HYPERGRAPH SEQUENCES AS A TOOL FOR SATURATION OF ULTRAPOWERS

M. E. MALLIARIS

ABSTRACT. Let T_1, T_2 be countable first-order theories, $M_i \models T_i$, and $\mathcal{D}$ any regular ultrafilter on $\lambda \geq \aleph_0$. A longstanding open problem of Keisler asks when T_2 is more complex than T_1, as measured by the fact that for any such $\lambda, \mathcal{D}$, if the ultrapower $(M_2)^\lambda/\mathcal{D}$ realizes all types over sets of size $\leq \lambda$, then so must the ultrapower $(M_1)^\lambda/\mathcal{D}$. In this paper, building on the author's prior work [11] [12] [13], we show that the relative complexity of first-order theories in Keisler's sense is reflected in the relative graph-theoretic complexity of sequences of hypergraphs associated to formulas of the theory. After reviewing prior work on Keisler's order, we present the new construction in the context of ultrapowers, give various applications to the open question of the unstable classification, and investigate the interaction between theories and regularizing sets. We show that there is a minimal unstable theory, a minimal TP_2 theory, and that maximality is implied by the density of certain graph edges (between components arising from Szemerédi-regular decompositions) remaining bounded away from $0, 1$. We also introduce and discuss *flexible* ultrafilters, a relevant class of regular ultrafilters which reflect the sensitivity of certain unstable (non low) theories to the sizes of regularizing sets, and prove that any ultrafilter which saturates the minimal TP_2 theory is flexible.

1. INTRODUCTION AND STATEMENT OF THE PROBLEM

In 1967 [7] Keisler asked about the structure of the following relation on countable theories.

Definition 1.1. *Given countable theories T_1, T_2, say that:*

(1) $T_1 \trianglelefteq_\lambda T_2$ *if for any $M_1 \models T_1$, $M_2 \models T_2$, and $\mathcal{D}$ a regular ultrafilter on λ,*
if $M_2^\lambda/\mathcal{D}$ is λ^+-saturated then $M_1^\lambda/\mathcal{D}$ must be λ^+-saturated.
(2) (Keisler's order) $T_1 \trianglelefteq T_2$ *if for all infinite λ, $T_1 \trianglelefteq_\lambda T_2$.*

Despite striking early work in the stable case, the structure of "Keisler's order" on unstable theories, in particular theories with the independence property, has remained open. The development of stability theory for unstable theories has not produced corresponding insight into the unstable Keisler classes; it appears that a fundamentally different sort of phenomenon is at work.

This paper arises from two sets of results (which will be precisely stated below). First, it was shown in [11] that saturation of regular ultrapowers depends on the realization of φ-types. Second, following [12] one can associate to any formula a sequence of hypergraphs, called the *characteristic sequence*; it was shown in [12]-[13] that classification-theoretic complexity of φ is reflected in graph-theoretic complexity of this sequence of hypergraphs, e.g. edge distribution, edge density, and what was called persistence of configurations. Here we bring these results to bear on Keisler's order. Since an underlying goal of the paper is to present the naturalness and interest of the model-theoretic structure which Keisler's order brings to light, we have also attempted to give a reasonably self-contained survey of the milestones of prior work. References are given throughout.

2. Discussion of the problem

Let us first motivate the problem. Keisler's criterion of choice, saturation of regular ultrapowers, is natural for two reasons. First, when the ultrapower is regular, the degree of its saturation depends only on the theory and not on the saturation of the index models, Theorem B below. Second, ultrapowers are a natural context for studying compactness, and Keisler's order can be thought of as studying the fine structure of compactness by asking: what families of consistent types are realized or omitted together in regular ultrapowers? Thus the relative difficulty of realizing the types of T_1 versus those of some T_2 in regular ultrapowers gives a measure of the combinatorial complexity of the types each T_i is able to describe.

Shelah in the 1970s gave a further series of beautiful and surprising results showing deep links between Keisler's order and the underlying structure of first-order theories. The dividing lines will be familiar to most readers. (More precise references, and some definitions, are given in the next section.)

