_{\text{ind}}(G, W) = (\mu^{V(G)} \otimes \lambda^{\binom{V(G)}{2}})(T_{\text{ind}}(G, W)) \).

Note also that if \( W \) is a \( \{0, 1\} \)-valued graphon and we interpret it as simply a measurable graph on \([0, 1]\), whenever all coordinates of \( x \) are distinct, we have \((x, y) \in T_{\text{ind}}(G, W)\) if and only if \( x \) is an embedding of \( G \) in \( W \).

In the same way that the usual (dense setting) Graph Removal Lemma [RS78, EFR86] (see also [Lov12, Lemma 11.64 and Theorems 15.24 and 15.25]) says that we can change a negligible fraction of edges to remove graphs that have negligible density, the following graphon version says that \( t_{\text{ind}}(G, W) = 0 \) can be turned into \( T_{\text{ind}}(G, W) \) morally empty by changing \( W \) only in a zero-measure set (see also Theorem 5.4 for the general case). In fact, Elek–Szegedy showed [ES12, Theorem 1] that the finite version of the Removal Lemma follows from a connection of limit theory via ultraproducts that we will see later.

**Theorem 3.4** (Graphon Removal Lemma [Pet13, Theorem 1]). If \( W \) is a graphon over \( \Omega = (X, A, \mu) \), then there exists a graphon \( W' \) over \( \Omega \) such that \( W = W' \) a.e. and for every finite graph \( G \) such that \( t_{\text{ind}}(G, W) = 0 \), we
have $T_{ind}(G, W') \subseteq \mathcal{D}_{V(G)}$, where

$$
\mathcal{D}_{V(G)} \overset{\text{def}}{=} \left\{ (x, y) \in X^{V(G)} \times [0,1)^{V(G)} \mid \exists v, w \in V(G), (v \neq w \land x_v = x_w) \right\}
$$

(3)

denotes the diagonal set with respect to the $x$ variables.

Furthermore, if $W$ is $\{0,1\}$-valued, then $W'$ can also be taken to be $\{0,1\}$-valued.

The final concept needed to state our graphon dichotomy theorem is that of an almost stable graphon defined below.

**Definition 3.5.** Recall that a half-graph of order $n$ in a graph $G$ (see Figure 2) is pair of sequences $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$ of vertices of $G$ such that $\{x_i, y_j\} \in E(G)$ if and only if $i \leq j$.

We say that a tree of height $n$ in a graph $G$ (see Figure 3) is pair of sequences $(x_\sigma \mid \sigma \in \{0,1\}^n)$ and $(y_\tau \mid m \in \{0,1,\ldots,n-1\}, \tau \in \{0,1\}_m)$ such that for every $\sigma = (\sigma_i)_{i=1}^n \in \{0,1\}^n$ and every $m < n$, $\{x_\sigma, y_{\sigma[|m]}\} \in E(G)$ if and only if $\sigma_{m+1} = 1$ (trees of height $n$ are also known under several different names in the literature).

Recall also that a graph is called $n$-stable (or more formally, its edge relation is $n$-stable) if it does not contain any half-graphs of order $n$.

A graphon $W$ is almost stable if there exists $n \in \mathbb{N}$ such that every finite graph $G$ containing a half-graph of order $n$ satisfies $p(G, W) = 0$.

Our use of stability will be mainly to provide a bound on the height of trees, since by [Hod93, Lemma 6.7.9], an $n$-stable graph does not contain any trees of height $2^{n+2} - 2$ and, conversely, if a graph does not contain a tree of height $n$, then it is $(2^{n+2} - 2)$-stable.

Figure 2: Half-graph of order 7. Between two distinct $x_i$'s or between two distinct $y_i$'s there is no edge/non-edge requirement.
Theorem 3.6. The following are equivalent for a graphon $W$ over $\Omega = (X, \mathcal{A}, \mu)$.

i. $W$ contains a subgraphon that is either constant equal to 1 or constant equal to 0.

ii. There exists a positive measure set $U \subseteq X$ such that either $W|_{U \times U} = 1$ a.e. or $W|_{U \times U} = 0$ a.e.

iii. $W$ contains an almost stable subgraphon.

Discussion 3.7. As we noted before, not every subgraphon of $W$ is of the form $W|_{U \times U}$ and thus the equivalence of items (i) and (ii) is not trivial.

However, if $P$ is property of graphons that is closed under taking subgraphons, then a graphon $W$ has a subgraphon satisfying $P$ if and only if it has a positive measure $U \subseteq X$ such that $W|_{U \times U}$ satisfies $P$. The backward implication is obvious, and the forward implication can be seen easily from Lemma 3.3: if the subgraphon $W'$ satisfying $P$ corresponds to a measurable function $f$ of positive integral, then for some $\epsilon > 0$, the set $U_\epsilon \overset{\text{def}}{=} \{x \in X \mid f(x) > \epsilon\}$ has positive measure and $W|_{U_\epsilon \times U_\epsilon}$ satisfies $P$ as it is a subgraphon of $W'$. 
As we will see in Theorem 4.3, the importance of item (ii) is that if $(H_n)_{n \in \mathbb{N}}$ converges to $W$, then subgraphons of form $W|_{U \times U}$ can be pulled back to $(H_n|_{U_n})_{n \in \mathbb{N}}$ without changing the sequence $(H_n)_{n \in \mathbb{N}}$.

The main ingredient to prove this theorem is the following lemma whose main idea can be seen as a graphon analogue of the construction of $\epsilon$-good sets in [MS14, MS21], but with $\epsilon = 0$ and is much easier for measure theoretic reasons.

**Lemma 3.8.** Let $W$ be an almost stable graphon over $\Omega = (X, \mathcal{A}, \mu)$. Then there exists a positive measure set $U \subseteq X$ such that either $W|_{U \times U} = 1$ a.e. or $W|_{U \times U} = 0$ a.e.

**Proof.** By [LS10, Theorem 4.1], we know that $W$ is $\{0,1\}$-valued almost everywhere, so we can change it in a zero-measure set so that it is $\{0,1\}$-valued.

By Theorem 3.4, we can further replace $W$ with a $\{0,1\}$-valued graphon $W'$ such that there exists $n \in \mathbb{N}$ such that every finite graph $G$ containing a half-graph of order $n$ satisfies $T_{ind}(G,W') \subseteq D_{V(G)}$, that is, $W'$ as a measurable graph over $\Omega$ is $n$-stable (except for potential half-graphs that collide vertices).

For $x \in X$, let $N_{W'}(x) \overset{\text{def}}{=} \{y \in X \mid W'(x,y) = 1\}$ denote the "neighborhood" of $x$ in $W'$ and let $X'$ be the set of $x \in X$ such that $N_{W'}(x)$ is measurable with respect to the completion $(\mathcal{A}',\mu')$ of $(\mathcal{A},\mu)$. Fubini’s Theorem gives $\mu'(X') = 1$.

We now construct sequences $(X_\sigma)_\sigma$ and $(y_\sigma)_\sigma$ indexed by finite strings over $\{0,1\}$ inductively in the length of $\sigma$ as follows.

1. Set $X_\emptyset \overset{\text{def}}{=} X'$.

2. Given $X_\sigma$, if there exists $z \in X'$ such that $0 < \mu'(N_{W'}(z) \cap X_\sigma) < \mu'(X_\sigma)$, then set $y_\sigma \overset{\text{def}}{=} z$, $X_{\sigma 1} \overset{\text{def}}{=} X_\sigma \cap N_{W'}(z)$ and $X_{\sigma 0} \overset{\text{def}}{=} X_\sigma \setminus N_{W'}(z)$; otherwise stop the construction.

Let also $Y$ be the (countable) set of all $y_\sigma$ that get defined in the construction above. By induction, it follows that if $X_\sigma$ is defined for every $\sigma \in \{0,1\}^t$ of a fixed length $t$, then $\{X_\sigma \mid \sigma \in \{0,1\}^t\}$ forms a measurable partition of $X'$ into sets of positive measure (hence non-empty).

---

1As we will see in Theorem 5.11 for general universal theories, this step is not actually necessary, but it simplifies the proof in the graphon case.
Furthermore, if \( x_\sigma \in X_\sigma \setminus Y \) (\( \sigma \in \{0,1\}^t \)) then \( (x_\sigma \mid \sigma \in \{0,1\}^t) \) and \( \{y_\tau \mid m \in \{0,1,\ldots,t-1\}, \tau \in \{0,1\}^m\} \) form a tree of height \( t \) in \( W' \).

By [Hod93, Lemma 6.7.9], we know that \( n \)-stable graphs do not contain trees of height \( 2n+2-2 \), so the construction above must stop otherwise it would produce an off-diagonal copy of a graph containing a half-graph of order \( n \) (the fact that the copy is off-diagonal follows since no \( x_\sigma \) is equal to any \( y_\tau \)). Thus there exists \( \tilde{\sigma} \) such that for every \( z \in X' \), we have \( \mu'(N_{W'}(z) \cap X_{\tilde{\sigma}})/\mu'(X_{\tilde{\sigma}}) \in \{0,1\} \). Let

\[
Z_i \overset{\text{def}}{=} \left\{ z \in X_{\tilde{\sigma}} \mid \frac{\mu'(N_{W'}(z) \cap X_{\tilde{\sigma}})}{\mu'(X_{\tilde{\sigma}})} = i \right\} \quad (i \in \{0,1\})
\]

and since \( X_{\tilde{\sigma}} = Z_0 \cup Z_1 \) and \( \mu'(X_{\tilde{\sigma}}) > 0 \), there exists \( i_0 \in \{0,1\} \) such that \( \mu'(Z_{i_0}) > 0 \). Finally, taking \( U \in A \) such that \( \mu'(Z_{i_0} \triangle U) = 0 \) gives \( \mu(U) = \mu'(Z_{i_0}) > 0 \) and \( W|_{U \times U} = W'|_{U \times U} = i_0 \) a.e. \( \square \)

**Discussion 3.9.** As we mentioned before, the proof of Lemma 3.8 can be seen as the construction of a 0-good set \( X_{\tilde{\sigma}} \) in the graphon \( W' \), that is, a positive measure set \( U \subseteq X \) such that almost every \( x \in X \) is either adjacent to almost all of \( U \) or almost none of \( U \) in the sense that

\[
\frac{1}{\mu(U)} \int_U W(x,y) \, d\mu(y) \in \{0,1\}.
\]

Another important concept in [MS14, MS21] that can be generalized to graphons is that of excellent sets. Let us say that a 0-excellent set in a graphon \( W \) is a 0-good set\(^2\) \( U \) such that for every 0-good set \( V \) we either have almost all edges between \( U \) and \( V \) or we have almost no edges between \( U \) and \( V \) in the sense that

\[
\frac{1}{\mu(U \times V)} \int_{U \times V} W(x,y) \, d\mu(x,y) \in \{0,1\}.
\]

Under this definition, it is easy to generalize the proof of Lemma 3.8 to prove that every 0-good set \( U \) in an almost stable graphon \( W \) contains some 0-excellent set: one can simply repeat the inductive construction by starting

\(^2\)In the finite case, we do not need to explicitly require excellent sets to be good as the goodness property follows from the excellent property when \( V \) is a single vertex (which is necessarily a good set in the finite). However, in the limit, a single vertex is not good as it does not have positive measure.
with $X_\Box \overset{\text{def}}{=} U$ an use 0-good sets $Y_\sigma$ for the internal nodes instead of single
vertices $y_\sigma$. Composing these two and with a transfinite induction, it then
follows that if $W$ is an almost stable graphon over $\Omega = (X, A, \mu)$, then there
exists a countable partition $(U_i)_{i \in I}$ of $X$ into positive measure sets such that
for each $i, j \in I$, there exists $b_{i,j} \in \{0, 1\}$ such that $W_{|{U_i \times U_j}} \equiv b_{i,j}$ a.e., that
is, $W$ is a $\{0, 1\}$-valued “countable step-graphon”. This can be seen as a
0-error version of the stable regularity lemma [MS14, MS21] in the limit.

Let us now show the stability dichotomy theorem for graphons.

Proof of Theorem 3.6. The implication $(i) \implies (iii)$ is trivial as the constant
0 and constant 1 graphons are almost stable.

For the implication $(ii) \implies (i)$, using Lemma 3.3 with the indicator
function $f \overset{\text{def}}{=} 1_U$ of $U$, we obtain a subgraphon $W_f$ that is either a.e. equal
to 1 or a.e. equal to 0.

For the final implication $(iii) \implies (ii)$, if $W'_f$ is an almost stable subgraphon
of $W$, then by Lemma 3.3 there exists $f : X \to [0, 1]$ with $\int_X f \, d\mu > 0$
such that $\phi_{W'} = \phi_{W_f}$ for the graphon $W_f$ over $\Omega_f = (X, A, \mu_f)$ given by
$W_f(x, y) \overset{\text{def}}{=} W(x, y)$. Let $V \overset{\text{def}}{=} \{x \in X \mid f(x) > 0\}$, let $g \overset{\text{def}}{=} 1_V$ be the
indicator function of $V$ and consider the subgraphon $W_g$ of $W$ over the space
$\Omega_g = (X, A, \mu_g)$ corresponding to $g$ via Lemma 3.3.

We claim that $W_g$ is almost stable. This is a standard measure theoretic
trick: for $\epsilon > 0$ and a finite graph $G$, let

\[
T_{\text{ind}}^\epsilon(G, W_f) \overset{\text{def}}{=} \{(x, y) \in T_{\text{ind}}(G, W_f) \mid \forall v \in V(G), f(x_v) > \epsilon\},
\]

\[
T_{\text{ind}}^\epsilon(G, W_g) \overset{\text{def}}{=} \{(x, y) \in T_{\text{ind}}(G, W_g) \mid \forall v \in V(G), f(x_v) > \epsilon\},
\]

then it is easy to see that

\[
(\mu_f^{V(G)} \otimes \lambda^{V(G)}_2)(T_{\text{ind}}^\epsilon(G, W_f))
\geq \left( \epsilon \cdot \frac{\mu(V)}{\int_X f \, d\mu} \right)^{|G|} (\mu_g^{V(G)} \otimes \lambda^{V(G)}_2)(T_{\text{ind}}^\epsilon(G, W_g)).
\]

On the other hand, since we have $T_{\text{ind}}(G, W_f) = \bigcup_{n \in \mathbb{N}^+} T_{\text{ind}}^{1/n}(G, W_f)$ $\mu_f$-a.e.
and $T_{\text{ind}}(G, W_g) = \bigcup_{n \in \mathbb{N}^+} T_{\text{ind}}^{1/n}(G, W_g)$ $\mu_g$-a.e., it follows that

$\forall G, (t_{\text{ind}}(G, W_f) = 0 \implies t_{\text{ind}}(G, W_g) = 0)$
and since $W_f$ is almost stable, we get that $W_g$ is almost stable.

By Lemma 3.8, there exists a measurable set $\tilde{U} \subseteq X$ such that $\mu_g(\tilde{U}) > 0$ and either $W_g|_{\tilde{U} \times \tilde{U}} = 1$ $\mu_g$-a.e. or $W_g|_{\tilde{U} \times \tilde{U}} = 0$ $\mu_g$-a.e. The result now follows by setting $U \overset{\text{def}}{=} \tilde{U} \cap V$ since $\mu_g(A) = \mu(A \cap V)/\mu(V)$ for every measurable set $A \subseteq X$.

**Remark 3.10.** Recall that the set $U$ produced by Lemma 3.8 actually has a stronger property than simply almost clique or almost anti-clique, namely, it is a 0-good set. By tracking down the application of this lemma in the proof of Theorem 3.6 above, we conclude that if a graphon $W$ contains some almost stable subgraphon $W'$ of measure $c$, then it contains positive measure sets $U$ and $V$ such that $\mu(V) \geq c$, $W|_{V \times V}$ is almost stable and $U$ is a 0-good set in $W|_{V \times V}$.

A natural question that arises is whether it is possible for a graphon to not contain any almost stable subgraphon. A trivial example is obtained by considering quasirandom graphons: for $p \in [0, 1]$, let $W_p$ be the constant $p$ graphon. It is not hard to see from Lemma 3.3 that $W_p$ is the only subgraphon of $W_p$. In fact, the content of one of the original graph quasirandomness equivalences [CGW89, $P_1 \iff P_4$] (see also [SS97, Theorem 3.4]) is precisely that these are the only graphons with this property. Since $W_p$ is not almost stable if $0 < p < 1$ (as $t_{\text{ind}}(G, W_p) = p|E(G)|/(1 - p)|E(G)| > 0$ for every finite graph $G$), it follows that none of its subgraphons are either.

One can then ask if this is not an artifact of the fact that $W_p$ has fractional values, that is, could it be that $\{0, 1\}$-valued graphons must necessarily contain some almost stable subgraphon? The next example answers this in the negative. We will also show in Lemma 8.8 that the recursive blow-up of $C_4$ (see Definition 8.5) is another such example.

**Example 3.11.** Let $\Omega$ be the space $[0, 1]^2$ equipped with the 2-dimensional Lebesgue measure over Borel sets and consider the graphon $W$ over $\Omega$ given by

$$W((x_1, x_2), (y_1, y_2)) \overset{\text{def}}{=} \begin{cases} 1, & \text{if } x_1 < y_1 \leftrightarrow x_2 < y_2 \text{ and } x_1, x_2, y_1, y_2 \text{ are distinct;} \\ 0, & \text{otherwise} \end{cases}$$

Clearly $W$ is $\{0, 1\}$-valued.

We claim that for every positive measure set $U \subseteq [0, 1]^2$, $W|_{U \times U}$ is not a.e. constant. By Theorem 3.6, this in particular means that $W$ does not have any almost stable subgraphon.
Note that if $W|_{U \times U} = 1$ a.e., then for every $n \in \mathbb{N}$, we must have $t_{\text{ind}}(K_n, W) \geq \lambda(U)^n$. On the other hand, if $W|_{U \times U} = 0$ a.e., then for every $n \in \mathbb{N}$, we must have $t_{\text{ind}}(K_n, W) \geq \lambda(U)^n$. In fact, by [CKP21, Theorem 6], for any graphon $W'$, we have
\[
\sup\{\lambda(U) \mid W'|_{U \times U} = 1 \text{ a.e.}\} = \lim_{n \to \infty} t_{\text{ind}}(K_n, W')^{1/n},
\]
\[
\sup\{\lambda(U) \mid W'|_{U \times U} = 0 \text{ a.e.}\} = \lim_{n \to \infty} t_{\text{ind}}(K_n, W')^{1/n}.
\]
However, it is easy to see that
\[
t_{\text{ind}}(K_n, W) = t_{\text{ind}}(K_n, W) = \frac{1}{n!}
\]
as $t_{\text{ind}}(K_n, W)$ is the probability that the relative order of the coordinates of $x$ matches that of the coordinates of $y$ when both are picked independently and uniformly in $[0,1]^n$ and $t_{\text{ind}}(K_n, W)$ is that these relative orders are the precise inverses of each other (a more detailed explanation will be given in Example 5.13). Since for every $c > 0$, there exists $n \in \mathbb{N}$ such that $1/n! < c^n$, the claim follows.

To visualize $W$, we can consider the standard measure-isomorphism $F$ modulo 0 from $[0,1]^2$ to $[0,1]$ that maps $(w,z) = (0.w_1w_2\cdots,0.z_{1}z_{2}\cdots)$ to $0.w_1z_1w_2z_2\cdots$ using the binary expansions of the numbers $w$ and $z$. The graphon $W'$ over $[0,1]$ given indirectly by $W'(F(x),F(y)) \overset{\text{def}}{=} W(x,y)$ then represents the same limit as $W$, see Figure 4.

**Discussion 3.12.** Another way of seeing Example 3.11 is as a “recursive half-graphon” that does not contain any almost stable graphon: we start by splitting the space $[0,1]^2$ into two parts $A_0 \overset{\text{def}}{=} [0,1] \times [0,1/2]$ and $A_1 \overset{\text{def}}{=} [0,1] \times (1/2,1]$ and put a half-graphon (see Figure 1) between these two parts by setting $W((x_1,y_1),(x_2,y_2)) \overset{\text{def}}{=} 1[x_1 < x_2]$ for every $(x_1,y_1) \in A_0$ and every $(x_2,y_2) \in A_1$. We then split each of the halves in two and proceed recursively splitting the space along the dyadics in the second coordinate. It is easy to see that this recursive construction gives the graphon of Example 3.11, which intuitively has half-graphons within every subgraphon. We will see in Example 5.13 that another way of interpreting this graphon is as the graphon of agreements of the quasirandom permuton.
Figure 4: Approximation of the graphon $W'$ of Example 3.11. The graphon $W'$ has a fractal structure, whose first 3 steps are represented in the picture.

4 Consequences for finite graphs

The objective of this section is to transfer Theorem 3.6 to the finite world. There are several different ways that one can construct different geometric limit objects that encode convergent sequences, each of the different approaches brings to light new connections between the finite and the infinite. The approach of Lovász–Szegedy [LS06] (see also [Lov12]) relied on Szemerédi’s Regularity Lemma [Sze78] and the graph cut-norm, the approach ofDiaconis–Janson [DJ08] uses the theory of exchangeable arrays (see [Kal05] for more on this theory), the approach of Elek–Szegedy [ES12] uses ultraproducts and, more recently, the approach of Doležal–Grešík–Hladký–Rocha–Rozhoň [DGH+21] uses weak∗ convergence when we think of the space of graphons as $L^\infty(\Omega^2)$.

To transfer Theorem 3.6, the ultraproduct method Elek–Szegedy [ES12] (and its generalization by Aroskar–Cummings [AC14]) will be of particular importance as it allows pulling back properties from the infinite to convergent sequences via Loś’s Theorem for ultraproducts. We describe this method informally here and defer formal definitions to Appendix A (we refer the combinatorially oriented reader to [ES12, §2.7] for an application-oriented introduction to (countable) ultrafilters and ultraproducts and to [CK90, Chapter 4]
Consider a sequence of graphs \((H_n)_{n \in \mathbb{N}}\) of increasing sizes and let \(V_n \overset{\text{def}}{=} V(H_n)\). Note that for each graph \(G\), the set \(T_{\text{ind}}(G, H_n)\) of embeddings of \(G\) in \(H_n\) can be seen as a subset of \(V_n^{\mathcal{V}(G)}\) and we have
\[
t_{\text{ind}}(G, H_n) = \mu_n^{\mathcal{V}(G)}(T_{\text{ind}}(G, H_n)) + o_G(1),
\]
where \(\mu_n^{\mathcal{V}(G)}\) is the normalized counting measure on \(V_n^{\mathcal{V}(G)}\) given by \(\mu_n^{\mathcal{V}(G)}(A) = \frac{|A|}{|H_n|^{\mathcal{V}(G)}}\) and the error term \(o_G(1)\) goes to 0 as \(n \to \infty\) for each fixed \(G\) and accounts for the fact that the normalization in \(t_{\text{ind}}\) is \((|H_n|)|G|\) instead of \(|H_n|^{\mathcal{V}(G)}|G|\). We then consider the ultraproduct \(H \overset{\text{def}}{=} \prod_{n \in \mathbb{N}} H_n/\mathcal{D}\) and note that Los’ Theorem (see [Hod93, Theorem 9.5.1]) implies that the set of embeddings \(T_{\text{ind}}(G, H)\) of \(G\) in \(H\) is an internal set of \(\prod_{n \in \mathbb{N}} V_n^{\mathcal{V}(G)}/\mathcal{D}\), namely, we have \(T_{\text{ind}}(G, H) = \prod_{n \in \mathbb{N}} T_{\text{ind}}(G, H_n)/\mathcal{D}\). Going one step further if \(\mu_n^{\mathcal{V}(G)}\) is the Loeb measure corresponding to \((\mu_n^{\mathcal{V}(G)})_{n \in \mathbb{N}}\), then we have
\[
\mu^{\mathcal{V}(G)}(T_{\text{ind}}(G, H)) = \mu^{\mathcal{V}(G)} \left( \prod_{n \in \mathbb{N}} T_{\text{ind}}(G, H_n)/\mathcal{D} \right) = \lim_{n \to \mathcal{D}} \mu_n^{\mathcal{V}(G)}(T_{\text{ind}}(G, H_n)) = \lim_{n \to \mathcal{D}} t_{\text{ind}}(G, H_n).
\]
If the sequence \((H_n)_{n \in \mathbb{N}}\) is convergent, this ultralimit must be equal to the actual limit \(\lim_{n \to \infty} t_{\text{ind}}(G, H_n)\); this means that the ultraproduct \(\prod_{n \in \mathbb{N}} H_n/\mathcal{D}\) along with the Loeb measures \(\mu_U\) for each finite set \(U\) successfully encode the “limit” of the sequence \((H_n)_{n \in \mathbb{N}}\). The “problem” with this encoding is that equipping \(\prod_{n \in \mathbb{N}} V_n^U/\mathcal{D}\) with the Loeb measure \(\mu_U\) gives a probability space that is far from being standard, namely it is non-separable. Moreover, if \(\sigma(U)\) is the \(\sigma\)-algebra of \(\mu^U\), then for \(U_1, U_2\) disjoint and non-empty, \(\sigma(U_1 \cup U_2)\) contains many more sets than the completion of the product \(\sigma(U_1) \otimes \sigma(U_2)\) with respect to the product measure \(\mu^{U_1} \otimes \mu^{U_2}\) (even though Fubini’s Theorem still holds: the \(\mu^U\)-measure of a set in \(\sigma(U)\) can be computed via iterated integrals with respect to \(\mu^{U_1}\) and \(\mu^{U_2}\), see Theorem A.2).