Theorem A. (Summary of known results on the structure of Keisler's order)

(1) *The theories without the finite cover property (FCP) are minimal in Keisler's order.* [14]
(2) *There is a dividing line between theories with and without FCP.* [7], [14]
(3) *The stable theories with FCP are an equivalence class in Keisler's order.* [14]
(4) *There is a dividing line between stable and unstable theories.* [14]
$$\vdots$$
(5) *There is a maximum class, namely, the theories which are λ^+-saturated iff the ultrafilter is λ^+-good. The strict order property is sufficient for maximality. In fact, SOP_3 is sufficient for maximality; however, the model-theoretic identity of the maximal class is not known.* [7], [14], [16], [17], [4]

Notice the coarseness of the order. Stability is a classic model-theoretic frontier, but the finite cover property crosscuts all of its usual refinements. Recent work of Shelah, Usvyatsov and Džamonja has weakened the sufficient condition for maximality from SOP to the strong order property SOP_3 (see [16] Theorem 2.9, [17] Theorem 3.5, as well as the introduction of [4]). However, the identity of the maximal class, as well as the structure of the order on unstable theories with the independence property, has remained open. Notice also that stability, fcp and strict order are all properties of formulas. In fact this is paradigmatic: the Keisler order reduces to the study of types in a single formula.

Theorem 2.1. (Malliaris [11]) *Let $\mathcal{D}$ be a regular ultrafilter on $\lambda \geq \aleph_0$ and let T be a countable theory, $M \models T$. Then $M^\lambda/\mathcal{D}$ is λ^+-saturated iff $M^\lambda/\mathcal{D}$ realizes all φ-types over sets of size $\leq \lambda$ for all $\mathcal{L}$-formulas φ.*

This mirrors a key move of "localization" in stability theory (T unstable iff it contains an unstable formula). The result suggests that it suffices to study the complexity of incidence graphs on individual formulas, which will be discussed in §4 below. By way of motivating this work, we first discuss in broad strokes (both coarse and subtle) what forty years of work on Keisler's order has brought to light.

3. WHAT IS KNOWN

The main results of this section are due to Keisler [7] and Shelah [14], Chapter VI; some are folklore about ultrapowers. An ultrapower is a reduced product where equivalence is computed modulo an ultrafilter $\mathcal{F}$ and the index models are taken to be the same, see for instance the article of Kochen [8] or the books of Comfort and Negrepontis [2] or Shelah [14], Chapter VI.

Definition 3.1. *(Regular ultrapowers) Let $\lambda \geq \kappa$ be infinite cardinals.*

(1) *A κ-regularizing set is any $X \subset \mathcal{P}(\lambda)$, $X = \langle X_i : i < \kappa \rangle$, satisfying:*
  - *X has the finite intersection property, i.e. for any $\sigma \in \mathcal{P}_{\aleph_0}(\kappa)$, $\bigcap_{i \in \sigma} X_i \neq \emptyset$*
  - *for any $t \in \lambda$, $|\{i < \kappa : t \in X_i\}| < \aleph_0$.*
(2) *An ultrafilter $\mathcal{D}$ on λ is κ-regular if it contains a κ-regularizing set. $\mathcal{D}$ is regular if it is λ-regular.*
(3) *A set $A \subset N := M^\lambda/\mathcal{D}$ is called small if $|A| \leq \lambda$. Any $p \in S(A)$ is a small type if A is small.*

For the remainder of this paper $\mathcal{D}$ will denote a regular ultrafilter on $\lambda \geq \aleph_0$.

Regular ultrapowers are "flat" in the sense that any small set in the ultrapower is contained in a product of finite subsets of the index models (see Definition 3.4). As a consequence, the amount of saturation in the ultrapower does not depend on the level of saturation of the index model, but only on its theory T.

Theorem B. *(Keisler [7] Corollary 2.1 p. 30; see also Shelah [14].VI.1) Suppose that $M_0 \equiv M_1$, the ambient language is countable, and $\mathcal{D}$ is a regular ultrafilter on λ. Then $M_0{}^\lambda/\mathcal{D}$ is λ^+-saturated iff $M_1{}^\lambda/\mathcal{D}$ is λ^+-saturated.*

This ensures that the quantification over all models of a theory in the Keisler order is justified.