To address this issue, Elek and Szegedy encode this ultraproduct probability space in the probability space \([0,1]^{r(k)}\) equipped with the Lebesgue measure (recall that \(r(k) \overset{\text{def}}{=} \{ A \subseteq [k] \mid A \neq \emptyset \}\)); this encoding is done via separable realizations, which can be seen as a structured version of...
Maharam’s Theorem [Mah42]. Informally, a separable realization of order \( k \in \mathbb{N}_+ \) is a measure-preserving function \( \Theta: \prod_{n \in \mathbb{N}} V_n^k / D \to [0, 1]^{r(k)} \) that preserves enough structure of the probability space so that:

i. For each \( m \in [k] \), there exists a restriction of \( \Theta \) of order \( m \), that is, a separable realization \( \Theta_m: \prod_{n \in \mathbb{N}} V_n^m / D \to [0, 1]^{r(m)} \) of order \( m \) such that the diagram

\[
\begin{array}{ccc}
\prod_{n \in \mathbb{N}} V_n^k / D & \xrightarrow{\Theta} & [0, 1]^{r(k)} \\
\downarrow \alpha^* & & \downarrow \alpha^* \\
\prod_{n \in \mathbb{N}} V_n^m / D & \xrightarrow{\Theta_m} & [0, 1]^{r(m)}
\end{array}
\]

commutes for every injection \( \alpha: [m] \hookrightarrow [k] \) (recall that \( \alpha^* \) is the “projection” given by \( \alpha^*(x)_u = x_{\alpha(u)} \)).

ii. For every \( m \geq k \), there exists a lifting of \( \Theta \) of order \( m \), that is, a measure-preserving \( \Theta_m: \prod_{n \in \mathbb{N}} V_n^m / D \to [0, 1]^{r(m,k)} \) such that the diagram

\[
\begin{array}{ccc}
\prod_{n \in \mathbb{N}} V_n^m / D & \xrightarrow{\Theta_m} & [0, 1]^{r(m,k)} \\
\downarrow \alpha^* & & \downarrow \alpha^* \\
\prod_{n \in \mathbb{N}} V_n^k / D & \xrightarrow{\Theta} & [0, 1]^{r(k)}
\end{array}
\]

commutes for every injection \( \alpha: [k] \hookrightarrow [m] \).

The main theorem of the ultraproduct method for hypergraphs is then the following, which we state only for the graph case (see also Theorem 6.1 for the general case).

**Theorem 4.1** (Elek–Szegedy [ES12]). For every sequence of graphs of increasing sizes \( (H_n)_{n \in \mathbb{N}} \) and every non-principal ultrafilter \( D \) over \( \mathbb{N} \), there exists a separable realization \( \Theta: \prod_{n \in \mathbb{N}} V(H_n)^2 / D \to [0, 1]^{r(2)} \) of order 2 and a measurable set \( \mathcal{N} \subseteq [0, 1]^{r(2)} \) such that

\[
\mu^2 \left( \Theta^{-1}(\mathcal{N}) \Delta \prod_{n \in \mathbb{N}} E^{H_n} / D \right) = 0,
\]

where \( E^{H_n} \overset{\text{def}}{=} \{ (v, w) \in V(H_n)^2 \mid H_n \models E(v, w) \} \) is the set of edges of \( H_n \) as ordered pairs.
With the theorem above, one can then define the graphon
\[ W(x, y) \stackrel{\text{def}}{=} \lambda(\{ z \in [0, 1] \mid (x, y, z) \in \mathcal{N} \}) \]
and from the properties of restrictions and liftings of the separable realization \( \Theta \), it follows that if \( G \) is a finite graph with \( V(G) = [m] \), then
\[
\mu^m \left( T_{\text{ind}} \left( G, \prod_{n \in \mathbb{N}} H_n / \mathcal{D} \right) \right) = \mu^m (\Theta_m^{-1}(T_{\text{ind}}(G, \mathcal{N}))) = \lambda(T_{\text{ind}}(G, \mathcal{N})) = \lambda(T_{\text{ind}}(G, W)) = t_{\text{ind}}(G, W),
\]
where
\[
T_{\text{ind}}(G, \mathcal{N}) \stackrel{\text{def}}{=} \left\{ x \in [0, 1]^{r(m, 2)} \mid \forall \{v, w\} \in \binom{V(G)}{2}, \right. \\
\left. (\{v, w\} \in E(G) \leftrightarrow (x_{\{v\}}, x_{\{w\}}, x_{\{v, w\}}) \in \mathcal{N}) \right\}.
\]
This completes the translation of convergent sequences into graphons.

In the next theorem, we explore this connection by pulling back the almost clique or almost empty graphon \( W|_{U \times U} \) provided by Theorem 3.6 through the separable realization of Theorem 4.1 to the ultraproduct and producing a linear-sized almost clique or almost empty graph in the convergent sequence. Naturally, we need the analogue of almost stability for increasing sequences of graphs.

**Definition 4.2.** We say that an increasing sequence of graphs \( (H_n)_{n \in \mathbb{N}} \) is **almost stable** if there exists \( \ell \in \mathbb{N} \) such that every finite graph \( G \) containing a half-graph of order \( \ell \) satisfies \( \lim_{n \to \infty} p(G, H_n) = 0 \).

The next theorem is the stability dichotomy for convergent sequences of graphs, which in plain English says that a convergent sequence of graphs \( (H_n)_{n \in \mathbb{N}} \) contains a **sequence** of linear-sized induced subgraphs that is either an almost clique or an almost anti-clique if and only if it contains a **subsequence** of linear-sized induced subgraphs that is almost stable. A posteriori, it is clear that these conditions are also equivalent to \( (H_n)_{n \in \mathbb{N}} \) containing a **subsequence** of linear-sized induced subgraphs that is either an almost clique or an almost anti-clique and equivalent to \( (H_n)_{n \in \mathbb{N}} \) containing a **sequence** of linear-sized induced subgraphs that is almost stable.
Theorem 4.3. The following are equivalent for a convergent sequence of graphs \((H_n)_{n\in\mathbb{N}}\).

i. There exist \(c > 0\) and sets \(U_n \subseteq V(H_n)\) such that \(|U_n| \geq c \cdot |H_n|\) for every \(n \in \mathbb{N}\) and \(\lim_{n \to \infty} p(\rho, H_n|_{U_n}) \in \{0, 1\}\).

ii. There exist a subsequence \((H_{n_t})_{t\in\mathbb{N}}\) of \((H_n)_{n\in\mathbb{N}}\) and sets \(U_{n_t} \subseteq V(H_{n_t})\) such that \(\limsup_{t\to\infty}|U_{n_t}|/|H_{n_t}| > 0\) and \((H_{n_t}|_{U_{n_t}})_{t\in\mathbb{N}}\) is almost stable.

Proof. The implication (i) \(\implies\) (ii) is trivial as \(\lim_{n \to \infty} p(\rho, H_n|_{U_n}) \in \{0, 1\}\) implies \((H_n|_{U_n})_{n\in\mathbb{N}}\) is almost stable.

For the implication (ii) \(\implies\) (i), fix any graphon \(W\) over some space \(\Omega = (X, \mathcal{A}, \mu)\) that is a limit of \((H_n)_{n\in\mathbb{N}}\) and note that since this sequence is convergent, every subsequence of \((H_n)_{n\in\mathbb{N}}\) converges to \(W\). By hypothesis, possibly passing to a further subsequence \((H_{n_t})_{t\in\mathbb{N}}\), there are sets \(U_{n_t} \subseteq V(H_{n_t})\) with \(\lim_{t\to\infty}|U_{n_t}|/|H_{n_t}| > 0\) such that \((H_{n_t}|_{U_{n_t}})_{t\in\mathbb{N}}\) is both almost stable and convergent and if \(\tilde{W}\) is a limit graphon of this sequence, then it is a stable subgraphon of \(W\). By Theorem 3.6, we conclude that there exists a positive measure set \(U \subseteq X\) such that \(W|_{U\times U} = 1\) a.e. or \(W|_{U\times U} = 0\) a.e.; we let \(b \in \{0, 1\}\) be the a.e. value of \(W|_{U\times U}\).

Let \(c \overset{\text{def}}{=} \mu(U)\). We claim now that if \(W'\) is another graphon over some space \(\Omega' = (X', \mathcal{A}', \mu')\) that is a limit of \((H_n)_{n\in\mathbb{N}}\), then there exists a positive measure set \(U' \subseteq X'\) such that \(W'|_{U'\times U'} = b\) a.e. and \(\mu'(U') \geq c\).

This is completely trivial from the Graphon Uniqueness Theorem [BCL10] (see also [Lov12, Theorem 13.10]), but we offer an ad hoc proof here: by Lemma 3.3 applied to \(W\) and the indicator function \(1_U\), there exist a sequence of graphs \((H'_n)_{n\in\mathbb{N}}\) converging to \(W\) and sets \(U'_n \subseteq V(H'_n)\) with \(\lim_{n \to \infty}|U'_n|/|H'_n| = c\) and \((H'_n|_{U'_n})_{n\in\mathbb{N}}\) converging to the constant \(b\) graphon. Since \((H'_n)_{n\in\mathbb{N}}\) also converges to \(W'\), by the same lemma, we get a measurable function \(f : X' \to [0, 1]\) with \(\int_X f \, d\mu' = c\) such that the graphon \(W_f\) over \(\Omega_f\) given by \(W_f(x, y) = W'(x, y)\) is \(\mu_f\)-a.e. equal to \(b\). Taking \(U' \overset{\text{def}}{=} \{x \in X' \mid f(x) > 0\}\) gives \(\mu'(U') \geq c\) and \(W'|_{U'\times U'} = b\) \(\mu'\)-a.e.

Therefore, we have shown that there exists \(c > 0\) such that every graphon \(W\) over some space \(\Omega = (X, \mathcal{A}, \mu)\) that is a limit of \((H_n)_{n\in\mathbb{N}}\) has a measurable set \(U \subseteq X\) such that \(\mu(U) \geq c\) and \(W|_{U\times U} = b\) a.e.

For each \(n \in \mathbb{N}\), let \(U_n^c \subseteq V(H_n)\) be a set that minimizes \(|p(\rho, H_n|_{U_n}) - b|\) over all possible sets \(U_n \subseteq V(H_n)\) such that \(|U_n| \geq (c/2) \cdot |H_n|\). To conclude the proof, it is sufficient to show that \(\lim_{n \to \infty} p(\rho, H_n|_{U_n^c}) = b\).
Suppose not. Then there exists a subsequence \((H_{m_\ell})_{\ell \in \mathbb{N}}\) of \((H_n)_{n \in \mathbb{N}}\) such that \(\lim_{\ell \to \infty} |p(\rho, H_{m_\ell}|_{U_{m_\ell}}) - b| > 0\) and by possibly passing to a further subsequence, we can also assume that \((H_{m_\ell}|_{U_{m_\ell}})_{\ell \in \mathbb{N}}\) is convergent.

We now let \(H \overset{\text{def}}{=} \prod_{\ell \in \mathbb{N}} H_{m_\ell} / D\) for some non-principal ultrafilter \(D\) over \(\mathbb{N}\) and let \(\Theta: \prod_{\ell \in \mathbb{N}} V(H_{m_\ell})^2 / D \to [0, 1]^{(2)}\) be a separable realization of order 2 and \(N\) be as in Theorem 4.1. We also let \(\Theta_1\) be a restriction of \(\Theta\) of order 1, let \(W'\) be the graphon over \([0, 1]\) defined by

\[
W'(x, y) \overset{\text{def}}{=} \lambda(\{z \in [0, 1] \mid (x, y, z) \in N\})
\]

and per our previous claim, let \(U' \subseteq [0, 1]\) be such that \(\lambda(U') \geq c\) and \(W'|_{U' \times U'} = b\) a.e. We define further \(\tilde{U} \overset{\text{def}}{=} \Theta^{-1}_1(U') \subseteq \prod_{\ell \in \mathbb{N}} V(H_{m_\ell}) / D\) and since \(\Theta_1\) is measure-preserving, it follows that \(\mu^1(\tilde{U}) \geq c\) for the Loeb measure \(\mu^1\).

Consider now the graph \(H|\tilde{U}\) and note that \(E^{H|\tilde{U}} = \Theta^{-1}(N \cap (U' \times U' \times [0, 1]))\) a.e. and since \(W'|_{U' \times U'} = b\) a.e., it follows that \(\mu^2(E^{H|\tilde{U}}) = b \cdot \mu^2(\tilde{U} \times \tilde{U}) = b \cdot \mu^1(\tilde{U})^2\), where the last equality follows from Fubini’s Theorem for Loeb measures, Theorem A.2. Let now \(U = \prod_{\ell \in \mathbb{N}} U_{m_\ell} / D\) be an internal set such that \(\mu^1(U \triangle \tilde{U}) = 0\). By Fubini’s Theorem again, it follows that \(\mu^2(E^{H|U}) = b \cdot \mu^1(U)^2\) so we must have

\[
\lim_{\ell \to D} \frac{|U_{m_\ell}|}{|H_{m_\ell}|} = \mu^1(U) \geq c,
\]

\[
\lim_{\ell \to D} p(\rho, H_{m_\ell}|_{U_{m_\ell}}) = \frac{\mu^2(E^{H|U})}{\mu^1(U)^2} = b.
\]

However, this is a contradiction because it implies that along some subsequence we have \(|U_{m_\ell}| / |H_{m_\ell}| \geq c/2\) and \(p(\rho, H_{m_\ell}|_{U_{m_\ell}}) \to b\) contradicting the fact that the former implies \(|p(\rho, H_{m_\ell}|_{U_{m_\ell}}) - b| \geq |p(\rho, H_{m_\ell}|_{U_{m_\ell}}) - b|\) and we have \(|p(\rho, H_{m_\ell}|_{U_{m_\ell}}) - b| \not\to 0\). □

Discussion 4.4. Note that the convergence condition in Theorem 4.3 is necessary for a very simple reason: if we take a sequence of increasing graphs that alternates between complete graphs (say, when \(n\) is even) and empty graphs (say, when \(n\) is odd), it is clearly not convergent and any linear-sized induced subgraph also alternates between almost clique or almost anti-clique.
Discussion 4.5. Naively, one might conjecture that if the sequence \((H_n)_{n \in \mathbb{N}}\) itself is almost stable and we know the order of its stability, say, we know that \(\lim_{n \to \infty} p(G, H_n) = 0\) for every finite graph \(G\) containing a half-graph of order \(\ell\), then one would be able to know bounds on the relative size \(c\) of the sets \(U_n\) depending only on \(\ell\). However, this is not the case since if \(H_{n,m}\) is the disjoint union of \(m\) cliques of size \(n\), then for each fixed \(m \in \mathbb{N}_+\), the sequence \((H_{n,m})_{n \in \mathbb{N}}\) is convergent, does not contain any half-graphs of order 2 and the maximum asymptotic relative size of an almost clique or almost anti-clique is \(1/m\).

This also shows the necessity of requiring almost cliques or almost anti-cliques as opposed to cliques or anti-cliques: the diagonal sequence \((H_{n,n})_{n \in \mathbb{N}}\) is convergent but the largest cliques or anti-cliques in \(H_{n,n}\) have size \(n = \sqrt{|H_{n,n}|}\). However, the edge density in the sequence \((H_{n,n})_{n \in \mathbb{N}}\) itself goes to zero so it is an almost anti-clique.

Discussion 4.6. A posteriori, the example of Discussion 4.4 shows that we cannot get Theorem 4.3 by simply applying the removal lemma followed by the stable regularity lemma [MS14, AFP18, MS21] to each of the \(H_{n,\ell}|_{U_n}\), with a precision \(\epsilon_{n_t} > 0\) as such argument does not use the required property of convergence of the sequence \((H_n)_{n \in \mathbb{N}}\) in any way. The reason why the stable regularity lemma is not enough is that when applied to \(\epsilon > 0\), it provides some \(c(\epsilon) > 0\) such that every sufficiently large \(H\) has some set \(U \subseteq V(H)\) of size at least \(c(\epsilon) \cdot |H|\) that has edge density either at least \(1 - \epsilon\) or at most \(\epsilon\). However, to obtain the almost clique or almost anti-clique, we need to make \(\epsilon_{n_t} \to 0\) which also destroys our guaranteed lower bound on the relative size of the sets: \(c(\epsilon_{n_t}) \to 0\).

We now proceed to transfer the stability dichotomy to countable graphs.

Theorem 4.7. The following are equivalent for a graph \(G\) with \(V(G) = \mathbb{N}_+\).

i. There exist a set \(U \subseteq \mathbb{N}_+\) and an increasing sequence \((n_t)_{t \in \mathbb{N}}\) of positive integers such that \(\lim_{\ell \to \infty} p(\rho, G|_{U \cap \{n_t\}}) \in \{0, 1\}\) and \(\lim_{\ell \to \infty} |U \cap \{n_t\}|/n_t > 0\).

ii. There exist a set \(U \subseteq \mathbb{N}_+\) and an increasing sequence \((n_t)_{t \in \mathbb{N}}\) of positive integers such that \((G|_{U \cap \{n_t\}})_{t \in \mathbb{N}}\) is almost stable and \(\lim_{\ell \to \infty} |U \cap \{n_t\}|/n_t > 0\).

Proof. The implication (i) \(\implies\) (ii) is trivial as \(\lim_{\ell \to \infty} p(\rho, G|_{U \cap \{n_t\}}) \in \{0, 1\}\) implies \((G|_{U \cap \{n_t\}})_{t \in \mathbb{N}}\) is almost stable.
For the implication (ii) \(\implies\) (i), by possibly passing to a subsequence of \((n_\ell)_{\ell\in\mathbb{N}}\), we may further assume that \((G|_{[n_\ell]})_{\ell\in\mathbb{N}}\) is convergent, so by Theorem 4.3, there exist \(c > 0\) and sets \(U_\ell \subseteq [n_\ell]\) such that \(|U_\ell| \geq c \cdot n_\ell\) for every \(\ell \in \mathbb{N}\) and \(\lim_{\ell \to \infty} p(\rho, G|_{U_\ell}) \in \{0, 1\}\).

Define then the sequence \((m_t)_{t\in\mathbb{N}}\) recursively by

\[
m_0 \overset{\text{def}}{=} n_0, \quad m_{t+1} \overset{\text{def}}{=} \min\{n_\ell \mid \ell \in \mathbb{N} \wedge n_\ell \geq 2^t \cdot m_t\}
\]

and for each \(t \in \mathbb{N}\), let \(\ell_t \in \mathbb{N}\) be such that \(m_t = n_{\ell_t}\). Let also

\[
U \overset{\text{def}}{=} \bigcup_{t\in\mathbb{N}} U_{\ell_t} \cap ([m_t] \setminus [m_{t-1}]),
\]

where \(m_{-1} \overset{\text{def}}{=} 0\).

Note that

\[
|(U \cap [m_t]) \triangle U_{\ell_t}| \leq m_{t-1} \leq 2^{t+1} \cdot |U_{\ell_t}|,
\]

hence \(\lim_{t \to \infty} |U \cap [m_t]|/|U_{\ell_t}| = 1\), which implies that

\[
\liminf_{t \to \infty} \frac{|U \cap [m_t]|}{m_t} = \liminf_{t \to \infty} \frac{|U_{\ell_t}|}{n_{\ell_t}} \geq c > 0,
\]

and

\[
\lim_{t \to \infty} p(\rho, G|_{U \cap [m_t]}) = \lim_{t \to \infty} p(\rho, G|_{U_{\ell_t}}) \in \{0, 1\},
\]

completing the proof. \(\square\)

Discussion 4.8. One might naïvely hope that in the countable case one would be able to produce an almost clique or almost anti-clique \(U\) of positive density (as opposed to positive upper density as in Theorem 4.7), but a simple counter-example shows this is not possible: if \(G\) is the graph over \(\mathbb{N}_+\) with edge set

\[
E(G) \overset{\text{def}}{=} \{\{v, w\} \mid \lfloor \sqrt{\log_2(v)} \rfloor \equiv \lfloor \sqrt{\log_2(w)} \rfloor \equiv 0 \pmod{2}\}
\]

then \(G\) is stable (as it is a union of cliques) and does not have any positive density almost clique or anti-clique simply because for each \(\epsilon > 0\) and each \(n_0 \in \mathbb{N}_+\), there exist \(n, n' \geq n_0\) such that the edge density of the marginals \(G|_{[n]}\) and \(G|_{[n']}\) are at most \(\epsilon\) away from 0 and 1, respectively.
5 Trivial sub-objects in theons

In this section, we state and prove the stability dichotomy theorem for theons, Theorem 5.11, which is a generalization of Theorem 3.6 of Section 3 for universal theories over finite relational languages. Before we do so, let us give a gentle introduction to the theories of flag algebras [Raz07] and theons [CR20b], which generalize the theory of graphons to universal theories.

First, given finite models $M$ and $N$ of a universal theory $T$ over a finite relational language $L$, we let $T_{\text{ind}}(M,N)$ be the set of all model embeddings of $M$ in $N$ (i.e., injective functions $f: V(M) \rightarrow V(N)$ that preserve all relations and negations of relations) and let

$$t_{\text{ind}}(M,N) \equiv \frac{|T_{\text{ind}}(M,N)|}{(|N|)|M|}$$

be the normalized number of embeddings of $M$ in $N$, called the labeled (induced) density of $M$ in $N$. We also define the (induced) density of $M$ in $N$ as the normalized number of induced submodels of $N$ that are isomorphic to $M$ given by

$$p(M,N) \equiv \frac{|\{U \subseteq V(N) \mid |N|_U \cong M\}|}{(|N|)!} = \frac{|M|!}{|\text{Aut}(M)|} \cdot t_{\text{ind}}(M,N),$$

where $\text{Aut}(M)$ is the group of automorphisms of $M$. For each $n \in \mathbb{N}$, we let $\mathcal{M}_n[T]$ be the set of models of $T$ of size $n$ up to isomorphism and we let $^3 \mathcal{M}[T] = \bigcup_{n \in \mathbb{N}} \mathcal{M}_n[T]$. Note that the fact that $T$ is a universal theory implies that $\mathcal{M}[T]$ is closed under substructures, which in turn implies that for every $N \in \mathcal{M}[T]$ and every $n \leq |N|$, we have $\sum_{M \in \mathcal{M}_n[T]} p(M,N) = 1$. For a set $V$, we also let $\mathcal{K}_V[T]$ be the set of models of $T$ whose vertex set is $V$ (we do not factor isomorphisms for $\mathcal{K}_V[T]$).

Recall that for universal theories $T_1$ and $T_2$ over finite relational languages $\mathcal{L}_1$ and $\mathcal{L}_2$, respectively, an open interpretation (also known under the name definition) from $T_1$ to $T_2$ is a function $I$ (denoted $I: T_1 \rightarrow T_2$) that maps each predicate symbol $P \in \mathcal{L}_1$ to an open (i.e., quantifier-free) formula $I(P)(x_1, \ldots, x_{k(P)})$ of $\mathcal{L}_2$, where $k(P)$ is the arity of $P$ and such that

$^3$In the framework of limits, it is very convenient to assume that the vertex set of a structure/model can be empty and thus $\mathcal{M}_0[T]$ is included in this union.
for each axiom $\forall \vec{x}, F(\vec{x})$ of $T_1$, we have $T_2 \vdash \forall \vec{x}, I(F)(\vec{x})$ when we declare $I$ to commute with logical connectives. An open interpretation $I: T_1 \sim T_2$ contra-variantly naturally defines maps $K_V[T_2] \to K_V[T_1]$ for each set $V$ given by $(I(M) \models P(\vec{x})) \iff (M \models I(P)(\vec{x}))$ for each $P \in L_1$.

Two open interpretations $I_1, I_2: T_1 \sim T_2$ are equivalent if $T_2 \vdash \forall \vec{x}, (I_1(P)(\vec{x}) \leftrightarrow I_2(P)(\vec{x}))$ for every predicate symbol $P \in L_1$. Equivalently, the open interpretations $I_1, I_2: T_1 \sim T_2$ are equivalent if they induce the same maps $K_V[T_2] \to K_V[T_1]$ for every set $V$ (in fact, it is enough to know that this is true for $V = [k]$, where $k$ is the maximum arity of a predicate of $T_1$). We let $\text{INT}$ be the category of universal theories in finite relational languages whose morphisms are open interpretations up to equivalence. Note that if $I: T_1 \sim T_2$ is an isomorphism of $\text{INT}$, then $p(I(M), I(N)) = p(M, N)$ for every $M, N \in M[T_2]$, which means that isomorphic theories are indistinguishable for the purposes of densities of submodels. Isomorphisms in the category $\text{INT}$ are also known under the name interdefinitions.