Fact 3.2. *Let $\lambda \geq \aleph_0$ be an infinite cardinal. Regular ultrafilters on λ always exist.*

Proof. Uniquely in this proof, let us write I for the index set of the filter to avoid confusion. Let $f : \mathcal{P}_{\aleph_0}(\lambda) \to I$ be a bijection. For $\eta \in \lambda$, define $X_\eta = \{i \in I : \eta \in f^{-1}(i)\}$. Then $\{X_\eta : \eta \in \lambda\}$ is a regularizing set of size λ. X has the finite intersection property and does not contain $\emptyset$, so it can be extended to a nonprincipal ultrafilter by Zorn's lemma. $\square$

3.1. **Distributions.** Let us spell out the "distribution" of a type $p \in S(A)$, $A \subset N$ small, across the index models M as a way of illustrating how types are realized or omitted in regular ultrapowers.

Throughout this paper, we will suppress overlines, writing $\varphi(x; y)$ when $\ell(x), \ell(y)$ need not be 1.

Convention 3.3. *For the purposes of this paper,*

  - *$\mathcal{D}$ is a regular ultrafilter on λ, and $t \in \lambda$ is an element of the index set.*
  - *T is a countable theory, M is a model of T and $N := M^\lambda/\mathcal{D}$.*
  - *"Small" means of cardinality $\leq \lambda$.*
  - *Write $M[t]$ for the model M considered as the index model at index t.*
  - *For each ultrapower $M^\lambda/\mathcal{D}$, fix a lifting: $[a]_\mathcal{D} \in M^\lambda/\mathcal{D} \mapsto a \in M^\lambda$. The parameter $a \in N$ is thus identified with $(\prod_{t<\lambda} a[t])/\mathcal{D}$, and the projections of elements and sets $a[t] \in M[t]$, $X[t] \subset M[t]$ are well defined.*

Definition 3.4. *(Distributions) Fix $T, M \models T, \lambda \geq \aleph_0$, $\mathcal{D}$ regular on λ, $N := M^\lambda/\mathcal{D}$, and a small type $p \in S(A)$, $A \subset N$. A distribution $d : \mathcal{P}_{\aleph_0}(p) \to \mathcal{D}$ of the type p is a monotonic assignment of each finite subset of p to an element of $\mathcal{D}$, such that d refines the Łoś map and the image of d is a regularizing set. More precisely:*

(1) *For each $u \in \mathcal{P}_{\aleph_0}(p)$, $d(u) \in \mathcal{D}$.*
(2) *d is monotonic, that is, for all $\varphi_i, \varphi_j \in p$, $d(\{\varphi_i, \varphi_j\}) \subset d(\{\varphi_i\})$.*
(3) *For each $u \in \mathcal{P}_{\aleph_0}(p)$, d is a refinement of the Łoś map. That is,*

$$d(u) \subset \left\{ t < \lambda : M[t] \models \exists x \left(\bigwedge_{\varphi(x;a) \in u} \varphi(x; a[t]) \right) \right\}$$

(4) *For each $t \in \lambda$, $|\{u : t \in d(u)\}| < \aleph_0$.*

Observation 3.5. *For any small type p in a regular ultrapower, a distribution exists.*

Proof. Let us sketch a possible construction:

- Write p as $\{\varphi_i(x; a_i) : i < \lambda\}$, where each φ_i is a formula of $\mathcal{L}$ and the parameters a_i are from A.
- Let $d_0 : p \to \mathcal{D}$ be the Łoś map, i.e. $\varphi_i(x; a_i) \mapsto \{t < \lambda : M[t] \models \exists x(\varphi_i(x; a_i[t]))\}$.
- Let $X = \langle X_i : i < \lambda \rangle$ be a regularizing set in $\mathcal{D}$. Define $d_1 : p \to \mathcal{D}$ by $d_1(\{\varphi_i\}) = d_0(\{\varphi_i\}) \cap X_i$.
- To finish, we extend the definition to $d : \mathcal{P}_{\aleph_0}(p) \to \mathcal{D}$ by:

$$d(\{\varphi_{i_1}, \dots \varphi_{i_n}\}) := \left\{ t : \bigwedge_{k \leq n} t \in d_1(\{\varphi_{i_k}\}) \right\} \cap \left\{ t : M[t] \models \exists x \bigwedge_{k \leq n} \varphi_{i_k}(x; a_{i_k}[t]) \right\}$$

The first set is equal to $\bigcap_{k \leq n} d_1(\{\varphi_{i_k}\})$, which is large because $\mathcal{D}$ has the finite intersection property; the second is large by Łoś' theorem. $\square$

Remark 3.6. *Paired with Observation 3.10 below, this basic construction shows the combinatorial issues at stake in realizing, or omitting, a small type in a regular ultrapower. Namely, let $\{\varphi_{i_1}\}, \dots \{\varphi_{i_n}\}$ be singleton elements of $\mathcal{P}_{\aleph_0}(p)$ whose images under d all contain t. Then for all $k \in \{1, \dots n\}$,*

$$(1) \qquad\qquad M[t] \models \exists x \left(\varphi_{i_k}(x; a_{i_k}[t]) \right)$$

But unless $t \in d(\{\varphi_{j_1}, \dots \varphi_{j_r}\})$, for $\{j_1, \dots j_r\} \subset \{i_1, \dots i_n\}$, it need not be the case that:

$$(2) \qquad\qquad M[t] \models \exists x \left(\bigwedge_{\ell \leq r} \varphi_{j_\ell}(x; a_{j_\ell}[t]) \right)$$

The important class of distributions which satisfy $d(u) \cap d(v) = d(u \cup v)$ are called multiplicative, Definition 3.7. A multiplicative distribution of p exists just in case p is realized; see Observation 3.10.

3.2. **Some examples.** Remarks on some typical cases:

Example 1: Algebraically closed fields. Let M be an algebraically closed (hence infinite) field, $N := M^\lambda/\mathcal{D}$. For some small $A \subset N$ small let $p(x) \in S(A)$ be the type describing an element which does not satisfy any nontrivial polynomial with coefficients in A. So we can write $p(x) = \{\neg f_i(x; \overline{a}_i) : i < \lambda\}$, where each f is a finite conjunction of polynomial equations with coefficients in the finite set $\overline{a}_i$. A distribution d assigns finitely many of the f_i to each index model $M[t]$. We look in $M[t]$ for an element $c[t]$ satisfying the finitely many relevant $\neg f_i(x; \overline{a_i}[t])$, which will always exist. Then $c := \prod_{t<\lambda} c[t]/\mathcal{D}$ will satisfy the type p, because it avoids each f_i on the large set $d(f_i)$, by construction. This gives an easy proof that in any regular ultrapower of M, the transcendence degree over the prime field will be at least λ^+. Indeed, $M^\lambda/\mathcal{D}$ will always be λ^+-saturated for any infinite λ and regular ultrafilter $\mathcal{D}$ on λ.

Example 2: The random (Rado) graph. The language contains equality and a binary edge relation R. The axioms say that the graph is infinite, and for each set of $2n$ distinct elements $y_1, \ldots y_n$, $z_1, \ldots z_n$,

$$\exists x \left(\bigwedge_{i \leq n} xRy_i \wedge \bigwedge_{j \leq n} \neg xRz_j \right)$$

Again, let $M \models T$ and $N := M^\lambda/\mathcal{D}$. By quantifier elimination, a small type p in N can be written as $p = \{xRa_i \wedge \neg xRb_i : i < \lambda\}$. Let d be a distribution, so:

$$t \in d(\{xRa_i \wedge \neg xRb_i\}) \implies M[t] \models \exists x \, (xRa_i[t] \wedge \neg xRb_i[t])$$

The distribution may fail to be multiplicative because of "collisions" between parameters in the index models. That is:

$$M[t] \models \quad \exists x \left(\bigwedge_{i \leq n} xRa_i[t] \wedge \neg xRb_i[t] \right)$$

$$\iff \qquad\qquad M[t] \models \left\{ \bigcup_{i \leq n} a_i[t] \right\} \cap \left\{ \bigcup_{j \leq n} b_j[t] \right\} = \emptyset$$

Let us write $A[t]$ for the set $\{a_i[t] : t \in d(u), u \in \mathcal{P}_{\aleph_0}(p), xRa_i \wedge \neg xRb_i \in u\} \subset M[t]$, and likewise for $B[t]$. The type p will be realized just in case there exists a distribution d in which, for almost every index model $M[t]$, $A[t] \cap B[t] = \emptyset$.

Equivalently, an ultrafilter will be able to realize all small types in models of the random (Rado) graph iff for any pair of disjoint small sets $A, B \subset N$ it is possible to expand each index model by a new monadic predicate $X[t]$ so that $X := \prod_t X[t]/\mathcal{D}$ separates A and B.