It will be more convenient to work with canonical theories, which are theories in which every relation contains only injective tuples, that is, theories that entail

$$\forall x_1, \ldots, x_{k(P)}, \left( \bigvee_{1 \leq i < j \leq k(P)} x_i = x_j \rightarrow \neg P(x_1, \ldots, x_{k(P)}) \right)$$

(4)

for every predicate symbol $P$. By [CR20b, Theorem 2.3] (see also [AC14, §2.2]), every universal theory is isomorphic (in $\text{INT}$) to a canonical theory. From this point forward, unless explicitly mentioned otherwise, all theories are assumed to be canonical. For a finite relational language, we let $T_L$ be the pure canonical theory over $L$, that is, the theory whose axioms are precisely the ones in (4) for each $P \in L$; the models of $T_L$ are sometimes referred to as canonical structures in $L$.

The notion of convergence is now defined analogously to the graph case: a sequence $(N_n)_{n \in \mathbb{N}}$ of finite models of a canonical theory $T$ is called convergent if it is increasing in the sense that for every $n \in \mathbb{N}$, $|N_n| < |N_{n+1}|$ and if for every $M \in M[T]$, the limit $\lim_{n \to \infty} p(M, N_n)$ exists. Again, another way of seeing this is as convergence in the (compact and metrizable) product topology of $[0, 1]^{M[T]}$ of the sequence $(p(-, N_n))_{n \in \mathbb{N}}$.

The simplest way of encoding the limit of a convergent sequence $(N_n)_{n \in \mathbb{N}}$ is syntactically/algebraically by defining $\phi \in [0, 1]^{M[T]}$ by $\phi(M) \overset{\text{def}}{=} \lim_{n \to \infty} p(M, N_n)$. 

27
The theory of flag algebras then describes which points of \([0, 1]^{\mathcal{M}[T]}\) can arise as limits of convergent sequences. It turns out that this description boils down to some polynomial restrictions plus a positivity condition. Namely, let \(\mathbb{R}\mathcal{M}[T]\) be the vector space of formal \(\mathbb{R}\)-linear combinations of elements of \(\mathcal{M}[T]\). We then extend each \(\phi \in [0, 1]^{\mathcal{M}[T]}\) linearly to a function \(\phi : \mathbb{R}\mathcal{M}[T] \to \mathbb{R}\) (which we denote by abuse with the same letter) as

\[
\phi \left( \sum_{M \in \mathcal{M}[T]} c_M M \right) \overset{\text{def}}{=} \sum_{M \in \mathcal{M}[T]} c_M \phi(M).
\]

Let \(\mathcal{K}[T]\) be the linear subspace of \(\mathbb{R}\mathcal{M}[T]\) spanned by elements of the form

\[
M - \sum_{N \in \mathcal{M}_n[T]} p(M, N) N
\]

for \(n \geq |M|\) and let \(\mathcal{A}[T] \overset{\text{def}}{=} \mathbb{R}\mathcal{M}[T]/\mathcal{K}[T]\). It is straightforward to see that if \(\phi = \lim_{n \to \infty} p(-, N_n)\) for some convergent sequence \((N_n)_{n \in \mathbb{N}}\), then \(\mathcal{K}[T] \subseteq \ker(\phi)\), which means that we can think of \(\phi\) instead as a linear map \(\phi : \mathcal{A}[T] \to \mathbb{R}\) by factoring out \(\mathcal{K}[T]\). It turns out that \(\mathcal{A}[T]\) becomes an \(\mathbb{R}\)-algebra when equipped with the (bilinear) product operation defined by

\[
M_1 \cdot M_2 \overset{\text{def}}{=} \sum_{M \in \mathcal{M}_n[T]} p(M_1, M_2; M) M,
\]

for \(n \geq |M_1| + |M_2|\), where

\[
p(M_1, M_2; M) \overset{\text{def}}{=} \left\{ \left( U_1, U_2 \right) \in 2^{V(M)} \times 2^{V(M)} \mid M |_{U_1} \cong M_1 \wedge M |_{U_2} \cong M_2 \wedge U_1 \cap U_2 = \emptyset \right\} /\binom{|M|}{|M_1|} \binom{|M|-|M_1|}{|M_2|},
\]

and the unit of \(\mathcal{A}[T]\) is the equivalence class of the element \(\sum_{M \in \mathcal{M}_n[T]} M\) for any given \(n \in \mathbb{N}\). Furthermore, any \(\phi\) coming from a convergent sequence respects this product operation, in other words, \(\phi\) is necessarily in the set \(\text{Hom}(\mathcal{A}[T], \mathbb{R})\) of \(\mathbb{R}\)-algebra homomorphisms from \(\mathcal{A}[T]\) to \(\mathbb{R}\). In fact, by letting

\[
\text{Hom}^+(\mathcal{A}[T], \mathbb{R}) \overset{\text{def}}{=} \{ \phi \in \text{Hom}(\mathcal{A}[T], \mathbb{R}) \mid \forall M \in \mathcal{M}[T], \phi(M) \geq 0 \}
\]
be the set of positive homomorphisms, any $\phi$ coming from a convergent sequence is necessarily a positive homomorphism. The main theorem below of flag algebra theory (sometimes referred to as Existence Theorem for flag algebras) says that in fact the set $\text{Hom}^+(A[T], \mathbb{R})$ is precisely the set of all limits of convergent sequences.

**Theorem 5.1** (Lovász–Szegedy [LS06], Razborov [Raz07]). Let $T$ be a universal theory in a finite relational language.

If $(N_n)_{n \in \mathbb{N}}$ is a convergent sequence of finite models of $T$, then there exists $\phi \in \text{Hom}^+(A[T], \mathbb{R})$ such that $\lim_{n \to \infty} p(M, N_n) = \phi(M)$ for every $M \in \mathcal{M}[T]$. Conversely, if $\phi \in \text{Hom}^+(A[T], \mathbb{R})$, then there exists a convergent sequence $(N_n)_{n \in \mathbb{N}}$ of $T$ such that $\lim_{n \to \infty} p(M, N_n) = \phi(M)$ for every $M \in \mathcal{M}[T]$. 

Note that because of the minimalist nature of the flag algebraic description, uniqueness here is obvious: $\phi_1, \phi_2 \in \text{Hom}^+(A[T], \mathbb{R})$ represent the limit of the same convergent sequence $(N_n)_{n \in \mathbb{N}}$ if and only if $\phi_1 = \phi_2$. For this reason, it is very convenient to use the set $\text{Hom}^+(A[T], \mathbb{R})$ when talking about limits of finite models of the theory $T$.

For a semantic/geometric description of the limit objects we use the theory of theons [CR20b], which generalizes the theory of graphons to describe limits of finite models of canonical theories. 

Given an atomless standard probability space $\Omega = (X, \mathcal{A}, \mu)$ and a set $V$, let

$$E_V(\Omega) \overset{\text{def}}{=} X^r(V)$$

and equip it with the completion of the product measure, which by abuse we also denote by $\mu$. We also define the *diagonal set* as (cf. Equation (3))

$$D_V(\Omega) \overset{\text{def}}{=} \{ x \in E_V(\Omega) \mid \exists v, w \in V, (v \neq w \land x_v = x_w) \}.$$ 

Clearly, the diagonal has zero-measure (and this is precisely the reason why we need to work with canonical theories so that no information is lost). Again, we will typically take $\Omega$ to be $[0, 1]$ equipped with the Lebesgue measure over Borel sets and in this case, we will omit $\Omega$ from the notation.

We will also be abusing the notation slightly by identifying the spaces $E_V(\Omega \times \Omega)$ and $E_V(\Omega) \times E_V(\Omega)$ naturally via the correspondence $E_V(\Omega \times \Omega) \ni x \leftrightarrow (y, z) \in E_V(\Omega) \times E_V(\Omega)$ given by $y_A \overset{\text{def}}{=} (x_A)_1$ and $z_A \overset{\text{def}}{=} (x_A)_2$. 

29
For a predicate symbol $P$, a $P$-on over $\Omega$ is a measurable subset of $\mathcal{E}_{k(P)}(\Omega)$, where $k(P)$ is the arity of $P$. An Euclidean structure in a finite relational language $\mathcal{L}$ over $\Omega$ is a function $\mathcal{N}$ that maps each predicate symbol $P \in \mathcal{L}$ to a $P$-on $\mathcal{N}_P \subseteq \mathcal{E}_{k(P)}(\Omega)$.

Analogously to the way that solution sets are defined, given an open formula $F(x_1, \ldots, x_n)$ in $\mathcal{L}$ and an Euclidean structure $\mathcal{N}$ in $\mathcal{L}$ over $\Omega$, the truth set $T(F, \mathcal{N}) \subseteq \mathcal{E}_n(\Omega)$ of $F$ is defined by

i. $T(x_i = x_i, \mathcal{N}) \overset{\text{def}}{=} \mathcal{E}_n(\Omega)$.

ii. $T(x_i = x_j, \mathcal{N}) \overset{\text{def}}{=} \emptyset$, if $i \neq j$.

iii. $T(P(x_{i_1}, \ldots, x_{i_k}(P)), \mathcal{N}) \overset{\text{def}}{=} \emptyset$, if $i: [k(P)] \rightarrow [n]$ is not injective.

iv. $T(P(x_{i_1}, \ldots, x_{i_k}(P)), \mathcal{N}) \overset{\text{def}}{=} (i^*)^{-1}(\mathcal{N}_P)$, if $i: [k(P)] \rightarrow [n]$ is injective (recall that $i^* : \mathcal{E}_n(\Omega) \rightarrow \mathcal{E}_{k(P)}(\Omega)$ is given by $i^*(x)_A \overset{\text{def}}{=} x_{i(A)}$).

v. $T(\neg, \mathcal{N})$ commutes with logical connectives (so, e.g., $T(\neg F, \mathcal{N}) \overset{\text{def}}{=} \mathcal{E}_n(\Omega) \setminus T(F, \mathcal{N})$ and $T(F_1 \land F_2, \mathcal{N}) \overset{\text{def}}{=} T(F_1, \mathcal{N}) \cap T(F_2, \mathcal{N})$).

One might complain that items (ii) and (iii) above should not be defined as the empty set but rather as particular subsets of the diagonal $D_n(\Omega)$, but since all information on the diagonal will be lost regardless, the definition uses the empty set for simplicity.

Truth sets allow us to define the set of copies of a finite canonical structure $M$ as follows: if $V(M) \overset{\text{def}}{=} [n]$, then the open diagram $D_{\text{open}}(M)(x_1, \ldots, x_n)$ of $M$ is the conjunction of all formulas of the form

$$x_i \neq x_j \quad \text{with } i \neq j,$$

$$P(x_{i_1}, \ldots, x_{i_k}) \quad \text{with } M \models P(i_1, \ldots, i_k),$$

$$\neg P(x_{i_1}, \ldots, x_{i_k}) \quad \text{with } M \models \neg P(i_1, \ldots, i_k).$$

Equivalently, it is the open formula that completely encodes the quantifier-free type (over the empty set) of the tuple $(1, \ldots, n)$ in $M$ (recall that the language is finite). The set of labeled (induced) copies of $M$ in $\mathcal{N}$ is defined as $T_{\text{ind}}(M, \mathcal{N}) \overset{\text{def}}{=} T(D_{\text{open}}(M), \mathcal{N})$. If the vertex set $V(M)$ of $M$ is not $[n]$,
then we simply relabel its vertices with a bijection $\alpha : V(M) \mapsto [n]$, where $n \overset{\text{def}}{=} |M|$ to get a canonical structure $N$ with vertex set $[n]$ such that $$(M \models P(v_1, \ldots, v_n)) \iff (N \models P(\alpha(v_1), \ldots, \alpha(v_k(P))))$$
and define $T_{\text{ind}}(M, N) \overset{\text{def}}{=} \alpha^*(T_{\text{ind}}(N, N)) \subseteq E_{V(M)}(\Omega)$ (it is easy to see that this does not depend on the choice of $\alpha$). The labeled (induced) density and the (induced) density of $M$ in $N$ are then defined respectively as

$$t_{\text{ind}}(M, N) \overset{\text{def}}{=} \mu(T_{\text{ind}}(M, N)), \quad \phi_N(M) \overset{\text{def}}{=} \frac{|M|!}{|\text{Aut}(M)|} \cdot t_{\text{ind}}(M, N).$$

Finally, given a canonical theory $T$ over $\mathcal{L}$ and an Euclidean structure $\mathcal{N}$ in $\mathcal{L}$ over $\Omega$, we say that $\mathcal{N}$ is a (weak) $T$-on if every $\mathcal{L}$-structure $M$ that is not a model of $T$ satisfies $t_{\text{ind}}(M, \mathcal{N}) = 0$ and we say that $\mathcal{N}$ is a strong $T$-on if every $\mathcal{L}$-structure $M$ that is not a model of $T$ satisfies $T_{\text{ind}}(M, \mathcal{N}) \subseteq D_{V(M)}(\Omega)$. We say that a weak or strong $T$-on $\mathcal{N}$ is a limit of a convergent sequence $(N_n)_{n \in \mathbb{N}}$ of models of $T$ if $\lim_{n \to \infty} p(M, N_n) = \phi_N(M)$ for every model $M$ of $T$ (see Theorem 7.8 below for an equivalent definition in terms of axioms of $T$).

The main theorem of the theory of theons is naturally the Existence Theorem for theons below.

**Theorem 5.2** ([CR20b, Theorem 3.4], see also [AC14, §3.1]). Let $T$ be a canonical universal theory in a finite relational language and $\Omega$ be an atomless probability space. If $(N_n)_{n \in \mathbb{N}}$ is a convergent sequence of models of $T$, then there exists a $T$-on $\mathcal{N}$ over $\Omega$ that is a limit of $(N_n)_{n \in \mathbb{N}}$. Conversely, every $T$-on over $\Omega$ is a limit of a convergent sequence of models of $T$.

**Remark 5.3.** If $T_{\text{Graph}}$ is the theory of graphs, then a $T_{\text{Graph}}$-on $\mathcal{N}$ is not exactly the same object as a graphon $W$, but there is a (not one-to-one) correspondence preserving densities of finite graphs given by

$$W_{\mathcal{N}}(x, y) \overset{\text{def}}{=} \lambda(\{z \in [0, 1] \mid (x, y, z) \in \mathcal{N}\}),$$

$$\mathcal{N} \overset{\text{def}}{=} \{x \in \mathcal{E}_2 \mid x_{1,2} < W(x_{1}, x_{2})\}.$$
Theorem 5.4 (Induced Euclidean Removal Lemma [CR20b, Theorem 3.3]). If $\mathcal{N}$ is a $T$-on over $\Omega = (X, \mathcal{A}, \mu)$, then there exists a strong $T$-on $\mathcal{N}'$ over $\Omega$ such that $\mu(\mathcal{N}_P \triangle \mathcal{N}'_P) = 0$ for every predicate symbol $P$.

Remark 5.5. Theorem 5.4 above can also be used to ensure that all open formulas that are false a.e. in $\mathcal{N}$ become false everywhere off-diagonal in $\mathcal{N}'$. Namely, given a $T$-on $\mathcal{N}$ over $\Omega = (X, \mathcal{A}, \mu)$, we let $\text{Th}(\varphi_\mathcal{N})$ be the canonical theory whose finite models are precisely those $\mathcal{M}$ such that $\varphi_\mathcal{N}(\mathcal{M}) > 0$. Note that $\mathcal{N}$ is also a (weak) $\text{Th}(\varphi_\mathcal{N})$-on, so we can apply Theorem 5.4 above to get a strong $\text{Th}(\varphi_\mathcal{N})$-on $\mathcal{N}'$ whose peons differ from those of $\mathcal{N}$ only by zero-measure sets. If $\mu(T(F, \mathcal{N})) = 0$ for some open formula $F(x_1, \ldots, x_n)$, then for any $\mathcal{L}$-structure $\mathcal{M}$ with $V(\mathcal{M}) = [n]$ and $\mathcal{M} \models F(1, \ldots, n)$, we must have $t_{\text{ind}}(\mathcal{M}, \mathcal{N}) = 0$ and thus $T_{\text{ind}}(\mathcal{M}, \mathcal{N}') \subseteq D_n(\Omega)$, which in turn implies $T(F, \mathcal{N}') \subseteq D_n(\Omega)$.

As expected from the graphon case, the same convergent sequence can converge to different theons and this is completely characterized by the Theon Uniqueness Theorem [CR20b, Theorems 3.9 and 3.11 and Proposition 7.7], which has a very technical statement. Fortunately, we will only need a consequence of it concerning open interpretations, Proposition 5.6 below. But before we state it, we need some preliminary definitions and properties.

First, open interpretations behave naturally with respect to convergence: it is not hard to see that if $I : T_1 \twoheadrightarrow T_2$ is an open interpretation and $(\mathcal{N}_n)_{n \in \mathbb{N}}$ is a convergent sequence of models of $T_2$, then $(I(\mathcal{N}_n))_{n \in \mathbb{N}}$ is a convergent sequence of models of $T_1$. It turns out that there are natural operations that encode this operation for limit objects. Namely, for flag algebras, Razborov [Raz07, Theorem 2.6] showed that the linear map $\pi^I : \mathcal{A}[T_1] \to \mathcal{A}[T_2]$ given by

$$\pi^I(M) = \sum \{M' \in \mathcal{M}_{|M'}[T_2] \mid I(M') \cong M\}$$

is an $\mathbb{R}$-algebra homomorphism and if $\phi$ is the limit of $(\mathcal{N}_n)_{n \in \mathbb{N}}$, then the composition $\phi_I^T \overset{\text{def}}{=} \phi \circ \pi^I \in \text{Hom}^+(\mathcal{A}[T_1], \mathbb{R})$ is the limit of $(I(\mathcal{N}_n))_{n \in \mathbb{N}}$. For theons [CR20b, Remark 6], if $\mathcal{N}$ is a $T_2$-on that is the limit of $(\mathcal{N}_n)_{n \in \mathbb{N}}$, then the $T_1$-on $I(\mathcal{N})$ defined via truth sets by $I(\mathcal{N})_P \overset{\text{def}}{=} T(P, \mathcal{N})$ for every predicate symbol $P$ is the limit of $(I(\mathcal{N}_n))_{n \in \mathbb{N}}$. We can combine these results neatly as $\phi(I(\mathcal{N})) = \phi_I^\mathcal{N}$, or in plain English, the limit encoded by $I(\mathcal{N})$ is the same as the interpreted limit of $\mathcal{N}$ via $I$.

One natural question that arises is whether theons can be lifted through open interpretations in the following sense: if $\mathcal{N}$ is a $T_1$-on and $\phi \in \text{Hom}^+(\mathcal{A}[T_2], \mathbb{R})$
is such that $\phi^I = \phi_N$ for some open interpretation $I : T_1 \rightsquigarrow T_2$, then is there a $T_2$-on $\mathcal{H}$ such that $\phi_\mathcal{H} = \phi$ and $I(\mathcal{H}) = \mathcal{N}$ a.e.? In plain English, if $\mathcal{N}$ encodes the limit $\phi^I$, then is it of the form $\mathcal{N} = I(\mathcal{H})$ a.e. for some limit $\mathcal{H}$ encoding $\phi$?

While the answer to this question is no (see [CR20b, Example 45]), the following proposition says that if we allow ourselves to add "dummy variables", the answer becomes yes.

**Proposition 5.6** ([CR20a, Proposition 4.3]). Let $I : T_1 \rightsquigarrow T_2$ be an open interpretation, let $\phi \in \text{Hom}^+(\mathcal{A}[T_2], \mathbb{R})$ and let $\mathcal{N}$ be a $T_1$-on over $\Omega$ such that $\phi^I = \phi_\mathcal{N}$. Then there exists a $T_2$-on $\mathcal{H}$ over $\Omega \times \Omega$ such that $\phi_\mathcal{H} = \phi$ and $I(\mathcal{H})_P = \mathcal{N}_P \times \mathcal{E}_{k(P)}(\Omega)$ a.e. for every predicate symbol $P$ in the language of $T_1$.

We now define limit sub-objects in analogy to subgraphons.

**Definition 5.7.** Given a limit object $\phi \in \text{Hom}^+(\mathcal{A}[T], \mathbb{R})$, a (positive measure) limit sub-object of $\phi$ is a limit object $\psi \in \text{Hom}^+(\mathcal{A}[T], \mathbb{R})$ such that there exists a sequence $(N_n)_{n \in \mathbb{N}}$ converging to $\phi$ and sets $U_n \subseteq V(H_n)$ such that $(N_n| U_n)_{n \in \mathbb{N}}$ converges to $\psi$ and $\lim_{n \to \infty} |U_n|/|N_n| > 0$; when we want to be more specific, for $c \overset{\text{def}}{=} \lim_{n \to \infty} |U_n|/|N_n| > 0$ we say that $\psi$ is a measure $c$ limit sub-object of $\phi$.

Similarly to the graphon case, if $\mathcal{N}$ is a $T$-on over $\Omega = (X, \mathcal{A}, \mu)$ with $\phi_\mathcal{N} = \phi$, then not every sub-object of $\phi$ can be represented by conditioning the vertex variables $x_{\{v\}}$ to be in some positive measure set $U \subseteq X$.

More precisely, given a $T$-on $\mathcal{N}$ over a space $\Omega = (X, \mathcal{A}, \mu)$ and a positive measure set $U \subseteq X$, we let $\mu_U$ be the measure over $(X, \mathcal{A})$ defined by $\mu_U(A) \overset{\text{def}}{=} \mu(A \cap U)/\mu(U)$ and for a measure-isomorphism $F$ modulo 0 from $\Omega_U \overset{\text{def}}{=} (X, \mathcal{A}, \mu_U)$ to $\Omega$, we let $\mathcal{N}_U$ be the $T$-on over $\Omega_U$ defined by

$$ (\mathcal{N}_U)_P \overset{\text{def}}{=} \{ x \in \mathcal{E}_{k(P)}(\Omega_U) \mid x^F \in \mathcal{N}_P \}, $$

where

$$ x^F_a \overset{\text{def}}{=} \begin{cases} x_a, & \text{if } |A| = 1; \\ F(x_a), & \text{if } |A| \geq 2. \end{cases} \quad (5) $$

Under this definition, not every sub-object $\psi$ of $\phi_\mathcal{N}$ is of the form $\phi_{\mathcal{N}_U}$ for some choice of $(U, F)$ as above. However, just as in the graphon case, this
description is not far from correct, we only need to “rescale” the underlying measure by a weight function.

**Lemma 5.8.** Let \( \mathcal{N} \) be a \( T \)-on over \( \Omega = (X, \mathcal{A}, \mu) \), let \( c > 0 \) and let \( \psi \in \text{Hom}^+(\mathcal{A}[T], \mathbb{R}) \). The following are equivalent.

i. There exist a convergent sequence \( (N_n)_{n \in \mathbb{N}} \) converging to \( \phi_N \) and sets \( U_n \subseteq V(N_n) \) with \( \lim_{n \to \infty} |U_n|/|N_n| = c \) such that \( (N_n|_{U_n})_{n \in \mathbb{N}} \) converges to \( \psi \), that is, \( \psi \) is a measure \( c \) limit sub-object of \( \phi_N \).

ii. There exists a measurable function \( f : X \to [0, 1] \) with \( \int_X f \, d\mu = c \) such that for the space \( \Omega_f \) defined by

\[
\mu_f(A) \overset{\text{def}}{=} \frac{\int_A f(x) \, d\mu(x)}{c},
\]

there exists a measure-isomorphism \( F \mod 0 \) from \( \Omega_f \) to \( \Omega \).

iii. Item (ii) holds for every measure-isomorphism \( F \mod 0 \) from \( \Omega_f \) to \( \Omega \).

**Proof.** The implication (iii) \( \implies \) (ii) is trivial.

For the other implications, we will use the operator \( \pi^{(U, I)} \) of the theory of flag algebras [Raz07, Theorem 2.6]. Let \( \widehat{T} \) be the theory obtained from \( T \) by augmenting it with a unary predicate symbol \( U \) and for each \( n \in \mathbb{N} \), let

\[
\mathcal{M}_n[\widehat{T}] \overset{\text{def}}{=} \{ M \in \mathcal{M}_n[\widehat{T}] \mid M \models \forall x, U(x) \}
\]

be the set of all models of \( \widehat{T} \) of size \( n \) in which all vertices satisfy \( U \). Let \( u \overset{\text{def}}{=} \sum_{M \in \mathcal{M}_{[\widehat{T}]}^U} M \) and let \( \mathcal{A}_u[\widehat{T}] \) be the localization of \( \mathcal{A}[\widehat{T}] \) with respect to the multiplicative system \( \{ u^n \mid n \in \mathbb{N} \} \). Finally, let \( I : T \sim \widehat{T} \) be the structure-erasing interpretation that acts identically on \( T \). By [Raz07, Theorem 2.6], the linear map \( \pi^{(U, I)} : \mathcal{A}[T] \to \mathcal{A}_u[\widehat{T}] \) given by

\[
\pi^{(U, I)}(M) \overset{\text{def}}{=} \frac{M^U}{u^{[M]}},
\]

34
where \( M^U \in \mathcal{M}_\mathcal{M}[\hat{T}] \) is the model of \( \hat{T} \) obtained from \( M \) by declaring all its vertices to satisfy \( U \) is an \( \mathbb{R} \)-algebra homomorphism. The intuition is that if \( \psi \in \text{Hom}^+(\mathcal{A}[\hat{T}], \mathbb{R}) \) is such that \( \psi(u) > 0 \), then \( \psi \) has a non-negligible fraction of “vertices” satisfying \( U \) and the composition \( \psi \circ \pi^{(U,I)} \in \text{Hom}^+(\mathcal{A}[T], \mathbb{R}) \) is the limit object of \( T \) induced by the “vertices” of \( \psi \) satisfying \( U \) (hence the need for the localization).

Let us prove the implication (i) \( \implies \) (iii).

For each \( n \in \mathbb{N} \), let \( \hat{N}_n \) be the model of \( \hat{T} \) obtained from \( N_n \) by declaring the predicate symbol \( U \) to be true exactly in the set \( U_n \) and by possibly passing to a subsequence, we may suppose that \( (\hat{N}_n)_{n \in \mathbb{N}} \) converges to some homomorphism \( \xi \in \text{Hom}^+(\mathcal{A}[\hat{T}], \mathbb{R}) \). Note that since \( \lim_{n \to \infty} |U_n|/|N_n| = c \), we have \( \xi(u) = c \). Note further that since \( (N_{\mathcal{A}_n})_{n \in \mathbb{N}} \) and \( (N_{\mathcal{A}_n}|U_n)_{n \in \mathbb{N}} \) converge to \( \phi_N \) and \( \psi \), respectively and \( I(\hat{N}_n) = N_n \), we must have \( \phi_N = \xi I \) and \( \psi = \xi \circ \pi^{(U,I)} \). By Proposition 5.6, there exists a \( \hat{T} \)-on \( \hat{N} \) over \( \Omega \times \Omega \) such that \( \phi_{\hat{N}} = \xi \) and \( I(\hat{N})_P = N_P \times E_{k(P)}(\Omega) \) a.e., for every predicate symbol \( P \) in the language of \( T \). Since \( \hat{N}_U \subseteq E_1(\Omega) \times E_1(\Omega) \cong X \times X \), we can define the function \( f : X \to [0, 1] \) by

\[
f(x) \overset{\text{def}}{=} \mu(\{ y \in X \mid (x,y) \in \hat{N}_U \})
\]

(8)

(defined it arbitrarily when the set above is not measurable) and Fubini’s Theorem ensures that \( f \) is measurable.

Note also that

\[
c = \xi(u) = \sum_{M \in \mathcal{M}_{\mathcal{A}}[\hat{T}]} t_{\text{ind}}(M, \hat{N}) = \mu(T(U, \hat{N})) = \mu(\hat{N}_U) = \int_X f \, d\mu.
\]

(9)

Define \( \Omega_f \overset{\text{def}}{=} (X, \mathcal{A}, \mu_f) \) with \( \mu_f \) given by (6) and the \( T \)-on \( \mathcal{N}_f \) by (7) for an arbitrary measure-isomorphism \( F \) modulo 0 from \( \Omega_f \) to \( \Omega \) and note that for every \( M \in \mathcal{M}_{\mathcal{A}}[T] \), we have

\[
\phi_{\mathcal{N}_f}(M) = \frac{|M|!}{|\text{Aut}(M)|} \cdot t_{\text{ind}}(M, \mathcal{N}_f) = \frac{|M^U|!}{|\text{Aut}(M^U)|} \cdot \frac{t_{\text{ind}}(M^U, \hat{N})}{\mu(\hat{N}_U)^n}
\]

(10)

so \( \phi_{\mathcal{N}_f} = \psi \) as required.
For the final implication (ii) \(\implies\) (i), we define the \(\hat{T}\)-on \(\hat{\mathcal{N}}\) over \(\Omega \times \Omega\) from \(\mathcal{N}\) by letting \(\hat{\mathcal{N}}_P \stackrel{\text{def}}{=} \mathcal{N}_P \times \mathcal{E}_{k(P)}(\Omega)\) for every predicate symbol of \(T\) and letting \(\hat{\mathcal{N}}_U \subseteq \mathcal{E}_1(\Omega) \times \mathcal{E}_1(\Omega)\) be any measurable set such that (8) holds. By also letting \(\xi \stackrel{\text{def}}{=} \hat{\phi}\), we can deduce the equalities in (10) in a different order:

\[
\psi(M) = \phi_{\mathcal{N}_f}^M = \frac{|M|!}{|\text{Aut}(M)|} \cdot t_{\text{ind}}(M, \mathcal{N}_f)
\]

\[
= \frac{|M_U|!}{|\text{Aut}(M_U)|} \cdot t_{\text{ind}}(M_U, \hat{\mathcal{N}}) \frac{\mu(\hat{\mathcal{N}})}{\mu(U_n)} = (\xi \circ \pi(U,I))(M).
\]

Similarly, the equalities in (9) also hold deduced in a different order:

\[
c = \int_X f \, d\mu = \mu(\hat{\mathcal{N}}_U) = \mu(T(U, \hat{\mathcal{N}})) = \xi(u).
\]

Finally, we let \((\hat{\mathcal{N}}_n)_{n \in \mathbb{N}}\) be a sequence of models of \(\hat{T}\) converging to \(\hat{\mathcal{N}}\), let \(N_n \stackrel{\text{def}}{=} I(\hat{\mathcal{N}}_n)\) and \(U_n \stackrel{\text{def}}{=} U\hat{\mathcal{N}}_n \stackrel{\text{def}}{=} \{v \in V(\hat{\mathcal{N}}_n) \mid \hat{\mathcal{N}}_n \models U(v)\}\) and note that \(\lim_{n \to \infty} |U_n|/|N_n| = \xi(u) = c\) and for every \(M \in \mathcal{M}[T]\), we have

\[
\lim_{n \to \infty} p(M, N_n|U_n) = \lim_{n \to \infty} p(M_U, \hat{\mathcal{N}}_n) \cdot \left(\frac{|N_n|}{|U_n|}\right)^{|M|} = \left(\frac{\xi(M_U)}{\xi(u|M|)}\right) = (\xi \circ \pi(U,I))(M) = \psi(M),
\]

concluding the proof.

From this theorem, Lemma 3.3 on subgraphons follows trivially.

Proof of Lemma 3.3. Follows directly from Lemma 5.8 via the correspondence between \(T_{\text{Graphons}}\) and graphons of Remark 5.3.

For general universal theories, the role of complete or empty graphons (i.e., \(W\) constant equal to 0 or 1) is played by trivial limits defined below.

**Definition 5.9.** A limit \(\phi \in \text{Hom}^+(\mathcal{A}[T], \mathbb{R})\) is called **trivial** if there exists a \(T\)-on \(\mathcal{N}\) with \(\phi_\mathcal{N} = \phi\) and each \(P\)-on \(\mathcal{N}_P\) either has measure 0 or 1. Equivalently, a limit \(\phi \in \text{Hom}^+(\mathcal{A}[T], \mathbb{R})\) is trivial if and only if it is of the form \(\phi = \psi^I\) for some open interpretation \(I: T \rightsquigarrow T_0\) and the unique \(\psi \in \text{Hom}^+(\mathcal{A}[T_0], \mathbb{R})\), where \(T_0\) is the trivial universal theory, that is, the theory over the empty language without any axioms.
Before we can finally state the stability dichotomy theorem for limits of arbitrary universal theories, we also need to define stability in this more general setting.

**Definition 5.10.** Recall that for a formula $F(\vec{x}, \vec{y})$ with a particular partition of its free variables into two parts $\vec{x}$ and $\vec{y}$, a half-graph of order $n$ with respect to $F(\vec{x}, \vec{y})$ in a structure $M$ is a pair of sequences $(\vec{a}_1, \ldots, \vec{a}_n)$ and $(\vec{b}_1, \ldots, \vec{b}_n)$ of tuples of vertices of $M$ with $|\vec{a}_i| = |\vec{x}|$, $|\vec{b}_i| = |\vec{y}|$ and such that $M \models F(\vec{a}_i, \vec{b}_j)$ if and only if $i \leq j$. A tree of height $n$ with respect to $F(\vec{x}, \vec{y})$ in a structure $M$ is a pair of sequences $(\vec{a}_\sigma \mid \sigma \in \{0, 1\}^n)$ and $(\vec{b}_m \mid m \in \{0, 1, \ldots, n-1\}, \tau \in \{0, 1\}^m)$ such that $|\vec{a}_\sigma| = |\vec{x}|$, $|\vec{b}_m| = |\vec{y}|$ and for every $\sigma = (\sigma_i)_{i=1}^n \in \{0, 1\}^n$ and every $m < n$, $M \models F(\vec{x}_{\sigma_1}, \vec{y}_{\sigma_{|m|}})$ if and only if $\sigma_{m+1} = 1$.

We say that $F(\vec{x}, \vec{y})$ is almost stable in a limit $\phi \in \text{Hom}^+(\mathcal{A}[T], \mathbb{R})$ if there exists $n \in \mathbb{N}$ such that every finite model $M$ of $T$ containing a half-graph of order $n$ with respect to $F(\vec{x}, \vec{y})$ satisfies $\phi(M) = 0$. Equivalently, letting

$$H_{n,F}(\vec{x}_1, \ldots, \vec{x}_n, \vec{y}_1, \ldots, \vec{y}_n) \overset{\text{def}}{=} \bigwedge_{1 \leq i < j \leq n} F(\vec{x}_i, \vec{y}_j) \wedge \bigwedge_{1 \leq j < i \leq n} \neg F(\vec{x}_i, \vec{y}_j) \quad (11)$$

be the formula encoding a half-graph of order $n$ with respect to $F(\vec{x}, \vec{y})$, the formula $F(\vec{x}, \vec{y})$ is almost stable in $\phi \in \text{Hom}^+(\mathcal{A}[T], \mathbb{R})$ if there exists $n \in \mathbb{N}$ such that for every (not-necessarily injective) substitution $H$ of the variables of the formula $H_{n,F}$, the set $T(H, \mathcal{N})$ has measure 0 for some (equivalently, every) $T$-on $\mathcal{N}$ such that $\phi = \phi_\mathcal{N}$.

It will also be convenient to define a weak version of almost stability: we say that $F(\vec{x}, \vec{y})$ is almost weakly stable in $\phi \in \text{Hom}^+(\mathcal{A}[T], \mathbb{R})$ if there exists $n \in \mathbb{N}$ such that $T(H_{n,F}, \mathcal{N})$ has measure 0 for some (equivalently, every) $T$-on $\mathcal{N}$ such that $\phi = \phi_\mathcal{N}$. Thus the difference between stability and weak stability is whether the tuples of the half-graph are allowed to repeat vertices or not.

Our stability dichotomy theorem for limits of universal theories will be particularly concerned with the case when $F(\vec{x}, \vec{y})$ is a $(1, k(P) - 1)$-split of a predicate symbol $P$ (whose arity $k(P)$ is at least 2), that is, we have

$$F(x, y_1, \ldots, y_{k(P)-1}) \overset{\text{def}}{=} P(y_1, \ldots, y_{i-1}, x, y_i, y_{i+1}, \ldots, y_{k(P)-1})$$

for some $i \in [k(P)]$. 
Theorem 5.11. The following are equivalent for a $T$-on $\mathcal{N}$ over a space $\Omega = (X, \mathcal{A}, \mu)$.

i. $\phi_\mathcal{N}$ contains a trivial sub-object $\psi$.

ii. There exists a positive measure $U \subseteq X$ such that for every measure-isomorphism $F$ modulo 0 from $\Omega_U$ to $\Omega$, the sub-object $\phi_{\mathcal{N}|_U}$ is trivial.

iii. $\phi_\mathcal{N}$ contains a sub-object $\psi$ in which every $(1, k(P) - 1)$-split of every predicate symbol $P$ is almost weakly stable.

The same observations of Discussion 3.7 can be made here: the equivalence between items (i) and (ii) is not immediate since not every sub-object is of the form that appears in the latter item. However, by an argument analogous to that in Discussion 3.7, if $P$ is a property of limits that is closed under sub-objects, then $\phi_\mathcal{N}$ has a sub-object satisfying $P$ if and only if there exists $U \subseteq X$ such that $\phi_{\mathcal{N}|_U}$ satisfies $P$ for every measure-isomorphism $F$ modulo 0 from $\Omega_U$ to $\Omega$.

Naturally, the main ingredient to prove the theorem above is a generalization of Lemma 3.8 for theons.

Lemma 5.12. Let $\mathcal{N}$ be a $T$-on over a space $\Omega = (X, \mathcal{A}, \mu)$ such that every $(1, k(P) - 1)$-split of every predicate symbol $P$ is almost weakly stable in $\phi_\mathcal{N}$. Then there exists a positive measure set $U \subseteq X$ such that for every measure-isomorphism $F$ modulo 0 from $\Omega_U$ to $\Omega$, the sub-object $\phi_{\mathcal{N}|_U}$ is trivial.

Proof. In this proof, we will work with measurability with respect to the $\sigma$-algebra corresponding to the completion of the measure $\mu$. Note that the result still follows for the original $\sigma$-algebra by simply changing the final set $U$ in a zero-measure set. Note also that it is enough to show the existence of some $U$ and $F$ such that $\phi_{\mathcal{N}|_U}$ is trivial as if $F'$ is any other measure-isomorphism modulo 0 from $\Omega_U$ to $\Omega$, then $\phi_{\mathcal{N}|_U}$ is also trivial.

Let $\mathcal{L}$ be the language of $T$. The proof is by induction in the sum $\sum_{P \in \mathcal{L}} k(P)$ of the arities of the predicate symbols.

If the language $\mathcal{L}$ is empty, the result is trivial.

Suppose then that $\mathcal{L}$ is non-empty, let $P \in \mathcal{L}$ and let $k \overset{\text{def}}{=} k(P)$ be its arity.

If $P$ is a unary predicate, then applying the result inductively for $\mathcal{L} \setminus \{P\}$, we get a measurable set $U' \subseteq X$ with $\mu(U') > 0$ such that for every $Q \in \mathcal{L} \setminus \{P\}$, we have $\mu(U'(\mathcal{N}|_{U'}^F)) \in \{0, 1\}$. Since $\mu(U') > 0$, at least one of
\( \mathcal{N}_P \cap U' \) or \( U' \setminus \mathcal{N}_P \) has positive \( \mu \)-measure, so letting \( U \) be any of these having positive measure gives the desired result.

Suppose now that \( k \geq 2 \). By Theorem 5.4 and Remark 5.5, we can replace \( \mathcal{N} \) with a \( T \)-on such that there exists \( n \in \mathbb{N} \) such that every \((1, k - 1)\)-split \( S(x, \tilde{y}) \) of \( P \) satisfies \( T(H_{n,S},\mathcal{N}) \subseteq \mathcal{D}_{nk}(\Omega) \) (here we are abusing the notation a bit by saying that the variables of \( H_{n,S} \) are indexed by \([nk]\)).

Let us consider the natural \((1, k - 1)\)-split of \( P \) given by \( P(x, \tilde{y}) \overset{\text{def}}{=} P(x, y_1, \ldots, y_{k-1}) \), which we will denote simply by \( P \) and for convenience of notation, let \( V_k \overset{\text{def}}{=} [k] \setminus \{1\} \).

For each \( a \in \mathcal{E}_1(\Omega) \cong X \) and \( b \in \mathcal{E}_{V_k}(\Omega) \setminus \mathcal{D}_{V_k}(\Omega) \), let

\[
\mathcal{N}_P(a, b) \overset{\text{def}}{=} \{ c \in X^{r(k) \setminus (r(1) \cup r(V_k))} \mid (a, b, c) \in \mathcal{N}_P \setminus \mathcal{D}_k(\Omega) \}
\]

be the set of points that complete \((a, b)\) to a point of \( \mathcal{N}_P \setminus \mathcal{D}_k(\Omega) \) (note that if \( a \in \{b_1, \ldots, b_{k-1}\} \), then \( \mathcal{N}_P(a, b) \) is immediately empty). Define further

\[
\mathcal{N}_P^1(b) \overset{\text{def}}{=} \{ a \in \mathcal{E}_1(\Omega) \mid \mu(\mathcal{N}_P(a, b)) = 1 \},
\]

\[
\mathcal{N}_P^0(b) \overset{\text{def}}{=} \{ a \in \mathcal{E}_1(\Omega) \mid \mu(\mathcal{N}_P(a, b)) = 0 \}
\]

and let \( B_0 \) be the set of all \( b \in \mathcal{E}_{V_k}(\Omega) \setminus \mathcal{D}_{V_k}(\Omega) \) such that both \( \mathcal{N}_P^1(b) \) and \( \mathcal{N}_P^0(b) \) are measurable. Fubini’s Theorem gives \( \mu(B_0) = 1 \).

Given a finite collection of points \( \{b_1, \ldots, b_t\} \subseteq B_0 \), let

\[
C(\{b_1, \ldots, b_t\}) \overset{\text{def}}{=} \{ (b_i)_v \mid i \in [t], v \in V_k \} \subseteq X
\]

be the set of coordinates of the \( b_i \) that are indexed by singletons and let

\[
B(\{b_1, \ldots, b_t\}) \overset{\text{def}}{=} \{ b \in B_0 \mid \forall v \in V_k, b_{(v)} \notin C(\{b_1, \ldots, b_t\}) \} \subseteq \mathcal{E}_{V_k}(\Omega)
\]

be the set of \( b \in B_0 \) whose coordinates indexed by a singletons do not appear in the set \( C(\{b_1, \ldots, b_t\}) \). We also let \( B(\emptyset) \overset{\text{def}}{=} B_0 \). Note that \( \mu(B(\{b_1, \ldots, b_t\})) = \mu(B_0) = 1 \).

We now construct sequences \( (A_\sigma)_\sigma \) and \( (b_\sigma)_\sigma \) indexed by finite strings over \( \{0, 1\} \) inductively in the length-lexicographic order \( \leq_{\text{LL}} \) as follows.

1. Set \( A_\emptyset \overset{\text{def}}{=} X \).
2. For a string $\sigma$, given $A_\sigma$ and $b_\tau$ for all $\tau <_{LL} \sigma$, if there exist $j \in \{0, 1\}$ and $b \in B(\{b_\tau \mid \tau <_{LL} \sigma\})$ such that $0 < \mu(N_P^j(b) \cap A_\sigma) < \mu(A_\sigma)$, then set

$$b_\sigma \stackrel{\text{def}}{=} b,$$

$$A_{\sigma j} \stackrel{\text{def}}{=} (A_\sigma \cap N_P^j(b)) \setminus C(\{b_\tau \mid \tau \leq_{LL} \sigma\}),$$

$$A_{\sigma(1-j)} \stackrel{\text{def}}{=} A_\sigma \setminus (N_P^j(b) \cup C(\{b_\tau \mid \tau \leq_{LL} \sigma\}));$$

otherwise, stop the construction.

By induction in the construction, it follows that if $A_\sigma$ is defined for every $\sigma \in \{0, 1\}^t$ of a fixed length $t$, then $\{A_\sigma \mid \sigma \in \{0, 1\}^t\}$ is a collection of pairwise disjoint sets of positive measure whose union has measure 1 (hence each of these sets is non-empty). Furthermore, if $a_\sigma \in A_\sigma (\sigma \in \{0, 1\}^t)$, then we can find a tree of height $t$ in $\mathcal{N}$ as follows. Let

$$U \stackrel{\text{def}}{=} \{0, 1\}^t \cup \bigcup_{m=0}^{t-1} (\{0, 1\}^m \times V_k)$$

and define $x \in \mathcal{E}_U(\Omega)$ as follows.

a. For each $\sigma \in \{0, 1\}^t$, let $x_{\{\sigma\}} \stackrel{\text{def}}{=} a_\sigma$.

b. For each $m \in \{0, 1, \ldots, t-1\}$, each $\tau \in \{0, 1\}^m$ and each $V \in r(V_k)$, let

$$x_{\{(\tau, v) \mid v \in V\}} \stackrel{\text{def}}{=} (b_\tau)_V,$$

that is, for the injection $\alpha_\tau: V_k \rightarrow U$ given by $\alpha_\tau(v) \stackrel{\text{def}}{=} (\tau, v)$, we have $\alpha_\tau^*(x) = b_\tau$.

c. For each $\sigma \in \{0, 1\}^t$ and each $m \in \{0, 1, \ldots, t-1\}$, let

$$c_{\sigma, m} \in \begin{cases} N_P(a_\sigma, b_{\sigma[m+1]}), & \text{if } \sigma_{m+1} = 1, \\ X_{r(k)} \setminus (r(1)) \cup r(V_k)) \setminus N_P(a_\sigma, b_{\sigma[m]}), & \text{if } \sigma_{m+1} = 0, \end{cases}$$

and for each $V \in r(V_k)$, let

$$x_{\{(\sigma \cup \{m\}, v) \mid v \in V\}} \stackrel{\text{def}}{=} (c_{\sigma, m})_{\{1\} \cup V}.$$
d. Define all other coordinates of \( x \) arbitrarily.

Let us make some observations about this construction. First, it is straightforward to check that no coordinate of \( x \) is defined more than once. Second, by induction in the construction, it follows that each \( A_\sigma \subseteq A_{|\sigma|}^{m+1} \); this means that the element \( c_{\sigma,m} \) of item (c) is guaranteed to exist from the definition of \( A_{|\sigma|}^{m+1} \). Third, all coordinates of \( x \) that are indexed by singletons are defined in items (a) and (b) and since these coordinates must be either \( a_\sigma \) for different \( \sigma \in \{0,1\}^t \) or coordinates indexed by singletons of some \( b_\tau \) for some \( \tau \in \{0,1\}^m \) with \( m \in \{0,1,\ldots,t-1\} \), it follows by construction that they must all be distinct, that is, we must have \( x \notin D_U(\Omega) \). Finally, if 

\[
G_t(\vec{x},\vec{y}) = \bigwedge_{\sigma \in \{0,1\}^t} \bigwedge_{m=0}^{t-1} -1^{1-\sigma_{m+1}} P(x_\sigma,\vec{y}|_{\sigma|[m]}),
\]

then \( x \in T(G,\mathcal{N}) \) after an appropriate bijective relabeling of variables.

From [Hod93, Lemma 6.7.9], we know that if a model has a tree of height \( 2^{n+2} - 2 \) with respect to \( P(x,\vec{y}) \), then it must have a half-graph of order \( n \), where the \( \vec{a} \) and \( \vec{b} \) parts of the half-graph are picked from the \( \vec{a} \) and \( \vec{b} \) parts of the tree, respectively, with all of them distinct. In particular, if \( t \geq 2^{n+2} - 2 \), then there exists an injection \( \alpha : [nk] \rightarrow U \) such that \( \alpha^*(x) \in T(H_{n,P},\mathcal{N}) \).

Since \( x \notin D_U(\Omega) \) and \( T(H_{n,P},\mathcal{N}) \subseteq D_{nk}(\Omega) \), the construction must stop before constructing all \( A_\sigma \) with \( |\sigma| = 2^{n+1} - 2 \).

Let then \( \bar{\sigma} \) be the last string considered by the construction and let \( A \overset{\text{def}}{=} A_{\bar{\sigma}} \) and \( B \overset{\text{def}}{=} B(\{b_\tau \mid \tau < \text{LL } \bar{\sigma}\}) \). We know that for every \( b \in B \), we have

\[
\frac{\mu(N^1_P(b) \cap A)}{\mu(A)} = 1 \frac{\mu(N^0_P(b) \cap A)}{\mu(A)} \in \{0,1\}.
\]

We now setup our induction: let \( \mathcal{L}' \overset{\text{def}}{=} (\mathcal{L} \setminus \{P\}) \cup \{P'\} \), where \( P' \) is a new predicate symbol of arity \( k(P') \overset{\text{def}}{=} k-1 \), let \( \bar{F} \) be a measure-isomorphism modulo 0 from \( \Omega_A \) to \( \Omega \) and define the \( T_{\mathcal{L}'} \)-on \( \mathcal{N}' \) over \( \Omega_A \) by letting \( \mathcal{N}'_Q \overset{\text{def}}{=} (\mathcal{N}|_A)_Q \) for every \( Q \in \mathcal{L}' \setminus \{P'\} \) and letting

\[
\mathcal{N}'_{P'} \overset{\text{def}}{=} \left\{ \tau^*(x) \left| x^{\bar{F}} \in B \land \frac{\mu(N^1_P(x^{\bar{F}}) \cap A)}{\mu(A)} = 1 \right\} \right.,
\]
where $\iota: [k - 1] \to V_k$ is the relabeling $\iota(v) = v + 1$ and $x^F$ is given by (5).

We claim that every $(1, k(Q) - 1)$-split of a predicate symbol $Q \in \mathcal{L}'$ is almost weakly stable in $\phi_{N'}$. For $Q \neq P'$ this obviously follows from the same property for $\phi_{N'}$ as $\mu_A$ is absolutely continuous with respect to $\mu$. For $P'$, if $S'(x, \vec{y}) \overset{\text{def}}{=} P'(y_1, \ldots, y_{i_0 - 1}, x, y_{i_0}, \ldots, y_{k - 2}) (i_0 \in [k - 1])$ is a $(1, k - 2)$-split of $P'$, then we want to show that $T(H_{n, S'}, N')$ has measure zero. To prove this, let $S(x, \vec{y})$ be the $(1, k - 1)$-split of $P$ given by $P(y_1, \ldots, y_{i_0}, x, y_{i_0 + 1}, \ldots, y_{k - 1})$ and let us change the indexing of the variables of the formulas $H_{n, S'}$ and $H_n, S$ as follows:

\[
H_{n, S'}(x_{i, j} \mid i \in V_k, j \in [n]) \overset{\text{def}}{=} \bigwedge_{1 \leq j_1 \leq j_2 \leq n} P'(x_{\beta_{j_1, j_2}(1)}, \ldots, x_{\beta_{j_1, j_2}(k - 1)}) \\
\quad \land \bigwedge_{1 \leq j_2 < j_1 \leq n} P'(x_{\beta_{j_1, j_2}(1)}, \ldots, x_{\beta_{j_1, j_2}(k - 1)}),
\]

\[
H_{n, S}(x_{i, j} \mid i \in [k], j \in [n]) \overset{\text{def}}{=} \bigwedge_{1 \leq j_1 \leq j_2 \leq n} P(x_{\gamma_{j_1, j_2}(1)}, \ldots, x_{\gamma_{j_1, j_2}(k)}) \\
\quad \land \bigwedge_{1 \leq j_2 < j_1 \leq n} P(x_{\gamma_{j_1, j_2}(1)}, \ldots, x_{\gamma_{j_1, j_2}(k)}),
\]

where the injections $\beta_{j_1, j_2}: [k - 1] \to V_k \times [n]$ and $\gamma_{j_1, j_2}: [k] \to [k] \times [n]$ are given by

\[
\beta_{j_1, j_2}(v) \overset{\text{def}}{=} \begin{cases} (v + 1, j_1), & \text{if } v \neq i_0, \\
(v + 1, j_2), & \text{if } v = i_0, \end{cases} \quad \gamma_{j_1, j_2}(v) \overset{\text{def}}{=} \begin{cases} (v, j_1), & \text{if } v \neq i_0 + 1, \\
(v, j_2), & \text{if } v = i_0 + 1. \end{cases}
\]

Suppose now that $z \in T(H_{n, S'}, N') \setminus \mathcal{D}_{V_k \times [n]}(\Omega_A)$ is such that all of its coordinates indexed by singletons are in $A$ and for every injection $\alpha: V_k \to V_k \times [n]$, we have $\alpha^*(z) \in B$. Then we can define a point $\widehat{z} \in T(H_{n, S}, N) \setminus \mathcal{D}_{[k] \times [n]}(\Omega)$ as follows.

A. For each $i \in V_k$ and each $j \in [n]$, let $\widehat{z}_{i(j)} \overset{\text{def}}{=} z_{i(j)}$.

B. For each $V \subseteq V_k \times [n]$ with $|V| \geq 2$, let $\widehat{z}_V \overset{\text{def}}{=} \vec{F}(z_V)$.

C. For each $j_1, j_2 \in [n]$ with $j_1 \leq j_2$, since $\beta^*_{j_1, j_2}(z) \in N'_{P'}$, the definition of $N'_{P'}$ implies that the point $w \overset{\text{def}}{=} (\beta_{j_1, j_2} \circ \iota^{-1})^*(z)^F \in \mathcal{E}_{V_k}(\Omega)$ satisfies $\mu(N'_P(w) \cap A) = \mu(A) > 0$, so we can let $\widehat{z}_{(i, j_{ij1})} \in N'_{P}(w) \cap A$ be different from all coordinates defined so far and define the coordinates.
\( \tilde{z}_V \) with \( \{(1,j_1)\} \subset V \subseteq \text{im}(\gamma_{j_1,j_2}) \) based on a point in \( \mathcal{N}_P(\tilde{z}_{(1,j_1)}), w \) so that \( \gamma_{j_1,j_2}^*(\tilde{z}) \in \mathcal{N}_P \).

D. Analogously, for each \( j_1, j_2 \in [n] \) with \( j_2 < j_1 \), since \( \beta_{j_1,j_2}^*(z) \notin \mathcal{N}_P' \), the definition of \( \mathcal{N}_P' \) implies that the point \( w \equiv (\beta_{j_1,j_2} \circ \iota^{-1})(z) \in \mathcal{E}_k(\Omega) \) satisfies
\[
\mu(\mathcal{N}_P^0(w) \cap A) = \mu(A) - \mu(\mathcal{N}_P^1(w) \cap A) = \mu(A) > 0,
\]
so we can let \( \tilde{z}_{(1,j_1)} \in \mathcal{N}_P^0(w) \cap A \) be different from all coordinates defined so far and define the coordinates \( \tilde{z}_V \) with \( \{(1,j_1)\} \subset V \subseteq \text{im}(\gamma_{j_1,j_2}) \) based on a point in the complement of \( \mathcal{N}_P(\tilde{z}_{(1,j_1)}), w \) so that \( \gamma_{j_1,j_2}^*(\tilde{z}) \notin \mathcal{N}_P \).

E. Finally, we define all other coordinates of \( \tilde{z} \) arbitrarily.

Since in items (C) and (D) we ensured that coordinates were not repeated, we get \( \tilde{z} \in T(H_{n,S}, \mathcal{N}) \setminus \mathcal{D}_{[k]\times[n]}(\Omega) \), a contradiction. Thus \( T(H_{n,S'}, \mathcal{N}') \) has measure zero.

Therefore, the \((1,k-2)\)-split \( S' \) of \( P' \) is almost weakly stable.

By inductive hypothesis, it follows that there exists a measurable set \( U \subseteq A \) with \( \mu_A(U) > 0 \) and a measure-isomorphism \( \tilde{F}' \) modulo 0 from \( \Omega_U \) to \( \Omega_A \) such that \( \phi_{\mathcal{N}_U'F'} \) is trivial. It follows from the definition of \( \mathcal{N}' \) that \( \phi_{\mathcal{N}_U'F'} \) is trivial, completing the proof.

The proof of Theorem 5.11 from Lemma 5.12 below is analogous to its graphon counterpart Theorem 3.6 from Lemma 3.8.

**Proof of Theorem 5.11.** The implication (ii) \( \implies \) (i) is trivial and the implication (i) \( \implies \) (iii) follows since (all splits of) all open formulas are almost stable in a trivial sub-object.

For the final implication (iii) \( \implies \) (ii), again it is enough to show the existence of some \( U \) and \( \tilde{F} \) such that \( \phi_{\mathcal{N}_U'F'} \) is trivial as if \( F' \) is any other measure-isomorphism modulo 0 from \( \Omega_U \) to \( \Omega \), then \( \phi_{\mathcal{N}_U'F'} \) is also trivial.

Let then \( \psi \) be a sub-object in which every \((1,k(P)-1)\)-split of every predicate symbol is almost weakly stable. By Lemma 5.8, there exists \( f : X \rightarrow [0,1] \) with \( \int_X f \, d\mu > 0 \) such that for any measure-isomorphism \( F \) modulo 0 from \( \Omega_f \) to \( \Omega \), we have \( \psi = \phi_{\mathcal{N}_f'F} \).

43
Let $V \overset{\text{def}}{=} \{ x \in X \mid f(x) > 0 \}$, let $\tilde{F}$ be any measure-isomorphism modulo 0 from $\Omega_V$ to $\Omega$ and consider the sub-object $\tilde{\psi} \overset{\text{def}}{=} \phi_{\mathcal{N}|\tilde{F}}$ of $\phi_{\mathcal{N}}$.

The same measure theoretic trick of Theorem 3.6 gives that every $(1, k(P) - 1)$ split of every predicate symbol $P$ is almost weakly stable in $\tilde{\psi}$: given one such split $S(\tilde{x}, y)$, since it is almost weakly stable in $\psi$, we know that there exists $n \in \mathbb{N}$ such that $\mu_f(T(H_{n,S}, \mathcal{N}|\tilde{F})) = 0$ for the half-graph formula $H_{n,S}$ of (11). For each $\epsilon > 0$, let

$$T^n(\epsilon)(H_{n,S}, \mathcal{N}|\tilde{F}) \overset{\text{def}}{=} \{ x \in T_{\text{ind}}(H_{n,S}, \mathcal{N}|\tilde{F}) \mid \forall v, f(x_{\{v\}}) > \epsilon \},$$

$$T^n(\epsilon)(H_{n,S}, \mathcal{N}|\tilde{F}) \overset{\text{def}}{=} \{ x \in T_{\text{ind}}(H_{n,S}, \mathcal{N}|\tilde{F}) \mid \forall v, f(x_{\{v\}}) > \epsilon \}.$$

Then it follows that

$$\mu_f(T^n(\epsilon)(H_{n,S}, \mathcal{N}|\tilde{F})) \geq \left( \epsilon \cdot \frac{\mu(V)}{\int_X f \, d\mu} \right)^{n-k(P)} \mu_V(T^n(\epsilon)(H_{n,S}, \mathcal{N}|\tilde{F}))$$

and since we have $T(H_{n,S}, \mathcal{N}|\tilde{F}) = \bigcup_{m \in \mathbb{N}_+} T^{1/m}(H_{n,S}, \mathcal{N}|\tilde{F})$ and $T(H_{n,S}, \mathcal{N}|\tilde{F}) = \bigcup_{m \in \mathbb{N}_+} T^{1/m}(H_{n,S}, \mathcal{N}|\tilde{F})$, it follows that $\mu_V(T(H_{n,S}, \mathcal{N}|\tilde{F})) = 0$, so $S$ is almost weakly stable in $\tilde{\psi}$.

By Lemma 5.12, there exists a measurable set $U' \subseteq X$ such that $\mu_V(U') > 0$ and the sub-object $\phi_{\mathcal{H}}$ is trivial, where $\mathcal{H} \overset{\text{def}}{=} (\mathcal{N}|\tilde{F})|_{U'}$ for any given measure-isomorphism $F'$ modulo 0 from $\Omega_{U'}$ to $\Omega_V$. The result now follows by letting $U \overset{\text{def}}{=} U' \cap V$ and using the measure-isomorphism $\tilde{F} \circ F'$ modulo 0 from $\Omega_U$ to $\Omega$.

Let us now revisit Example 3.11.

**Example 5.13.** An alternative way of constructing the $\{0, 1\}$-valued graphon of Example 3.11 that does not have any linear-sized almost clique or almost anti-clique is as follows.

Given a permutation $\sigma \in S_n$ and a set $U \subseteq [n]$, the subpermutation induced by $U$ is the unique permutation $\tau \in S_{|U|}$ such that for every $i, j \in [|U|]$, we have $\tau(i) < \tau(j) \iff \tau(\iota_U(i)) < \tau(\iota_U(j))$, where $\iota_U : [|U|] \to [n]$ is the unique increasing function with $\text{im}(\iota_U) = U$; equivalently, we have $\tau = \iota_{\sigma(U)}^{-1} \circ \sigma \circ \iota_U$.

Consider now the theory $T_{\text{Perm}} \overset{\text{def}}{=} T_{\text{LinOrder}} \cup T_{\text{LinOrder}}$, where $T_{\text{LinOrder}}$ is the theory of (strict) linear orders, that is, $T_{\text{Perm}}$ is the theory of two linear
orders on the same base set. There is a natural correspondence between $S_n$ and $\mathcal{M}_n[T_{\text{Perm}}]$ in which $\sigma \in S_n$ corresponds to the model $M_\sigma \in \mathcal{M}_n[T_{\text{Perm}}]$, in which the first order $\prec_1$ is the natural order on $[n]$ and the second order $\prec_2$ is given by $i \prec_2 j \iff \sigma^{-1}(i) \prec_2 \sigma^{-1}(j)$. Furthermore, under this correspondence, subpermutations correspond to submodels (up to isomorphism).

It is straightforward to check that if $\sigma_n$ is distributed uniformly at random in $S_n$, then with probability 1 the sequence $(\sigma_n)_{n \in \mathbb{N}}$ is convergent (as models of $T_{\text{Perm}}$) and it converges to the $T_{\text{Perm}}$-on $H$ over $[0, 1]^2$ given by

$$H_{\prec_i} \overset{\text{def}}{=} \{ x \in \mathcal{E}_2([0, 1]^2) \mid \pi_i(x_{(1)}) < \pi_i(x_{(2)}) \} \quad (i \in [2]),$$

where $\pi_i : [0, 1]^2 \to [0, 1]$ is the projection onto the $i$th coordinate. The limit $\phi_H$ is called the quasirandom permuton and it is easy to see that $\phi_H(\sigma) = 1/|\sigma|!$ for every permutation $\sigma$.

Consider then the open interpretation $I : T_{\text{Graph}} \rightsquigarrow T_{\text{Perm}}$ corresponding to the construction of the graph of agreements of a permutation given by

$$I(E)(x_1, x_2) \overset{\text{def}}{=} x_1 \neq x_2 \land (x_1 \prec_1 x_2 \leftrightarrow x_1 \prec_2 x_2).$$

The interpreted $T_{\text{Graph}}$-on $I(H)$ over $[0, 1]^2$ is then given by

$$I(H)_E = \{ (x, y) \in \mathcal{E}_2 \times \mathcal{E}_2 \mid x_{(1)} < x_{(2)} \leftrightarrow y_{(1)} < y_{(2)} \},$$

which means that $\phi^I_H = \phi_{I(H)}$ is precisely the limit $\phi_W$ encoded by the graphon of Example 3.11.

Since a clique (anti-clique, resp.) in a graph $G$ of agreements of a permutation $\sigma$ corresponds to an increasing (decreasing, resp.) sequence in $\sigma$, we have

$$\phi^I_H(K_n) = \phi^I_H(\overline{K_n}) = \frac{1}{n!}.$$  

6 Consequences for finite models

In this section, we transfer Theorem 5.11 to the finite world just as we did in Section 4 for the theory of graphs. To do so, we will use the generalization of the ultraproduct method of Elek–Szegedy [ES12] by Aroskar–Cummings [AC14] below.
Theorem 6.1 (Elek–Szegedy [ES12], Aroskar–Cummings [AC14]). Let $T$ be a canonical universal theory in a finite relational language $\mathcal{L}$, let $(N_n)_{n \in \mathbb{N}}$ be a convergent sequence of models of $T$, let $\mathcal{D}$ be a non-principal ultrafilter over $\mathbb{N}$ and let $k \in \mathbb{N}_{+}$ be such that $k(P) \leq k$ for every $P \in \mathcal{L}$.

Then there exists a separable realization $\Theta: \prod_{n \in \mathbb{N}} V(N_n)^k/\mathcal{D} \to \mathcal{E}_k$ of order $k$ and measurable sets $N_P \subseteq \mathcal{E}_{k(P)}$ for each $P \in \mathcal{L}$ such that

$$\mu^{k(P)} \left( \Theta^{-1}_{k(P)}(N_P) \triangle \prod_{n \in \mathbb{N}} P_{N_n}/\mathcal{D} \right) = 0$$

for every $P \in \mathcal{L}$, every restriction $\Theta_{k(P)}$ of $\Theta$ of order $k(P)$ and where $\mu^{k(P)}$ is the Loeb measure on $\prod_{n \in \mathbb{N}} V(N_n)^{k(P)}/\mathcal{D}$.

In plain English, the theorem above says that the ultraproduct construction is the pre-image of the $T$-on $\mathcal{N}$ under the separable realization $\Theta$, except for a zero-measure error. The properties of restrictions and liftings of separable realizations then imply that for an open formula $F(x_1, \ldots, x_m)$, we have

$$\mu^m \left( F \left( \prod_{n \in \mathbb{N}} N_n/\mathcal{D} \right) \right) = \mu^m ( \Theta^{-1}_m(T(F, \mathcal{N})) ) = \lambda(T(F, \mathcal{N})),$$

and thus $(N_n)_{n \in \mathbb{N}}$ converges to $\phi_{\mathcal{N}}$.

Just as in the graph case, we can use this connection to pull back the set yielding a trivial sub-object in the theon world through the separable realization and produce a linear-sized “almost trivial” submodel in the convergent sequence. For this we make the following natural definitions.

Definition 6.2. An increasing sequence $(N_n)_{n \in \mathbb{N}}$ of structures in a language $\mathcal{L}$ is almost trivial if for every $P \in \mathcal{L}$, we have $\lim_{n \to \infty} |P_{N_n}|/|N_n|^{k(P)} \in \{0, 1\}$, i.e., either all but $o(|N_n|^{k(P)})$ amount of $k(P)$-tuples satisfy $P$ or at most an $o(|N_n|^{k(P)})$ amount of $k(P)$-tuples satisfy $P$.

We say that $F(\bar{x}, \bar{y})$ is almost stable in $(N_n)_{n \in \mathbb{N}}$ if there exists $m \in \mathbb{N}$ such that $\lim_{n \to \infty} |H_{m,F}(N_n)|/|N_n|^{|\bar{x}|+|\bar{y}|} = 0$, i.e., only an $o(|N_n|^{|\bar{x}|+|\bar{y}|})$ amount of tuples satisfy the formula $H_{n,F}$.

Note that the notion of almost stability for convergent sequences corresponds to almost weak stability in the limit; this is because solutions of $H_{m,F}$ that repeat variables can only account for at most $O(|N_n|^{|\bar{x}|+|\bar{y}|-1})$ tuples.
Theorem 6.3. The following are equivalent for a convergent sequence \((N_n)_{n \in \mathbb{N}}\) of models of a universal theory \(T\) in a finite relational language \(\mathcal{L}\).

i. There exist \(c > 0\) and sets \(U_n \subseteq V(N_n)\) such that \(|U_n| \geq c|N_n|\) for every \(n \in \mathbb{N}\) and \((N_n|_{U_n})_{n \in \mathbb{N}}\) is almost trivial.

ii. There exist a subsequence \((N_{n_k})_{k \in \mathbb{N}}\) of \((N_n)_{n \in \mathbb{N}}\) and sets \(U_{n_k} \subseteq V(N_{n_k})\) such that \(\limsup_{k \to \infty} |U_{n_k}|/|N_{n_k}| > 0\) and every \((1,k(P)−1)\)-split of every predicate symbol \(P \in \mathcal{L}\) is almost stable in \((N_{n_k}|_{U_{n_k}})_{k \in \mathbb{N}}\).

Proof. The implication \(i \implies (ii)\) follows since (all splits of) all open formulas are almost stable in an almost trivial sequence.

For the implication \((ii) \implies (i)\), let \(\phi \in \text{Hom}^+(\mathcal{A}[T], \mathbb{R})\) be the limit of \((N_n)_{n \in \mathbb{N}}\). By hypothesis and possibly passing to a further subsequence \((N_{n_k})_{k \in \mathbb{N}}\), there exist sets \(U_{n_k} \subseteq V(N_{n_k})\) with \(\lim_{k \to \infty} |U_{n_k}|/|N_{n_k}| > 0\) such that the sequence \((N_{n_k}|_{U_{n_k}})_{k \in \mathbb{N}}\) is convergent and any \((1,k(P)−1)\)-split of any predicate symbol \(P \in \mathcal{L}\) is almost stable in it. Since \((N_{n_k})_{k \in \mathbb{N}}\) also converges to \(\phi\), by Lemma 5.8, it follows that \(\phi\) contains a sub-object in which every \((1,k(P)−1)\)-split of every predicate symbol \(P \in \mathcal{L}\) is almost stable.

Let then \(\mathcal{H}\) be a \(T\)-on over some space \(\Omega = (X, \mathcal{A}, \mu)\) with \(\phi = \phi_{\mathcal{H}}\). By Theorem 5.11, there exists a positive measure \(U \subseteq X\) such that \(\psi \overset{\text{def}}{=} \phi_{\mathcal{H}|_{U}}\) is trivial for every measure-isomorphism \(F\) modulo 0 from \(\Omega_U\) to \(\Omega\).

Let \(c \overset{\text{def}}{=} \mu(U) > 0\). We claim that for every \(T\)-on \(\mathcal{H}'\) over some space \(\Omega' = (X', \mathcal{A}', \mu')\) with \(\phi = \phi_{\mathcal{H}'}\), there exists a measurable set \(U' \subseteq X'\) such that \(\phi_{\mathcal{H}'|_{U'}} = \psi\) for every measure-isomorphism \(F'\) modulo 0 from \(\Omega_{U'}\) to \(\Omega\) and \(\mu'(U') \geq c\).

This is completely trivial from the Theon Uniqueness Theorem [CR20b, Theorems 3.9 and 3.11 and Proposition 7.7], but an ad hoc proof analogous to the one in Theorem 4.3 can be obtained from Lemma 5.8: by this lemma applied to \(\mathcal{H}\), there exists a sequence \((N'_n)_{n \in \mathbb{N}}\) converging to \(\phi\) and sets \(U'_n \subseteq V(N'_n)\) with \(\lim_{n \to \infty} |U'_n|/|N'_n| = c\) and \((N'_n|_{U'_n})_{n \in \mathbb{N}}\) converging to \(\psi\). Applying this lemma again to \(\mathcal{H}'\), it follows that for some measurable function \(f: X' \to [0,1]\) with \(\int_{X'} f \, d\mu' = c\) and every measure-isomorphism \(F'\) modulo 0 from \(\Omega_f\) to \(\Omega\), we have \(\phi_{\mathcal{H}'|_{U'_n}} = \psi\). Taking \(U' \overset{\text{def}}{=} \{x \in X' \mid f(x) > 0\}\) gives \(\mu'(U') \geq c\) and \(\phi_{\mathcal{H}'|_{U'}} = \psi\) for every measure-isomorphism \(F'\) modulo 0 from \(\Omega_{U'}\) to \(\Omega\) (note that the rescaling \(f\) does not change densities of submodels within \(U'\) because \(\psi\) is trivial) completing the proof of the claim.
Since \( \psi = \phi_{\mathcal{H}_U} \) is trivial, for each \( P \in \mathcal{L} \), we know that 
\[
\def\dnp\sum_{P \in \mathcal{L}} \left| \frac{P^{N_n \mid U_n}}{U_n \mid k(P)} - b_P \right|
\]
over all possible sets \( U_n \subseteq V(N_n) \) with \( |U_n| \geq (c/2) \cdot |N_n| \). To conclude the proof, it is sufficient to show that \( \lim_{n \to \infty} d_n(U_n^c) = 0 \). Suppose not. Then there exists a subsequence \( (N_{m_\ell})_{\ell \in \mathbb{N}} \) of \( (N_n)_{n \in \mathbb{N}} \) such that \( \lim_{\ell \to \infty} d_{m_\ell}(U_{m_\ell}^c) > 0 \) and by possibly passing to a further subsequence, we can also assume that 
\[
(N_{m_\ell} \mid U_{m_\ell}^c)_{\ell \in \mathbb{N}}
\]
is convergent.

We now let \( N' = \prod_{\ell \in \mathbb{N}} N_{m_\ell} \) for some non-principal ultrafilter \( \mathcal{D} \) over \( \mathbb{N} \), let \( k \geq k(P) \) for every \( P \in \mathcal{L} \), let \( \Theta: \prod_{\ell \in \mathbb{N}} V(N_{m_\ell})^k \to \mathcal{E}_k \) and \( N' \) be as in Theorem 6.1 and per our previous claim, there exists a measurable set \( U' \subseteq [0,1] \) with \( \lambda(U') \geq c \) and \( \phi_{\mathcal{N}_{U'}} = \psi \) for every measure-isomorphism \( F' \) modulo 0 from \( ([0,1], \lambda_{U'}) \) to \( ([0,1], \lambda) \). Define further \( \hat{U} \equiv \Theta_1^{-1}(U') \subseteq \prod_{\ell \in \mathbb{N}} V(N_{m_\ell})/\mathcal{D} \), where \( \Theta_1 \) is a restriction of \( \Theta \) or order 1, and note that since \( \Theta_1 \) is measure-preserving, we have \( \mu^1(\hat{U}) \geq c \) for the Loeb measure \( \mu^1 \).

Consider now the model \( N|\hat{U} \) and note that for every predicate symbol \( P \in \mathcal{L} \) and every restriction \( \Theta_{k(P)} \) of \( \Theta \) of order \( k(P) \), we have
\[
P^{N \mid U} = \Theta_{k(P)}^{-1}(\{x \in N_P \mid \forall v \in [k(P)], x_v \in U'\}) \text{ a.e.}
\]
and since \( \lambda_{U'}((N_{U'}^{\ell})_{\ell \in \mathbb{N}}) = b_P \), it follows that
\[
\mu^{k(P)}(P^{N \mid U}) = b_P \cdot \mu^{k(P)}(\hat{U}^{k(P)}) = b_P \cdot \mu^1(\hat{U})^{k(P)},
\]
where the last equality follows from Fubini’s Theorem for Loeb measures, Theorem A.2.

Let now \( U \equiv \prod_{\ell \in \mathbb{N}} U_\ell / \mathcal{D} \) be an internal set such that \( \mu^1(U \triangle \hat{U}) = 0 \). By Fubini’s Theorem again, it follows that \( \mu^{k(P)}(P^{N \mid U}) = b_P \cdot \mu^1(U)^{k(P)} \), so we must have
\[
\lim_{\ell \to \mathcal{D}} \frac{|U_\ell|}{|N_{m_\ell}|} = \mu^1(U) \geq c,
\]
\[
\lim_{\ell \to \mathcal{D}} \frac{|P^{N_{m_\ell} \mid U_\ell}|}{|U_\ell|^{k(P)}} = \frac{\mu^{k(P)}(P^{N \mid U})}{\mu^1(U)^{k(P)}} = b_P \quad (P \in \mathcal{L}).
\]
However, this is a contradiction because it implies that along some subsequence we have $|U_\ell|/|N_{m_\ell}| \geq c/2$ and $d_{m_\ell}(U_\ell) \to 0$, contradicting the fact that the former implies $d_{m_\ell}(U_{m_\ell}^c) \leq d_{m_\ell}(U_\ell)$ and we have $d_{m_\ell}(U_{m_\ell}^c) \not\to 0$. \hfill $\Box$

Finally, with an argument similar to that of Theorem 4.7, we can prove a stability dichotomy for countable models.

**Theorem 6.4.** Let $T$ be a universal theory in a finite relational language $\mathcal{L}$. The following for a countable model $N$ of $T$ with $V(N) = \mathbb{N}_+$.

i. There exist a set $U \subseteq \mathbb{N}_+$ and an increasing sequence $(n_\ell)_{\ell \in \mathbb{N}}$ of positive integers such that $(N|_{U \cap [n_\ell]})_{\ell \in \mathbb{N}}$ is almost trivial and $\lim_{\ell \to \infty} |U \cap [n_\ell]|/n_\ell > 0$.

ii. There exist a set $U \subseteq \mathbb{N}_+$ and an increasing sequence $(n_\ell)_{\ell \in \mathbb{N}}$ of positive integers such that every $(1, k(P) - 1)$-split of every predicate symbol $P \in \mathcal{L}$ is almost stable in $(N|_{U \cap [n_\ell]})_{\ell \in \mathbb{N}}$ and $\lim_{\ell \to \infty} |U \cap [n_\ell]|/n_\ell > 0$.

**Proof.** The implication (i) $\implies$ (ii) is trivial as (all splits of) all open formulas are almost stable in almost trivial sequences.

For the implication (ii) $\implies$ (i), by possibly passing to a subsequence of $(n_\ell)_{\ell \in \mathbb{N}}$, we may further assume that $(N|_{[n_\ell]})_{n \in \mathbb{N}}$ is convergent, so by Theorem 6.3, there exist $c > 0$ and sets $U_\ell \subseteq [n_\ell]$ such that $|U_\ell| \geq c \cdot n_\ell$ for every $\ell \in \mathbb{N}$ and $(N|_{U_\ell})_{\ell \in \mathbb{N}}$ is almost trivial.

We then use the same recursive definition from Theorem 4.7 of the sequence $(m_\ell)_{\ell \in \mathbb{N}}$ by

$$m_0 \overset{\text{def}}{=} n_0, \quad m_{t+1} \overset{\text{def}}{=} \min\{n_\ell \mid \ell \in \mathbb{N} \land n_\ell \geq 2^t \cdot m_t\}$$

and for each $t \in \mathbb{N}$, let $\ell_t \in \mathbb{N}$ be such that $m_t = n_{\ell_t}$.

Finally, letting

$$U \overset{\text{def}}{=} \bigcup_{t \in \mathbb{N}} U_{\ell_t} \cap ([m_t] \setminus [m_{t-1}]),$$

where $m_{-1} \overset{\text{def}}{=} 0$ gives the result by an argument similar to that of Theorem 4.7. \hfill $\Box$
7 The approximate Erdős–Hajnal property

In this section, we study more systematically the approximate Erdős–Hajnal property defined below.

**Definition 7.1.** We say that a universal theory $T$ has **approximate Erdős–Hajnal property** (AEHP) if every limit $\phi \in \text{Hom}^+(\mathcal{A}[T], \mathbb{R})$ contains a trivial sub-object.

By Theorem 5.11(i) $\iff$ (ii), we have $T \in \text{AEHP}$ if and only if every $T$-on $\mathcal{N}$ over some space $\Omega = (X, \mathcal{A}, \mu)$ has a positive measure $U \subseteq X$ such that $\phi_{|U}$ is trivial for some (equivalently, every) measure-isomorphism $F$ modulo 0 from $\Omega_U$ to $\Omega$. See also Discussion 7.14 for an equivalent formulation in terms of convergent sequences.

**Discussion 7.2.** Before we proceed to showing basic properties of AEHP, let us note that Examples 3.11 and 5.13 already bring to light a curious difference between the usual Erdős–Hajnal Conjecture and its approximate version.

For the usual version, a perfect graph $G$ of size $n$ is guaranteed to contain either a clique or an anti-clique of size at least $\sqrt{n}$. This is because of the trivial bound $\alpha(G)\chi(G) \geq n$ involving the independence and chromatic numbers of $G$ and the fact that the chromatic and clique numbers of $G$ are the same. On the other hand, the stable Ramsey Theorem [MS14, MS21] only guarantees that any stable graph on $n$ vertices contains a clique or anti-clique of size $n^c$ for some fixed $c \in (0, 1)$ that depends only on the largest order of a half-graph of $G$. More generally, the Erdős–Hajnal Conjecture is believed to be true [FPS19] for hereditary graphs that whose neighborhoods of vertices have bounded Vapnik–Chervonenkis dimension [VC71] (these are also known as classes with NIP, i.e., without the independence property, in model theory).

However, for the approximate version, Theorem 4.3 implies that any convergent sequence of stable graphs contains a linear-sized almost clique or almost anti-clique, but since every graph of agreements of a permutation is a perfect graph, Example 5.13 says that there exists a convergent sequence of perfect graphs without any linear-sized almost clique or anti-clique.

Furthermore, it is easy to see that the theory $T$ of graphs of agreements of permutations has NIP, i.e., neighborhoods of vertices have bounded VC dimension: this is because any hereditary class of graphs without NIP is required to have at least $2^\Omega(n^2)$ different graphs with vertex set $[n]$ and $T$ has at most $(n!)^2$ different graphs with vertex set $[n]$ (as models of $T_{\text{Perm}}$ over $[n]$).
consist of two linear orders on \([n]\)).

This means that neither perfection nor NIP are enough to ensure that convergent sequences of graphs contain linear-sized almost cliques or anti-cliques. As we will see in Section 8, for graph theories, the approximate Erdős–Hajnal property is equivalent to forbidding some induced subgraph of some recursive blow-up of the 4-cycle.

Let us now prove some basic properties about \(\text{AEHP}\).

**Proposition 7.3.** If \(T' \vdash T\) and \(T \in \text{AEHP}\), then \(T' \in \text{AEHP}\).

**Proof.** This follows immediately since every \(T'\)-on is also a \(T\)-on. \(\square\)

Next we will study the universal theory analogue of the substitution operation studied for the original Erdős–Hajnal property (see also Remark 7.5 below and cf. [APS01] and [Chu14, §2]).

**Definition 7.4.** Let \(T_1\) and \(T_2\) be universal theories in the same language and let \(\forall x_1, \ldots, x_n, F(x_1, \ldots, x_n)\) be an axiom of \(T_1\).

We define the universal theory \(T_1^{F \rightarrow T_2}\) as the theory obtained from \(T_1\) by removing the axiom \(\forall x_1, \ldots, x_n, F(x_1, \ldots, x_n)\) and for every axiom of \(T_2\) of the form \(\forall y_1, \ldots, y_m, G(y_1, \ldots, y_m)\), adding the axiom

\[
\forall x_1, \ldots, x_{n+m-1}, F^G(x_1, \ldots, x_{n+m-1}),
\]

where \(F^G(x_1, \ldots, x_{n+m-1})\) is the formula

\[
G(x_n, \ldots, x_{n+m-1}) \lor \bigvee_{i=n}^{n+m-1} F(x_1, \ldots, x_{n-1}, x_i).
\]

**Remark 7.5.** When all predicate symbols of the language \(L\) have arity at most 2 and the theories \(T_1\) and \(T_2\) are of the form \(\text{Forb}_T(F)\) for some canonical theory \(T\) in \(L\) and some family \(F\) of finite models of \(T\), that is, the axioms of \(\text{Forb}_T(F)\) are those of \(T\) along with \(\forall \vec{x}, \neg D_{\text{open}}(F)(\vec{x})\) for each \(F \in F\), then the substitution operation can be done at the level of the family \(F\) (cf. [APS01] and [Chu14, §2]).

Namely, given finite \(L\)-structures \(F_1\) and \(F_2\) and some \(v \in V(F_1)\), the substitution \(F_1^v \rightarrow F_2\) of \(v\) by \(F_2\) in \(F_1\) is the \(L\)-structure obtained by replacing the vertex \(v\) with \(|F_2|\) copies of it inducing a copy of \(F_2\) (see Figure 5 for an example in the theory of graphs).

---

4 Note that any canonical theory can be reaxiomatized as \(\text{Forb}_{T_\kappa}(F)\) for some \(F\).
Note that if \( V(F_1) = [n] \), then the formula \( \neg D_{\text{open}}(F_1) \rightarrow \neg D_{\text{open}}(F_2) \) is equivalent to \( \neg D_{\text{open}}(F_1 \rightarrow F_2) \). Thus for families \( \mathcal{F}_1 \) and \( \mathcal{F}_2 \) of finite \( \mathcal{L} \)-structures and for \( F_1 \in \mathcal{F}_1 \) with \( V(F_1) = [n] \), the theory \( \text{Forb}_T(F_1) \rightarrow \text{Forb}_T(F_2) \) is equivalent to the theory \( \text{Forb}_T(\mathcal{F}') \), where

\[
\mathcal{F}' \overset{\text{def}}{=} (\mathcal{F}_1 \setminus \{F_1\}) \cup \{F_1 \rightarrow F_2 \mid F_2 \in \mathcal{F}_2\}.
\]

However, note that when \( \mathcal{L} \) has predicate symbols of arity at least 3, such easy description is not possible: the substitution operation will be completely agnostic about tuples containing at least two vertices of \( F_2 \) and at least one vertex of \( F_1 \) that is not \( v \).

Figure 5: Substitution operation for graphs.
Remark 7.6. Theories of the form \( \text{Forb}_T(\{F\}) \) with AEHP also bring to light models that have positive density in all limits without trivial sub-objects, namely, for each \( \phi \in \text{Hom}^+(\mathcal{A}[T], \mathbb{R}) \), let \( \mathcal{C}_P(\phi) \overset{\text{def}}{=} \mathcal{M}[\text{Th}(\phi)] \) be the set of finite models \( M \) of \( T \) (up to isomorphism) such that \( \phi(M) > 0 \) and let

\[
\mathcal{C}_P(T) \overset{\text{def}}{=} \bigcap_{\phi} \mathcal{C}_P(\phi),
\]

where the intersection is over all \( \phi \in \text{Hom}^+(\mathcal{A}[T], \mathbb{R}) \) that do not have any trivial sub-object (if \( T \) already has AEHP, this empty intersection is assumed to result \( \mathcal{M}[T] \) by convention).

We claim that \( \mathcal{C}_P(T) \) is exactly the class of finite models \( F \) of \( T \) (up to isomorphism) such that \( \text{Forb}_T(\{F\}) \in \text{AEHP} \). Both containments are more easily shown by their contrapositive. If \( \text{Forb}_T(\{F\}) \notin \text{AEHP} \), then there must be some \( \phi \in \text{Hom}^+(\mathcal{A}[\text{Forb}_T(\{F\})], \mathbb{R}) \) without any trivial sub-object, but for the axiom-erasing interpretation \( I: T \rightsquigarrow \text{Forb}_T(\{F\}) \) that acts identically on the language of \( T \), we have \( \phi^I \in \text{Hom}^+(\mathcal{A}[T], \mathbb{R}) \) and since \( \phi^I(F) = 0 \), we have \( F \notin \mathcal{C}_P(T) \). On the other hand, if \( F \) is a model of \( T \) that is not in \( \mathcal{C}_P(T) \), then there exists some \( \phi \in \text{Hom}^+(\mathcal{A}[T], \mathbb{R}) \) without any trivial sub-object such that \( \phi(F) = 0 \), but the latter condition implies that \( \phi \) can be seen as an element of \( \text{Hom}^+(\mathcal{A}[\text{Forb}_T(\{F\})], \mathbb{R}) \) and thus \( \text{Forb}_T(\{F\}) \notin \text{AEHP} \).

Note that we could have equivalently have defined \( \mathcal{C}_P(T) \) as the set of all finite models \( F \) that “persist” in the stronger sense that they have positive density in every sub-object \( \psi \) of every \( \phi \in \text{Hom}^+(\mathcal{A}[T], \mathbb{R}) \) that does not have trivial sub-objects. This seemingly stronger “persistence” definition yields precisely the same class of objects because of the quantification of \( \phi \) and the fact that if \( \phi \) does not have any trivial sub-object, then any sub-object of \( \phi \) also has this property.

Before we proceed, let us recall the definition of substitutionally closed theories from [CR20b, Definition 3.6].

Definition 7.7. Given an open formula \( F(x_1, \ldots, x_n) \) and an equivalence relation \( \sim \) on \( [n] \) with \( m \) equivalence classes \( C_1, \ldots, C_m \), we define the open formula \( F_\sim(y_1, \ldots, y_m) \) as the formula \( F(y_{\nu_1}, \ldots, y_{\nu_n}) \), where \( \nu \) is the unique function such that \( x_t \in C_{\nu_t} \) for every \( t \in [n] \).

A universal theory \( T \) is said to be substitutionally closed if for each axiom \( \forall \bar{x}, F(\bar{x}) \) and each equivalence relation \( \sim \), \( T \) proves \( \forall \bar{y}, F_\sim(\bar{y}) \) using only propositional rules and injective renamings of variables (but replacing two different variables with the same variable is disallowed).
The substitutional closure of $T$ is the theory whose axioms are $\forall \vec{y}, F_{\sim}(\vec{y})$ for each axiom $\forall x_1, \ldots, x_n, F(\vec{x})$ of $T$ and each equivalence relation $\sim$ on $[n]$.

Note that if $T'$ is the substitutional closure of $T$, then $T \vdash T'$ and $T' \vdash T$, that is, substitutional closedness is a property of the axiomatization of $T$ rather than its set of theorems. For substitutionally closed theories $T$, the next theorem from [CR20b] gives a simpler characterization of $T$-ons as Euclidean structures satisfying the axioms of $T$.

**Theorem 7.8** ([CR20b, Theorem 3.7]). Let $T$ be a canonical substitutionally closed universal theory in a finite relational language and $\mathcal{N}$ be an Euclidean structure on some space $\Omega = (X, \mathcal{A}, \mu)$ in the language of $T$. Then $\mathcal{N}$ is a weak (strong, respectively) $T$-on if and only if for every axiom $\forall x_1, \ldots, x_n, F(\vec{x})$ of $T$, we have $\mu(T(F, \mathcal{N})) = 1$ ($T(F, \mathcal{N}) \supseteq \mathcal{E}_n(\Omega) \setminus \mathcal{D}_n(\Omega)$, respectively).

**Remark 7.9.** It is easy to see that the forward direction of Theorem 7.8 does not require the substitutional closedness property. For the backward direction, the necessity of the property is illustrated in [CR20b, Example 37].

**Theorem 7.10.** Let $T_1$ and $T_2$ be canonical universal theories in the same finite relational language and let $\forall x_1, \ldots, x_n, F(x_1, \ldots, x_n)$ be an axiom of $T_1$. If $T_1, T_2 \in \text{AEHP}$ and $T_1^{F \rightarrow T_2}$ is canonical, then $T_1^{F \rightarrow T_2} \in \text{AEHP}$.

**Proof.** Let us first prove the case when all axioms of $T_1$ and $T_2$ are of the form

$$\forall x_1, \ldots, x_t, \left( \bigwedge_{1 \leq i < j \leq t} x_i \neq x_j \rightarrow A(\vec{x}) \right) \quad (12)$$

for some open formula $A$. Note that under these conditions $T_1$, $T_2$ and $T_1^{F \rightarrow T_2}$ are substitutionally closed as any replacement of two different variables with the same variable leads to a tautology.

Let $T \overset{\text{def}}{=} T_1^{F \rightarrow T_2}$. We need to show that every $T$-on $\mathcal{N}$ over some space $\Omega = (X, \mathcal{A}, \mu)$ contains a trivial sub-object. By possibly applying Theorem 5.4, we may suppose that $\mathcal{N}$ is a strong $T$-on. If $\mathcal{N}$ is a $T_1$-on, then this follows from $T_1 \in \text{AEHP}$, so suppose $\mathcal{N}$ is not a $T_1$-on. Since the only axiom of $T_1$ that is not an axiom of $T$ is $F$, by Theorem 7.8, we must have $\mu(T(F, \mathcal{N})) < 1$ and thus $\mu(T(\neg F, \mathcal{N})) > 0$. By Fubini's Theorem, there exists some $z \in \mathcal{E}_{n-1}(\Omega)$ such that the set

$$C(z) \overset{\text{def}}{=} \{ y \in X^{\{z\}} \mid \mu(C(z, y)) > 0 \},$$

54
has positive measure, where
\[ C(z, y) \overset{\text{def}}{=} \{ w \in X^{r(n)} \cap X^{r(n-1)} \cup \{ n \} \mid (z, y, w) \in T(\neg F, N) \}. \]

By identifying \( X^{\{ n \}} \) with \( X \), the set
\[ U \overset{\text{def}}{=} \{ y \in C(z) \mid \forall j \in [n - 1], y \neq z_{(j)} \} \]
also has positive measure.

Let \( \tilde{F} \) be a measure-isomorphism modulo 0 from \( \Omega_U \) to \( \Omega \). We claim that \( N|_{\tilde{F}} U \) is a \( T^2 \)-on. Suppose not. By Theorem 7.8, there exists some axiom of \( T^2 \) of the form \( \forall x_n, \ldots, x_{n+m-1}, G(x_n, \ldots, x_{n+m-1}) \) (we index the variables by \( V \overset{\text{def}}{=} \{ n, \ldots, n + m - 1 \} \) for convenience) such that \( \mu(T(G, N|_{\tilde{F}})) < 1 \). In particular, this means that there exists a point \( y \in \mathcal{E}_V(\Omega) \setminus \mathcal{D}_V(\Omega) \) such that \( y \notin T(G, N|_{\tilde{F}}) \) and \( y_{\{ v \}} \in U \) for every \( v \in V \). Then we can define a point \( \tilde{z} \in \mathcal{E}_{n+m-1}(\Omega) \setminus \mathcal{D}_{n+m-1}(\Omega) \) as follows.

a. For each \( A \in \mathcal{r}(n - 1) \), define \( \tilde{z}_A \overset{\text{def}}{=} z_A \).

b. For each \( A \in \mathcal{r}(V) \), define \( \tilde{z}_A \overset{\text{def}}{=} y_A^{\tilde{F}} \), where \( y_A^{\tilde{F}} \) is given by (5).

c. For each \( i \in V \), since \( y_{\{ i \}} \in U \subseteq C(z) \), let \( w^i \in C(z, y_{\{ i \}}) \) and define \( \tilde{z}_{A \cup \{ i \}} \overset{\text{def}}{=} w_{A \cup \{ i \}}^i \) for every \( A \in \mathcal{r}(n - 1) \).

d. Define all other coordinates arbitrarily.

The definition of \( U \) ensures that \( \tilde{z} \notin \mathcal{D}_{n+m-1}(\Omega) \). Furthermore, since \( (z, y_{\{ i \}}, w^i) \in T(\neg F, N) \) for every \( i \in V \) and \( y \in T(\neg G, N|_{\tilde{F}}) \), it follows that \( \tilde{z} \notin T(F^G, N) \), contradicting the fact that \( N \) is a strong \( T \)-on.

Therefore \( N|_{\tilde{F}} U \) is a \( T^2 \)-on and since \( T_2 \in \text{AEHP} \), it must contain a trivial sub-object, which must also be a sub-object of \( \phi_N \) (as \( \phi_N|_{\tilde{F}} \) is a sub-object of \( \phi_N \)).

Let us now prove the case in which all axioms of \( T_2 \) are of the form (12) but those of \( T_1 \) are not necessarily of this form. Let \( T_1' \) be the theory whose axioms are
\[ \forall y_1, \ldots, y_m, \left( \bigwedge_{1 \leq i < j \leq m} y_i \neq y_j \rightarrow A_{\neg}(y_1, \ldots, y_m) \right) \quad (13) \]
for every axiom \( \forall x_1, \ldots, x_t, A(x_1, \ldots, x_t) \) of \( T_1 \), every equivalence relation \( \sim \) on \([t]\) with \( m \) equivalence classes \( C_1, \ldots, C_m \) and without loss of generality, let us assume that we always enumerate these classes in a way that \( x_t \in C_m \). Let us use the notation \( A' \) for the open formula in (13). Note that \( T'_1 \vdash T_1 \), so \( T'_1 \in \text{AEHP} \) by Proposition 7.3.

Let us now focus our attention on the open formula \( F(x_1, \ldots, x_n) \) and let us enumerate all equivalence relations on \([n]\) as \( \sim_1, \ldots, \sim_\ell \).

We now define theories \( T^i \) for \( i \in \{0, \ldots, \ell\} \) inductively by letting \( T^0 \stackrel{\text{def}}{=} T'_1 \) and \( T^{i+1} \stackrel{\text{def}}{=} (T^i)_{F \sim_i \rightarrow T_2} \). A simple induction shows that \( T^i \) can be reaxiomatized so that all of its axioms are of the form (12) and thus by the previous case (and Proposition 7.3) another induction gives \( T^i \in \text{AEHP} \). On the other hand, it is straightforward to see that \( T^\ell \) is a reaxiomatization of \( T^{F \rightarrow T_1} \), so we get \( T^{F \rightarrow T_2} \in \text{AEHP} \) by Proposition 7.3.

Finally, for the case when both \( T_1 \) and \( T_2 \) are general, we can let \( T'_2 \) be the theory whose axioms are (13) but for every axiom of \( T_2 \) instead so that its axioms are all of the form (12). Then we clearly have \( T'_2 \vdash T_2 \) and \( T^{F \rightarrow T_2} \vdash T^{F \rightarrow T_2} \) so the result follows from two applications of Proposition 7.3 and the previous case.

Similarly to the results of Sections 4 and 6, the approximate Erdős–Hajnal property can also be pulled back to the finite world. Since the proofs are completely analogous to those of Section 6, we state these results without proof here.

**Theorem 7.11.** Let \( T \) be a universal theory in a finite relational language \( \mathcal{L} \) such that \( T \in \text{AEHP} \) and let \( (N_n)_{n \in \mathbb{N}} \) be a convergent sequence of structures in \( \mathcal{L} \).

Suppose there exists a subsequence \( (N^\ell_n)_{n \in \mathbb{N}} \) of \( (N_n)_{n \in \mathbb{N}} \) and sets \( U^\ell_n \subseteq V(N^\ell_n) \) such that \( \lim_{\ell \to \infty} |U^\ell_n|/|N^\ell_n| > 0 \) and for every finite \( \mathcal{L} \)-structure \( M \) that is not a model of \( T \), we have \( \lim_{\ell \to \infty} p(M, N^\ell_n|U^\ell_n) = 0 \).

Then there exist \( c > 0 \) and sets \( U_n \subseteq V(N_n) \) such that \( |U_n| \geq c|N_n| \) for every \( n \in \mathbb{N} \) and \( (N_n|U_n)_{n \in \mathbb{N}} \) is almost trivial.

**Theorem 7.12.** Let \( T \) be a universal theory in a finite relational language \( \mathcal{L} \) such that \( T \in \text{AEHP} \) and let \( N \) be a countable \( \mathcal{L} \)-structure with \( V(N) = \mathbb{N}_+ \).

Suppose there exist a set \( U \subseteq \mathbb{N}_+ \) and an increasing sequence \( (n^\ell)_{\ell \in \mathbb{N}} \) of positive integers such that for every finite \( \mathcal{L} \)-structure \( M \) that is not a model of \( T \), we have \( \lim_{\ell \to \infty} p(M, N|U \cap [n^\ell]) = 0 \) and \( \lim_{\ell \to \infty} |U \cap [n^\ell]|/n^\ell > 0 \).
Then there exist a set \( U \subseteq \mathbb{N}_+ \) and an increasing sequence \((n_\ell)_{\ell \in \mathbb{N}}\) of positive integers such that \((N|_{U \cap [n_\ell]})_{\ell \in \mathbb{N}}\) is almost trivial and \(\lim_{\ell \to \infty} |U \cap [n_\ell]|/n_\ell > 0\).

**Remark 7.13.** Differently from the case of Theorems 4.3, 4.7, 6.3 and 6.4, in Theorems 7.11 and 7.12 we do not get an equivalence as the “trivial” implication breaks down: an almost trivial sequence \((N_n)_{n \in \mathbb{N}}\) does not need to be a sequence of “almost” models of \(T\). For example, any universal theory \(T\) without infinite models (in flag algebra language, a degenerate theory) vacuously satisfies \(\text{AEHP}\) as it does not have any increasing sequence of models and by the same token it cannot have an increasing sequence of “almost” models. For a slightly less trivial example, if \(T\) is the theory of empty graphs and \((N_n)_{n \in \mathbb{N}}\) is an increasing sequence of complete graphs, then \((N_n)_{n \in \mathbb{N}}\) is (almost) trivial but does not contain any increasing subsequence of induced subgraphs that are “almost” empty graphs.

**Discussion 7.14.** A consequence of Theorem 7.11 is that \(T \in \text{AEHP}\) is equivalent to every convergent sequence of models of \(T\) having an almost trivial sequence of linear-sized induced submodels. Again, the convergence condition is essential (see Discussion 4.4) and requiring almost trivial as opposed to trivial is also essential (see Discussion 4.5).

As it was already observed in [Chu14, §5], without the conditions above the problem completely trivializes for graphs: if we require a universal theory \(T\) of graphs (i.e., \(T \vdash T_{\text{Graph}}\)) to be such that every sufficiently large model \(M\) of \(T\) either contains a clique or anti-clique of size strictly larger than \(\sqrt{|M|}\), then \(T\) must forbid some disjoint union of cliques and some complete partite graph. This stems from the graphs of Discussion 4.5: the largest cliques and anti-cliques in the graph \(H_{m,m}\) consisting of a disjoint union of \(m\) cliques of size \(m\) have size \(m\) so some induced subgraph of \(H_{m,m}\), which is necessarily a disjoint union of cliques, must be forbidden by \(T\). Similarly, the complement \(\overline{H}_{m,m}\) of \(H_{m,m}\) shows that \(T\) must forbid some complete partite graph.

### 8 Characterization via forbidden subgraphs

The purpose of this section is to completely characterize the approximate Erdős–Hajnal property \(\text{AEHP}\) for universal theories of graphs. Specifically, we show (Theorem 8.10) that universal theories of graphs with \(\text{AEHP}\) are precisely characterized as the ones that forbid some induced subgraph of
some recursive blow-up of the 4-cycle $C_4$ (defined below). Let us remind the reader that even the existence of such family characterizing AEHP for $T_{\text{Graph}}$ is a surprise: in general, it is not clear that given a universal theory $T$, there exists a family $\mathcal{C}$ such that any universal theory $T' \vdash T$ has AEHP if and only if it forbids some element of $\mathcal{C}$.

**Definition 8.1.** For $\ell \in \mathbb{N}$, the *recursive blow-up of the 4-cycle of height $\ell$* is the graph $C_4^\ell$ defined by $V(C_4^\ell) = [4]^\ell$ and in which two distinct vertices $\sigma, \tau \in [4]^\mathbb{N}$ are adjacent if and only if $\sigma_i - \tau_i \equiv \pm 1 \pmod{4}$, where $i \in [\ell]$ is the first position in which $\sigma$ and $\tau$ differ (see Figure 6).

We also let $\mathcal{C}_C$ be the set of all graphs (up to isomorphism) that are induced subgraphs of $C_4^\ell$ for some $\ell \in \mathbb{N}$.

The *recursive blow-up of the 4-cycle of countable height* is the graph $C_4^\omega$ defined by $V(C_4^\omega) = [4]^\mathbb{N}$ and in which two distinct vertices $\sigma, \tau \in [4]^\mathbb{N}$ are adjacent if and only if $\sigma_i - \tau_i \equiv \pm 1 \pmod{4}$, where $i \in \mathbb{N}$ is the first position in which $\sigma$ and $\tau$ differ.

**Remark 8.2.** It is easy to see that $\mathcal{C}_C$ can alternatively be described as the class of finite graphs $G$ that are induced subgraphs of $C_4^\ell$ for some $\ell \in \mathbb{N}$. If we wanted a smaller single graph $H$ whose class of finite induced subgraphs is $\mathcal{C}_C$, we could also take $H$ as the disjoint union $\bigsqcup_{\ell \in \mathbb{N}} C_4^\ell$ or as any direct limit $\lim \rightarrow C_4^\ell$ (in the categorical sense) relative to any direct system of embeddings $C_4^\ell \hookrightarrow C_4^k$ ($\ell \leq k$); both of these are countable graphs.

**Remark 8.3.** Let us note that there is not much particularly special about $C_4$ in the definition of $\mathcal{C}_C$. Namely, if $G \in \mathcal{C}_C$ contains at least one edge and one non-edge, then by analogously defining the recursive blow-ups $G^\ell$ and $G^\omega$ of $G$ of height $\ell \in \mathbb{N}$ and of countable height, respectively, it straightforward to check that $\mathcal{C}_C$ is precisely the set graphs that are induced subgraphs of some $G^\ell$ or alternatively, the set of finite induced subgraphs of $G^\omega$.

Let us now give an intuition of the steps required to show that a universal theory of graphs $T \vdash T_{\text{Graph}}$ has AEHP if and only if some $F \in \mathcal{C}_C$ is not a model of $T$.

First, recall from Remark 7.6 that $\mathcal{C}_P(T_{\text{Graph}})$ is the class of all graphs $G$ (up to isomorphism) that “persistently have positive density” in the sense that $\phi(G) > 0$ for every $\phi \in \text{Hom}^+(\mathcal{A}_\mathcal{T}_{\text{Graph}}, \mathbb{R})$ that does not have any trivial sub-object (i.e., every graphon without any subgraphon that is an almost clique or almost anti-clique). Recall also from Remark 7.6 that $\mathcal{C}_P(T_{\text{Graph}})$ can be described alternatively as the class of finite graphs $F$ (up to isomor-
Figure 6: Pictorial view of the recursive blow-up $C_4^4$ of the 4-cycle of height 4.

phism) such that $\text{Forb}_{T_{\text{Graph}}} \{ F \} \in \text{AEHP}$. Let now $C_M$ be the union of all classes of graphs $\mathcal{F}$ (up to isomorphism) that are minimal for the property that $\text{Forb}_{T_{\text{Graph}}} (\mathcal{F}) \in \text{AEHP}$, i.e., $C_M$ is the set of graphs that appear in some such minimal class. Our characterization can then be restated as the equality $\mathcal{C}_P (T_{\text{Graph}}) = C_M = C_C$.

To show these equalities, let us introduce one more class: let $\mathcal{C}_S$ be the smallest class of graphs (up to isomorphism) that contains all graphs of size at most 2 (i.e., the trivial graph $K_0$ with no vertices, the single vertex graph $K_1$, the edge $K_2$ and the non-edge $\overline{K}_2$) and that is closed under the substitution operation of Remark 7.5 (note that substitutions of the form $G^{v \rightarrow K_0}$ are isomorphic to $G|_{V(G)\backslash \{v\}}$, so we could have defined equivalently $\mathcal{C}_S$ as the
smallest class containing the edge, the non-edge and that is closed under both the substitution operation and taking induced subgraphs).

The proof of Theorem 8.10 can be informally summarized by the following steps.

1. By Remark 7.5 and Theorem 7.10 (and the fact that we trivially have $\text{Forb}_{T_{\text{Graph}}}(\{F\}) \in \text{AEHP}$ whenever $F$ has at most 2 vertices), it follows that $\mathcal{C}_S \subseteq C_P(T_{\text{Graph}})$.

2. In Lemma 8.7, we will show that $C_C \subseteq \mathcal{C}_S$ with an inductive argument. The combined inclusion $C_C \subseteq \mathcal{C}_S \subseteq C_P(T_{\text{Graph}})$ along with Proposition 7.3 then implies that if $F \in C_C$ is not a model of some universal theory of graphs $T$, then $T \in \text{AEHP}$ (as $T \vdash \text{Forb}_{T_{\text{Graph}}}(\{F\})$).

3. For the other implication, note that if all $F \in C_C$ are models of a universal theory of graphs $T$, then the limit $\phi_{C_4}$ of $(C^n_C)_{n \in \mathbb{N}}$ (see Definition 8.5) is a limit of $T$. By showing in Lemma 8.8 that $\phi_{C_4}$ does not have trivial sub-objects, we get $T \notin \text{AEHP}$ and the theorem follows. Another interpretation of this final step is that the fact that $\phi_{C_4}$ does not have trivial sub-objects implies that any collection of finite graphs $\mathcal{F}$ such that $\text{Forb}_{T_{\text{Graph}}}(\mathcal{F}) \in \text{AEHP}$ must necessarily have some element of $C_C$ (otherwise $\phi_{C_4}$ would be a limit of $\text{Forb}_{T_{\text{Graph}}}(\mathcal{F})$ as $C_C$ is downward closed). Since $C_C \subseteq C_P(T_{\text{Graph}})$, any minimal such collection $\mathcal{F}$ must be of the form $\{F\}$ for some $F \in C_C$ and thus $\mathcal{C}_M \subseteq C_C$, which along with the trivial containment $C_P(T_{\text{Graph}}) \subseteq \mathcal{C}_M$ gives the equality of all classes $\mathcal{C}_S = \mathcal{C}_C = \mathcal{C}_P(T_{\text{Graph}}) = \mathcal{C}_M$.

**Remark 8.4.** In the same way that $C_P(T)$ is defined for arbitrary universal theories $T$, we can also define $\mathcal{C}_M(T)$ as the union of all families $\mathcal{F}$ of finite models of $T$ (up to isomorphism) that are minimal for the property that $\text{Forb}_T(\mathcal{F}) \in \text{AEHP}$. Again we trivially have $C_P(T) \subseteq \mathcal{C}_M(T)$, but the other inclusion need not hold for general $T$. In fact, the equality $C_P(T) = \mathcal{C}_M(T)$ is equivalent to the statement that there exists a family $\mathcal{C}$ such that $T' \vdash T$ if and only if $T'$ forbids some model of $\mathcal{C}$ (namely, the family is $\mathcal{C} \overset{\text{def}}{=} C_P(T) = \mathcal{C}_M(T)$).

---

5There is a small difference between $\text{Forb}_{T_{\text{Graph}}}(\{K_0\})$ and $\text{Forb}_{T_{\text{Graph}}}(\{K_1\})$: the former has no models at all while the latter has only $K_0$ as its model (recall from Footnote 3 that we allow our models to have empty vertex set). However, since neither of them contain any increasing sequences of models, they satisfy $\text{AEHP}$ vacuously as they do not contain any limit object.
Definition 8.5. The limit recursive blow-up of $C_4$ is the limit object $\phi_{C_4} \in \text{Hom}^+(\mathcal{A}[T_{\text{Graph}}], \mathbb{R})$ that is the limit of the sequence $(C_4^n)_{n \in \mathbb{N}}$. It is straightforward to check that this sequence is convergent, but we can also alternatively define $\phi_{C_4}$ by giving an explicit $T_{\text{Graph}}$-on $\mathcal{N}_{C_4}$ representing it as follows. Let $\Omega \defeq ([4]^\mathbb{N}, \mathcal{A}, \nu)$ be the quaternary Cantor probability space, that is, $\mathcal{A}$ is the Borel $\sigma$-algebra of the product topology on $[4]^\mathbb{N}$ and $\nu$ is the unique Borel measure such that $\nu(K_\sigma) = 4^{-t}$ for every $t \in \mathbb{N}$ and every $\sigma \in [4]^{\{0,1,\ldots,t-1\}}$, where

$$K_\sigma \defeq \{ \tau \in [4]^\mathbb{N} \mid \forall i \in \{0,1,\ldots,t-1\}, \tau_i = \sigma_i \}. \quad (14)$$

The $T_{\text{Graph}}$-on $\mathcal{N}_{C_4}$ over $\Omega$ is defined by

$$\mathcal{N}_{E_{C_4}} \defeq \{ x \in \mathcal{E}_2(\Omega) \setminus \mathcal{D}_2(\Omega) \mid (x_{\{1\}})_i - (x_{\{2\}})_i \equiv \pm 1 \pmod{4} \},$$

where $i$ is the first position in which $x_{\{1\}}$ and $x_{\{2\}}$ differ.

The corresponding graphon $W_{C_4}$ over $\Omega$ as in Remark 5.3 is given by

$$W_{C_4}(x, y) \defeq \begin{cases} 1, & \text{if } x \neq y \text{ and } x_i - y_i \equiv \pm 1 \pmod{4}, \\ 0, & \text{otherwise}, \end{cases}$$

where $i$ is the first position in which $x$ and $y$ differ. By using the measure-isomorphism $F$ modulo 0 from $\Omega$ to $[0, 1]$ that maps $\sigma \in [4]^\mathbb{N}$ to $\sum_{i \in \mathbb{N}} \sigma_i \cdot 4^{-i-1}$, we obtain the equivalent graphon $\widehat{W_{C_4}}$ of Figure 7 given indirectly by $\widehat{W_{C_4}}(F(\sigma), F(\tau)) = W_{C_4}(\sigma, \tau)$. Under the interpretation that a $\{0,1\}$-valued graphon is simply a measurable graph, $W_{C_4}$ is just the recursive blow-up $C_{4}^{\omega}$ of the 4-cycle of countable height equipped with the quaternary Cantor probability measure.

Remark 8.6. As we will show in Lemma 8.8 below, $\phi_{C_4}$ does not contain any trivial sub-object and thus by Theorem 3.6, it does not contain any almost stable sub-object. In particular, this means that $\mathcal{C}_C$ must contain half-graphs of arbitrarily large order, which can be verified in an ad hoc fashion as follows.

First, it is easy to see that $\mathcal{C}_C$ is closed under substitutions as if $\alpha$ and $\beta$ are embeddings of $F_1, F_2 \in \mathcal{C}_C$ in $C_4^{\ell_1}$ and $C_4^{\ell_2}$, respectively and $v \in V(F_1)$, then defining the concatenation map $\gamma : V(F_1^{\langle v \rightarrow F_2 \rangle}) \to [4]^{\ell_1+\ell_2}$ by

$$\gamma(w) \defeq \begin{cases} (\alpha(w), 1^{\ell_2}), & \text{if } w \in V(F_1) \setminus \{v\}, \\ (\alpha(v), \beta(w)), & \text{if } w \in V(F_2) \end{cases}$$

61
Figure 7: Approximation of the graphon $\hat{W}^{C_4}$ of Definition 8.5. The graphon $\hat{W}^{C_4}$ has a fractal structure, whose first 3 steps are represented in the picture.

gives an embedding of $F_{v \rightarrow F_2}$ in $C_4^{\ell_1 + \ell_2}$. Thus, we have $C_S \subseteq C_C$.

Now, define a sequence $(\hat{H}_n)_{n \in \mathbb{N}}$ of clique-empty-half-graphs inductively by $\hat{H}_1 \overset{\text{def}}{=} K_2$ and

$$
\hat{H}_{n+1} \overset{\text{def}}{=} K_{2^v} \rightarrow \hat{K}_2 \rightarrow \hat{H}_n
$$

that is, starting from the edge $K_2 = \hat{H}_1$, we alternate substitution operations in $\hat{K}_2$ and in $K_2$ (obviously, the choices of the substituted vertex do not matter since $K_2$ and $\hat{K}_2$ are vertex-transitive). As the name suggests, $\hat{H}_n$ is a half-graph of order $n$ in which one of the sides forms a clique and the other forms an empty graph (see Figure 8) and since $\hat{H}_n \in C_S \subseteq C_C$, it follows that $C_C$ contains half-graphs of arbitrarily large order.

Recall that a finite graph $G$ is called prime if it cannot be obtained from smaller graphs via substitution, that is, $G$ is not of the form $F_{v \rightarrow F_2}$ for any graphs $F_1, F_2$ and $v \in V(F_1)$ with $|F_1|, |F_2| < |G|$.

**Lemma 8.7.** We have $C_C \subseteq C_S$. In particular, if $G$ is a finite graph such that $\phi_{C_4}(G) > 0$, then $G \in C_S$.

**Proof.** Let $G \in C_C$ and let us show that $G \in C_S$ by induction on the size $n$ of $G$. 

62
The base cases are when $G$ is a prime graph. In this case, we will show that $G$ must be an induced subgraph of $C_4 = C_4^1$. Let $\alpha$ be an embedding of $G$ in $C_4^\ell$ for some $\ell \in \mathbb{N}_+$. If $G$ has size at most 1, then it is trivially a subgraph of $C_4$. If not, let $\sigma$ be the longest string over $[4]$ that is a prefix of every string in $\text{im}(\alpha)$ and let $t$ be its length. For each $i \in [4]$, let $V_i \overset{\text{def}}{=} \{v \in V(G) \mid \alpha(v)_{t+1} = i\}$ and let $G_i \overset{\text{def}}{=} G|_{V_i}$. Let also $I \overset{\text{def}}{=} \{i \in [4] \mid V_i \neq \emptyset\}$ and let $H = C_4|_I$. Note that the structure of $C_4^\ell$ implies that $G$ can be obtained from $H$ by substituting each $i \in I$ by $G_i$. Since $|G_i| < |G|$ for every $i \in I$ and $G$ is prime, we must have $|H| = |G|$, that is, $|V_i| = 1$ for every $i \in I$ and thus the unique $\beta : V(G) \to [4]$ such that $v \in V_{\beta(v)}$ is an embedding of $G$ in $C_4$.

We claim that $G$ has size at most 2. Indeed, this follows because there are no prime graphs of size 3 and $C_4$ itself is not prime. Since $|G| \leq 2$, we trivially have $G \in C_S$.

For the inductive step, note that if $G$ is not prime, then it is of the form $F_1 \rightarrow F_2$ for some graphs $F_1, F_2$ and $v \in V(F_1)$ with $|F_1|, |F_2| < |G|$. By inductive hypothesis, we have $F_1, F_2 \in C_S$ and since $C_S$ is closed under substitutions, we get $G \in C_S$.

Finally, since $\phi_{C_4}$ is the limit of $(C_4^n)_{n \in \mathbb{N}}$, every $G$ with $\phi_{C_4}(G) > 0$ must
Lemma 8.8. \( \phi_{C_4} \) does not contain any trivial sub-object.

Proof. By [CKP21, Theorem 6] (see also Examples 3.11 and 5.13), to show that \( \phi_{C_4} \) does not have trivial sub-objects, we need to show that

\[
\lim_{n \to \infty} \phi_{C_4}(K_n)^{1/n} = \lim_{n \to \infty} \phi_{C_4}(\overline{K}_n)^{1/n} = 0.
\]

We claim that for every \( n \geq 2 \), we have

\[
\phi_{C_4}(K_n) = t_{\text{ind}}(K_n, \mathcal{N}^{C_4})
= \sum_{m \in \mathbb{N}} 4^m \cdot 4^{-nm} \cdot 4 \cdot 4^{-n} \cdot \sum_{t=1}^{n-1} \binom{n}{t} \cdot \phi_{C_4}(K_t) \cdot \phi_{C_4}(K_{n-t}) \tag{15}
= \frac{1}{4^{n-1} - 1} \cdot \sum_{t=1}^{n-1} \binom{n}{t} \cdot \phi_{C_4}(K_t) \cdot \phi_{C_4}(K_{n-t}).
\]

The first formula can be deduced by considering the measure of all copies of \( K_n \) in \( \mathcal{N}^{C_4} \) such that the largest common prefix \( \sigma \) of the vertex variables (which are strings in \([4]^{\mathbb{N}}\)) has length \( m \): there are exactly \( 4^m \) such \( \sigma \) and the set \( U_{\sigma} \subseteq E_n(\Omega) \) of points whose vertex variables all start with the prefix \( \sigma \) has measure \( 4^{-nm} \) (i.e., the vertex variables are in the set \( K_\sigma \) of (14)). Once in \( U_{\sigma} \), to yield a copy of \( K_n \), two vertex variables corresponding to different vertices \( i, j \in [n] \) that differ in the \((m+1)\)th position must satisfy \((x_i)_{m+1} = (x_j)_{m+1} \equiv \pm 1 \) (mod 4). This means that

\[
\mathcal{C} \overset{\text{def}}{=} \{ t \in [4] \mid \exists i \in [n], (x_{\{i\}})_{m+1} = t \}
\]

must induce a clique of size at least 2 in \( C_4 \) and in fact, of size 2 as \( C_4 \) is triangle-free. There are exactly 4 edges in \( C_4 \) and a requirement of the form \((x_{\{i\}})_{m+1} = c_i \) for each \( i \in [n] \) gives a conditional probability of \( 4^{-n} \) conditioned on \( U_{\sigma} \). Finally, the vertex variables must be split along the chosen edge of \( C_4 \) with \( t \) vertices to one side forming a \( K_t \) and \( n-t \) vertices to the other side forming a \( K_{n-t} \) and the recursive structure of \( \mathcal{N}^{C_4} \) allows us to compute the conditional probability of these events inductively.

64
With a similar argument, for every \( n \geq 2 \), we have

\[
\phi_{C_4}(K_n) = \sum_{m \in \mathbb{N}} 4^m \cdot 4^{-nm} \cdot 2 \cdot 4^{-n} \cdot \sum_{t=1}^{n-1} \binom{n}{t} \cdot \phi_{C_4}(K_t) \cdot \phi_{C_4}(K_{n-t})
\]

\[
= \frac{1}{2 \cdot (4^{n-1} - 1)} \cdot \sum_{t=1}^{n-1} \binom{n}{t} \cdot \phi_{C_4}(K_t) \cdot \phi_{C_4}(K_{n-t}).
\]

From this, a simple induction shows \( \phi_{C_4}(K_n) \leq \phi_{C_4}(K_{n}) \). Let \( c \) be the limit \( \lim_{n \to \infty} \phi_{C_4}(K_n)^{1/n} \) (which is guaranteed to exist by [CKP21, Theorem 6]) and suppose toward a contradiction that \( c > 0 \). Let \( n_0 \in \mathbb{N} \) be large enough so that

\[
\frac{3}{4} \cdot c \leq \phi_{C_4}(K_n)^{1/n} \leq \frac{5}{4} \cdot c
\]

for every \( n \geq n_0 \). Since \( c > 0 \), we can let

\[
a \overset{\text{def}}{=} \min \left\{ \phi_{C_4}(K_n) \cdot \left( \frac{3}{4} \cdot c \right)^{-n} \mid n \leq n_0 \right\} \cup \{1\},
\]

\[
b \overset{\text{def}}{=} \max \left\{ \phi_{C_4}(K_n) \cdot \left( \frac{5}{4} \cdot c \right)^{-n} \mid n \leq n_0 \right\} \cup \{1\},
\]

and note that since \( \phi_{C_4}(K_n) > 0 \) for every \( n \in \mathbb{N} \), it follows that \( a > 0 \). The definitions of \( a \) and \( b \) ensure that

\[
a \cdot \left( \frac{3}{4} \cdot c \right)^n \leq \phi_{C_4}(K_n) \leq b \cdot \left( \frac{5}{4} \cdot c \right)^n
\]

for \emph{every} \( n \in \mathbb{N} \) (as \( a \leq 1 \leq b \)).

Plugging these inequalities in (15), we get that for \( n \geq 2 \) we have

\[
a \cdot \left( \frac{3}{4} \cdot c \right)^n \leq \frac{1}{2 \cdot (4^{n-1} - 1)} \cdot b^2 \cdot \left( \frac{5}{4} \cdot c \right)^n \cdot \sum_{t=1}^{n-1} \binom{n}{t}
\]

\[
\leq \frac{1}{2 \cdot (4^{n-1} - 1)} \cdot b^2 \cdot \left( \frac{5}{4} \cdot c \right)^n \cdot 2^n,
\]

from which we conclude

\[
a \leq b^2 \cdot \left( \frac{5}{3} \right)^n \cdot \frac{2^n}{4^{n-1} - 1}.
\]
which by letting $n \to \infty$ yields $a = 0$, a contradiction. Therefore

$$\lim_{n \to \infty} \phi_{C_4}(K_n)^{1/n} = \lim_{n \to \infty} \phi_{C_4}(\overline{K}_n)^{1/n} = 0,$$

as desired. \qed

**Remark 8.9.** Similarly to Remark 8.3, the proof of Lemma 8.8 can be generalized to show that if $G \in \mathcal{C}_C$ has at least one edge and one non-edge, then the limit $\phi_G$ of the sequence $(G^n)_{n \in \mathbb{N}}$ of recursive blow-ups of $G$ does not contain any trivial sub-object.

We can finally put all pieces together to characterize AEHP for universal theories of graphs.

**Theorem 8.10.** The following are equivalent for a universal theory $T$ of graphs (i.e., $T \vdash T_{\text{Graph}}$).

i. We have $T \in \text{AEHP}$.

ii. There exists an induced subgraph $G \in \mathcal{C}_C$ of the recursive blow-up $\mathcal{C}_4^\omega$ of the 4-cycle of countable height such that $G$ is not a model of $T$.

iii. $\mathcal{C}_4^\omega$ is not a model of $T$.

In particular, we have $\mathcal{C}_S = \mathcal{C}_P(T_{\text{Graph}}) = \mathcal{C}_M = \mathcal{C}_C$.

**Proof.** The equivalence (ii) $\iff$ (iii) follows from Remark 8.2 and the fact that $T$ is universal.

For the implication (ii) $\implies$ (i), first note that if $F \in \mathcal{C}_C$, then Lemma 8.7 implies that $F \in \mathcal{C}_S$ so by Remark 7.5 and Theorem 7.10 (and the fact that trivially $\text{Forb}_{T_{\text{Graph}}}(\{F\}) \in \text{AEHP}$ whenever $|F| \leq 2$), we have $\text{Forb}_{T_{\text{Graph}}}(\{F\}) \in \text{AEHP}$. At this point we have $\mathcal{C}_C \subseteq \mathcal{C}_S \subseteq \mathcal{C}_P(T_{\text{Graph}})$.

On the other hand, if $T \vdash T_{\text{Graph}}$ is such that there exists $F \in \mathcal{C}_C$ that is not a model of $T$, then $T \vdash \text{Forb}_{T_{\text{Graph}}}(\{F\})$, so by Proposition 7.3, we have $T \in \text{AEHP}$.

We prove the implication (i) $\implies$ (ii) by the contra-positive: if every $G \in \mathcal{C}_C$ is a model of $T$, then $(\mathcal{C}_4^n)_{n \in \mathbb{N}}$ is a convergent sequence of models of $T$ whose limit $\phi_{C_4}$ does not have any trivial sub-object by Lemma 8.8, thus $T \notin \text{AEHP}$.
This implication shows that any family \( F \) that is minimal for the property \( \text{Forb}_{\text{Graph}}(F) \in \text{AEHP} \) must intersect \( C_C \). Since \( C_C \subseteq C_P(\text{Graph}) \), the minimality of \( F \) gives \( F = \{F\} \) for some \( F \in C_C \), thus \( C_M \subseteq C_C \), which along with the trivial inclusion \( C_P(\text{Graph}) \subseteq C_M \) and the already shown inclusion \( C_C \subseteq C_S \subseteq C_P(\text{Graph}) \) gives the equality \( C_S = C_P(\text{Graph}) = C_M = C_C \).

We conclude this section by showing that any universal theory \( T \) of graphs with \( \text{AEHP} \) also satisfies the usual Erdős–Hajnal property \( \text{(EHP)} \), that is, there exists \( c_T > 0 \) such that every graph of \( T \) of size \( n \) has a clique or anti-clique of size \( n^{c_T} \). Note that a priori it is not clear that the existence of linear-sized almost cliques or almost anti-cliques in convergent sequences of \( T \) should imply the existence of “polynomial-sized” cliques or anti-cliques in all graphs of \( T \). The proof of \( \text{AEHP} \implies \text{EHP} \) for graphs instead relies on the characterization of \( \text{AEHP} \) of Theorem 8.10 and analogue of Theorem 7.10 for \( \text{EHP} \) from [APS01, Theorem 1.1] (see also [Chu14, Theorem 2.3]) that inspired Theorem 7.10.

**Theorem 8.11.** If \( T \) is a universal theory of graphs with \( \text{AEHP} \), then \( T \) has \( \text{EHP} \).

**Proof.** We prove this by the contra-positive. Assume \( T \) does not have \( \text{EHP} \) and write \( T \) as \( \text{Forb}_{\text{Graph}}(F) \) for some \( F \) (see Remark 7.5).

For each \( F \in \mathcal{F} \), since \( T \) does not have \( \text{EHP} \), we know that \( \text{Forb}_{\text{Graph}}(\{F\}) \notin \text{EHP} \) (since all models of \( T \) are obviously models of \( \text{Forb}_{\text{Graph}}(\{F\}) \)). Let \( \mathcal{P}_F \) be the set of prime graphs that are induced subgraphs of \( F \). Since \( F \) can be obtained from the graphs in \( \mathcal{P}_F \) via substitution, by the contra-positive of [APS01, Theorem 1.1], there exists \( P_F \in \mathcal{P}_F \) such that the theory \( \text{Forb}_{\text{Graph}}(\{P_F\}) \) does not have \( \text{EHP} \).

Since \( \text{Forb}_{\text{Graph}}(\{K_0\}) \), \( \text{Forb}_{\text{Graph}}(\{K_1\}) \), \( \text{Forb}_{\text{Graph}}(\{K_2\}) \) and \( \text{Forb}_{\text{Graph}}(\{K_2\}) \) all have \( \text{EHP} \) (as the first theory has no models, the second is the theory whose unique model is \( K_0 \), the third is the theory of empty graphs and the fourth is the theory of complete graphs), we have \( P_F \notin \{K_0, K_1, K_2, K_2\} \).

Thus every graph \( F \) in \( \mathcal{F} \) has some prime subgraph \( P_F \) that is not \( K_0 \), \( K_1 \), \( K_2 \) or \( K_2 \), hence \( \mathcal{C}_S \subseteq \text{Forb}_{\text{Graph}}(F) = \mathcal{M}[T] \) as \( \mathcal{C}_S \) is the closure of \( \{K_0, K_1, K_2, K_2\} \) under substitutions.

By Theorem 8.10, it follows that every graph in \( \mathcal{C}_C = \mathcal{C}_S \) is a model of \( T \), hence \( T \notin \text{AEHP} \).  

---

The reader familiar with \( \text{EHP} \) may be more accustomed to the definition of \( \text{EHP} \) as a property of a graph \( H \) corresponding to \( \text{Forb}_{\text{Graph}}(\{H\}) \) having \( \text{EHP} \) as defined here.
9 Conclusion and open problems

In this paper we studied the asymptotic consequences of stability in the finite when coupled with the notion of convergence of densities, focusing particularly on producing linear-sized almost uniform sets in limits of convergent sequences of models. Once such uniform sets are produced in the limit, they can be pulled back to linear-sized almost uniform sets in convergent sequences of models or to positive upper-density almost uniform sets in countable models. We then studied which universal theories have the approximate Erdős–Hajnal property (AEHP), i.e., theories that must necessarily have linear-sized almost uniform sets in all of its limit objects (equivalently, in all of its convergent sequences) and we characterized the particular case of universal theories of graphs with AEHP as those that forbid some induced subgraph of some recursive blow-up of the 4-cycle.

A consequence of Theorems 4.7 and 6.4 is that any stable countable model must necessarily have an almost uniform set with positive upper density. As we mentioned in Discussion 4.8, one cannot hope to upgrade these theorems to produce almost uniform sets with positive density instead. A natural question is then what extra hypothesis would allow such upgrade? More concretely, a natural extra condition would be that of convergence of the marginals of the countable model, that is, is it true that a stable countable model $N$ such that $(N|_{[n]})_{n \in \mathbb{N}}$ is convergent must necessarily contain a positive density almost uniform set? On the one hand, this rules out the example of Discussion 4.8 as it does not have convergent marginals, but on the other hand, in the proofs of Theorems 4.7 and 6.4, it is not clear how to put together the sets $U_{\ell}$ returned by Theorems 4.3 and 6.3 into a single almost uniform set $U$ of positive density even in the presence of convergence of the marginals.

One of the interpretations of the usual Erdős–Hajnal Conjecture is that graphs that are not random have larger cliques or anti-cliques than the usual bound provided by Ramsey’s Theorem, in other words, the usual Erdős–Hajnal property can be seen as “failure of randomness”. In the case of the approximate Erdős–Hajnal property, this “failure of randomness” interpretation is even more prominent: every $T$-on $\mathcal{N}$ over a space $\Omega = (X, A, \mu)$ gives rise to a natural random exchangeable countable model $K$ of $T$ by sampling $x$ in $E_{\mathbb{N}^+}(\Omega)$ according to $\mu$ and letting

$$\mathcal{K} \models P(\alpha) \iff \alpha^*(x) \in \mathcal{N}_P.$$  \hfill (16)
Given a positive measure $U \subseteq X$, we can also define a natural random exchangeable countable model $K_U$ via (16) but taking $x$ in $E_{N, \mu}(\Omega)$ according to the product measure that uses $\mu_U$ for variables indexed by vertices and $\mu$ for all other variables. The natural quasirandomness property $U\text{Induce}[1]$ in [CR20a] (generalizing the graph quasirandomness property [CGW89, $P_4$]) requires that $K_U$ is equidistributed with $K$ for every positive measure $U \subseteq X$; informally, $\mathcal{N}$ is “random” in the sense of $U\text{Induce}[1]$ if restricting $\phi$ to any positive measure set yields the same limit $\phi$. In the context of $\text{AEHP}$, a limit $\phi$ that does not contain any trivial sub-object fails randomness in an even stronger sense: there exists a positive measure $U$ such that $K_U$ is a deterministic countable model (i.e., it is equal to some fixed $K$ with probability 1).

Elaborating further on this notion of weak randomness, we could call a limit object $\phi \in \text{Hom}^+(\mathcal{A}[T], \mathbb{R})$ weakly random if it satisfies the following weakening of $U\text{Induce}[1]$: every sub-object $\psi$ of $\phi$ satisfies $\text{Th}(\phi) = \text{Th}(\psi)$, that is, restricting to positive measure sets does not change which finite models have positive density. A consequence of the equality $\mathcal{C}_C = C_P(T_{\text{Graph}})$ proved in Section 8 is that the limit recursive blow-up of $C_4$ (and more generally, the limit recursive blow-up of any graph in $\mathcal{C}_C$ that has at least one edge and one non-edge) is weakly random. Just as in the theory of quasirandomness, it is natural to ask for equivalent characterizations of this weak randomness notion and higher arity generalizations of it.

In Theorem 8.11, we used the characterization of $\text{AEHP}$ from Theorem 8.10 to show that $\text{AEHP}$ implies $\text{EHP}$ for graphs, but the proof of Theorem 8.11 is non-constructive, so it is very natural to ask if a constructive proof is possible. More specifically, how does one find a clique or anti-clique of size $n^c$ knowing only that a linear-sized almost clique or almost anti-clique is guaranteed to exist in any convergent sequence? A very basic instance of this question is as follows: Theorem 3.6 provides an almost clique or anti-clique in a stable graphon by constructing a $0$-good set in the limit (see Discussion 3.9), which along with Theorem 8.11 implies that stable classes of graphs have $\text{EHP}$; on the other hand, this result is already known by stable Ramsey [MS14] but does not involve the known construction of $\epsilon$-good sets in the finite. It is natural to begin by trying to prove that stable classes of graphs have $\text{EHP}$ from the existence of $\epsilon$-good sets alone.

In Theorem 8.10, we characterized universal theories of graphs with the approximate Erdős–Hajnal property (AEHP) as precisely those that forbid some induced subgraph of a recursive blow-up of the 4-cycle. It is natural
to ask what happens for more complicated universal theories. For example, for universal theories of $k$-hypergraphs ($k \geq 3$), the substitution operation of Theorem 7.10 necessarily yields "agnostic edges" (see Remark 7.5) and recursive blow-ups of a hypergraph $H$ also have a similar degree of freedom: when we divide the space into parts $(V_i)_{i \in V(H)}$, how should we handle tuples containing at least two vertices in one part $V_i$ but not all vertices in $V_i$?

The behavior of $\text{AEHP}$ completely changes if we allow predicates to be asymmetric, namely, in a language $\mathcal{L} \overset{\text{def}}{=} \{E\}$ with a single binary predicate symbol $E$, if $\overline{K_2}$, $K_2$ and $A$ denote the anti-edge, the anti-parallel edges and the single edge, respectively (i.e., $V(\overline{K_2}) \overset{\text{def}}{=} V(K_2) \overset{\text{def}}{=} V(A) \overset{\text{def}}{=} [2]$, $E\overline{K_2} \overset{\text{def}}{=} \emptyset$, $E\overline{K_2} \overset{\text{def}}{=} \{(1, 2), (2, 1)\}$ and $E\overline{A} \overset{\text{def}}{=} \{(1, 2)\}$), then the canonical theories

$$\text{Forb}_{T_\mathcal{L}}(\{\overline{K_2}\}), \text{Forb}_{T_\mathcal{L}}(\{K_2\}), \text{Forb}_{T_\mathcal{L}}(\{A\})$$

clearly do not have $\text{AEHP}$: the first two because the sequence of transitive tournaments avoids linear-sized almost uniform sets and the last because $\text{Forb}_{T_\mathcal{L}}(\{A\}) \cong T_{\text{Graph}}$. On the other hand, $\text{Forb}_{T_\mathcal{L}}(\{\overline{K_2}, K_2, A\})$ does not have any models of size 2, so it trivially has $\text{AEHP}$ (as it has no limit object). This means that in general we cannot hope that for a universal theory $T$ there exists a family $C$ such that any universal theory $T' \vdash T$ has $\text{AEHP}$ if and only if it forbids some element of $C$, that is, in general we do not expect that $C_P(T) = C_M(T)$ (see Remark 8.4). A natural problem is then to characterize which theories have this "principality" property, and more generally, to study how different can $C_P(T)$ be from $C_M(T)$.

One might think that the example of the previous paragraph stems from the requirement of almost trivial being too strong for asymmetric predicates; after all, sets returned by Ramsey’s Theorem do not necessarily yield almost trivial sequences when asymmetric predicates are involved. Instead, one could define a property $\text{AEHP}'$ as $T \in \text{AEHP}'$ if every $\phi \in \text{Hom}^+(\mathcal{A}[T], \mathbb{R})$ has a finitely categorical sub-object $\psi$, that is, $\text{Th}(\psi)$ is finitely categorical in the model-theoretic sense (equivalently, for each $n \in \mathbb{N}$ there is exactly one model $M_n$ of size $n$ up to isomorphism such that $\psi(M_n) > 0$). Finitely categorical limits are precisely the limits of convergent sequences of sets that can be returned by Ramsey’s Theorem. To show failure of $\text{AEHP}'$ the sequence of transitive tournaments is not good as it converges to a finitely categorical limit. However, $\text{Forb}_{T_\mathcal{L}}(\{\overline{K_2}\})$ and $\text{Forb}_{T_\mathcal{L}}(\{K_2\})$ still do not have $\text{AEHP}'$ by using the quasirandom sequence of tournaments instead; thus "principality" still fails for $\text{AEHP}'$ over $\mathcal{L} \overset{\text{def}}{=} \{E\}$. It is not clear which of $\text{AEHP}$ or $\text{AEHP}'$ is
more appropriate in the presence of asymmetric predicates.

Acknowledgments

We are grateful to Avi Wigderson and Alexander Razborov for some useful comments on an earlier version of this manuscript.

References

[AC14] Ashwini Aroskar and James Cummings. Limits, regularity and removal for finite structures. Technical Report arXiv:1412.8084 [math.LO], arXiv e-print, 2014.

[AFP18] Nathanael Ackerman, Cameron Freer, and Rehana Patel. Stable regularity for relational structures. Technical Report arXiv:1712.09305 [math.LO], arXiv e-print, 2018.

[ALMM19] Noga Alon, Roi Livni, Maryanthe Malliaris, and Shay Moran. Private PAC learning implies finite Littlestone dimension. In STOC’19—Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, pages 852–860. ACM, New York, 2019.

[APS01] Noga Alon, János Pach, and József Solymosi. Ramsey-type theorems with forbidden subgraphs. volume 21, pages 155–170. 2001. Paul Erdős and his mathematics (Budapest, 1999).

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

[BCL10] Christian Borgs, Jennifer Chayes, and László Lovász. Moments of two-variable functions and the uniqueness of graph limits. Geom. Funct. Anal., 19(6):1597–1619, 2010.

[BLM20] Mark Bun, Roi Livni, and Shay Moran. An equivalence between private classification and online prediction. In 2020 IEEE 61st Annual Symposium on Foundations of Computer Science, pages 389–402. IEEE Computer Soc., Los Alamitos, CA, [2020] ©2020.
[CGW89] F. R. K. Chung, R. L. Graham, and R. M. Wilson. Quasi-random graphs. *Combinatorica*, 9(4):345–362, 1989.

[Chu14] Maria Chudnovsky. The Erdős-Hajnal conjecture—a survey. *J. Graph Theory*, 75(2):178–190, 2014.

[CK90] C. C. Chang and H. J. Keisler. *Model theory*, volume 73 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, third edition, 1990.

[CKP21] Oliver Cooley, Mihyun Kang, and Oleg Pikhurko. On a question of Vera T. Sós about size forcing of graphons. Technical Report arXiv:2103.09114 [math.CO], arXiv e-print, 2021.

[CR20a] L. N. Coregliano and A. A. Razborov. Natural quasirandomness properties. Technical Report arXiv:2012.11773 [math.CO], arXiv e-print, 2020.

[CR20b] L. N. Coregliano and A. A. Razborov. Semantic limits of dense combinatorial objects. *Uspekhi Mat. Nauk*, 75(4(454)):45–152, 2020.

[DF81] Persi Diaconis and David Freedman. On the statistics of vision: the Julesz conjecture. *J. Math. Psych.*, 24(2):112–138, 1981.

[DGH+21] Martin Doležal, Jan Grebík, Jan Hladký, Israel Rocha, and Václav Rozhoň. Relating the cut distance and the weak* topology for graphons. *J. Combin. Theory Ser. B*, 147:252–298, 2021.

[DJ08] Persi Diaconis and Svante Janson. Graph limits and exchangeable random graphs. *Rend. Mat. Appl. (7)*, 28(1):33–61, 2008.

[EFR86] P. Erdős, P. Frankl, and V. Rödl. The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent. *Graphs Combin.*, 2(2):113–121, 1986.

[EH89] P. Erdős and A. Hajnal. Ramsey-type theorems. volume 25, pages 37–52. 1989. Combinatorics and complexity (Chicago, IL, 1987).
[ER56] P. Erdős and R. Rado. A partition calculus in set theory. *Bull. Amer. Math. Soc.*, 62:427–489, 1956.

[ES12] Gábor Elek and Balázs Szegedy. A measure-theoretic approach to the theory of dense hypergraphs. *Adv. Math.*, 231(3-4):1731–1772, 2012.

[FPS19] Jacob Fox, János Pach, and Andrew Suk. Erdős-Hajnal conjecture for graphs with bounded VC-dimension. *Discrete Comput. Geom.*, 61(4):809–829, 2019.

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

[Kal05] Olav Kallenberg. *Probabilistic symmetries and invariance principles*. Probability and its Applications (New York). Springer, New York, 2005.

[Lov12] László Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012.

[LS06] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B*, 96(6):933–957, 2006.

[LS10] László Lovász and Balázs Szegedy. Regularity partitions and the topology of graphons. In *An irregular mind*, volume 21 of *Bolyai Soc. Math. Stud.*, pages 415–446. János Bolyai Math. Soc., Budapest, 2010.

[Mah42] Dorothy Maharam. On homogeneous measure algebras. *Proc. Nat. Acad. Sci. U.S.A.*, 28:108–111, 1942.

[MS14] M. Malliaris and S. Shelah. Regularity lemmas for stable graphs. *Trans. Amer. Math. Soc.*, 366(3):1551–1585, 2014.

[MS21] M. Malliaris and S. Shelah. Notes on the stable regularity lemma. *Bull. Symb. Log.*, 27(4):415–425, 2021.

[Pet13] F. Petrov. General removal lemma. Technical Report arXiv:1309.3795 [math.CO], arXiv e-print, 2013.
A Ultraproduct method

In this section, we present the formal definition of separable realizations from [ES12] (see also [AC14]). Throughout this section, we assume that we have a fixed sequence \((V_n)_{n \in \mathbb{N}}\) of finite sets of increasing sizes (intended to be the vertex sets of a convergent sequence of models) and we have fixed a non-principal ultrafilter \(\mathcal{D}\) over \(\mathbb{N}\).
**Definition A.1** (Loeb measure). Given a finite set \( U \), let \( \tau(U) \) be the Boolean algebra of internal subsets of \( \prod_{n \in \mathbb{N}} V_n^U / \mathcal{D} \). Let also \( \mu^U : \tau(U) \to [0, 1] \) be the finitely additive measure defined by the ultralimit

\[
\mu^U \left( \prod_{n \in \mathbb{N}} A_n / \mathcal{D} \right) \overset{\text{def}}{=} \lim_{n \to \mathcal{D}} \mu^U_n(A_n),
\]

where \( \mu^U_n(A) \overset{\text{def}}{=} \frac{|A|}{|V_n^U|} \) is the normalized counting measure on \( V_n^U \).

A \( \mu^U \)-nullset is a set \( N \subseteq \prod_{n \in \mathbb{N}} V_n^U \) such that for every \( \epsilon > 0 \), there exists \( B \in \tau(U) \) such that \( N \subseteq B \) and \( \mu^U(B) \leq \epsilon \). Let \( \sigma(U) \) be the collection of sets \( A \subseteq \prod_{n \in \mathbb{N}} V_n^U / \mathcal{D} \) that differ from some set in \( \tau(U) \) only by a \( \mu^U \)-nullset.

A standard saturation argument (see [ES12, Lemma 2.4]) shows that if \( A_m \in \tau(U) \) (\( m \in \mathbb{N} \)) are internal sets, then there exists an internal set \( B \supseteq \bigcup_{m \in \mathbb{N}} A_m \) with \( \mu^U(B) = \lim_{m \to \infty} \mu^U(\bigcup_{m' \leq m} A_{m'}) \). This in particular implies that \( \sigma(U) \) is a \( \sigma \)-algebra and that \( \mu^U \) is a finite pre-measure on \( \tau(U) \) and thus Carathéodory’s Theorem shows that \( \mu^U \) can be uniquely extended to a (complete) measure on \( \sigma(U) \), called **Loeb measure** and which we denote also by \( \mu^U \) by abuse.

As mentioned in Section 4, the probability space \((\prod_{n \in \mathbb{N}} V_n^U / \mathcal{D}, \sigma(U), \mu^U)\) is far from being standard, namely, it is non-separable. Furthermore, even the structure between these spaces is somewhat counter-intuitive: for \( U_1, U_2 \) disjoint and non-empty, \( \sigma(U_1 \cup U_2) \) is much larger than the completion of the product \( \sigma(U_1) \otimes \sigma(U_2) \). Nevertheless, the following analogue of Fubini’s Theorem still holds.

**Theorem A.2** (Fubini’s Theorem for Loeb measures). If \( A \in \sigma(U) \) and \( U' \subseteq U \), then for \( \mu_U \)-almost every \( x \in \prod_{n \in \mathbb{N}} V_n^{U'} / \mathcal{D} \), the set

\[
A(x) \overset{\text{def}}{=} \left\{ y \in \prod_{n \in \mathbb{N}} V_n^{U \setminus U'} / \mathcal{D} \mid (x, y) \in A \right\}
\]

is in \( \sigma(U \setminus U') \), the function \( x \mapsto \mu^{U \setminus U'}(A(x)) \) (defined arbitrarily when \( A(x) \) is not in \( \sigma(U \setminus U') \)) is measurable with respect to \( \sigma(U') \) and

\[
\mu^U(A) = \int_{\prod_{n \in \mathbb{N}} V_n^{U'} / \mathcal{D}} \mu^{U \setminus U'}(A(x)) \, d\mu^{U'}(x).
\]

Recall that an injection \( \alpha : U_1 \to U_2 \) defines contra-variantly the “projections” \( \alpha^* : \prod_{n \in \mathbb{N}} V_n^{U_2} / \mathcal{D} \to \prod_{n \in \mathbb{N}} V_n^{U_1} / \mathcal{D} \) and \( \alpha^* : [0, 1]^{r(U_2)} \to [0, 1]^{r(U_1)} \) via \( \alpha^*(x)_u = x_{\alpha(u)} \).
Definition A.3 (Separable realizations). Given finite sets $U' \subseteq U$, define the $\sigma$-algebra

$$\sigma(U', U) \overset{\text{def}}{=} (\iota^*)^{-1}(\sigma(U')) \overset{\text{def}}{=} \{(\iota^*)^{-1}(A) \mid A \in \sigma(U')\},$$

where $\iota: U' \to U$ is the inclusion map. We also let $\sigma(U', U)^*$ be the $\sigma$-algebra generated by $\{\sigma(U'', U) \mid U'' \subseteq U'\}$.

Given $k \in \mathbb{N}_+$, a separable realization of order $k$ is a measure-preserving function $\Theta: \prod_{n \in \mathbb{N}} V_{n}^k / \mathcal{D} \to [0, 1]^{r(k)}$ such that

i. For every $U \in r(k)$ and every Lebesgue measurable $A \subseteq [0, 1]$, the set $(\pi_U \circ \Theta)^{-1}(A)$ is in $\sigma(U, [k])$ and is independent from $\sigma(U, [k])^*$, where $\pi_U: [0, 1]^{r(k)} \to [0, 1]$ is the projection onto the $U$ coordinate.

ii. For every permutation $\sigma \in S_k$, we have $\sigma^* \circ \Theta = \Theta \circ \sigma^*$.

Given $m \in [k]$, a restriction of $\Theta$ of order $m$ is a measure-preserving function $\Theta_m: \prod_{n \in \mathbb{N}} V_{n}^m / \mathcal{D} \to [0, 1]^{r(m)}$ such that for every injection $\alpha: [m] \hookrightarrow [k]$, we have $\alpha^* \circ \Theta = \Theta_m \circ \alpha^*$.

Given $m \geq k$, a lifting of $\Theta$ of order $m$ is a measure-preserving function $\Theta_m: \prod_{n \in \mathbb{N}} V_{n}^m / \mathcal{D} \to [0, 1]^{r(m,k)}$ such that for every injection $\alpha: [k] \hookrightarrow [m]$, we have $\alpha^* \circ \Theta_m = \Theta \circ \alpha^*$.

It is straightforward to check that the properties of separable realizations, restrictions and liftings imply that $\Theta_m \circ \sigma^* = \sigma^* \circ \Theta_m$ for every $\sigma \in S_m$ both for liftings and restrictions. This in particular implies that restrictions of order $m$ are also separable realizations of order $m$. Furthermore, the definitions of restrictions and liftings themselves already show their uniqueness: restrictions must be defined by

$$\Theta_m(x)_A \overset{\text{def}}{=} \Theta(y)_A, \quad (17)$$

where $y \in \prod_{n \in \mathbb{N}} V_{n}^k / \mathcal{D}$ is any point in $(\iota^*)^{-1}(x)$ and $\iota: [m] \hookrightarrow [k]$ is the inclusion map and liftings must be defined by

$$\Theta_m(x)_A \overset{\text{def}}{=} \Theta(\alpha^*(x))_{[\ell]}, \quad (18)$$

where $\alpha: [k] \hookrightarrow [m]$ is any injection with $\alpha([\ell]) = A$. It is then straightforward to check that (17) and (18) give a restriction and a lifting, respectively (see [ES12, Lemma 3.2]).