We will see that there are ultrafilters which fail to have this property; in fact there is a Keisler equivalence class strictly between algebraically closed fields and the random graph.

Example 3: The finite cover property. Let M be the prime model for the following theory: the language contains equality and a binary equivalence relation E, and the theory says that E has a single class of size n for each $n < \omega$. Let $N = M^\lambda/\mathcal{D}$.

Let $A \subset N$ be an infinite set contained in a single E-equivalence class, and $p(x, A) := \{xEa \wedge x \neq a : a \in A\}$. Let $B \subset N$ be a set of representatives of distinct E-equivalence classes, and

$q(x, B) := \{\neg xEb : b \in B\}$. Then it is easy to see that the type q is always realized, as any finite fragment assigned to $M[t]$ by a distribution is satisfiable. For d a distribution, write $A[t]$ for the set $\{a_i[t] : (\exists u \in \mathcal{P}_{\aleph_0}(p))(t \in d(u) \ \land \ xEa_i \land x \neq a_i \ \in u)\}$. For p (see first line of paragraph) the following are equivalent:

(1) There exists $c \in N$ such that $c \models p$.
(2) Fixing some $a_0 \in A$, $\{x \in N : xEa_0 \land x \neq a_0\} \supsetneq A$.
(3) For any $a \in A$, there exists a distribution $d' : p \to \mathcal{D}$ whose associated $A[t]$ satisfy, almost everywhere,
    $|A[t]| < |\{x \in M[t] : xEa[t] \land x \neq a[t]\}|$.

In fact, extrapolating from condition (2) one can see that N realizes all such types over small sets A' iff the size of every nonstandard equivalence class is large, i.e. $\geq \lambda^+$. We shall see that the size of pseudofinite sets is sensitive to the ultrafilter $\mathcal{D}$ in Theorem D below. Thus regular ultrapowers of theories with the finite cover property will not always be λ^+-saturated.

3.3. **Multiplicative refinements.** Let us describe a class of ultrafilters, the *good* ultrafilters, which are subtle enough to untangle any type in a countable language. Because there exists a first order theory whose $\mathcal{D}$-ultrapowers are λ^+-saturated iff $\mathcal{D}$ is good (Observation C), we see that there must be a *maximum*, rather than simply maximal, class in the Keisler order.

Definition 3.7. *(Multiplicativity)*

(1) *A function $f : \mathcal{P}_{\aleph_0}(\lambda) \to \mathcal{D}$ is* multiplicative *if $f(u) \cap f(v) = f(u \cup v)$, and* monotonic *if $f(u \cup v) \subset f(u)$.*
(2) *If every monotonic $f : \mathcal{P}_{\aleph_0}(\lambda) \to \mathcal{D}$ has a multiplicative refinement, then $\mathcal{D}$ is called λ^+-good.*

The existence of λ^+-good ultrafilters on λ is a theorem of Kunen [9].

Fact 3.8. *Every nonprincipal ultrafilter is $\aleph_1$-good. When T is countable, this means that every nonprincipal ultrapower of $M \models T$ is $\aleph_1$-saturated.*

See for instance [14].VI.2.

Definition 3.9. *A distribution $d : p \to \mathcal{D}$ is* accurate *if for each index $t < \lambda$ and each finite subset $\{\varphi_{i_1}, \dots \varphi_{i_n}\} \subset D(t) := \{\varphi_j : t \in d(\varphi_j)\}$, we have that $t \in d(\{\varphi_{i_1}, \dots \varphi_{i_n}\})$ iff $M[t] \models \exists x \bigwedge_{k \leq n} \varphi_{i_k}$.*

Observation 3.10. *Choose $T, M, \lambda, \mathcal{D}$, $N := M^\lambda/\mathcal{D}$, $A \subset N$ small, $p \in S(A)$. Then the following are equivalent.*

(1) *Some distribution d of p has a multiplicative refinement.*
(2) *Every accurate distribution d of p has a multiplicative refinement.*
(3) *The type p is realized in N.*

Proof. (2) $\implies$ (1) The construction of Observation 3.5 above shows that accurate distributions always exist.

(1) $\implies$ (3) Let d' be the multiplicative refinement. Then the formulas $\varphi_1, \dots \varphi_n$ assigned to index model $M[t]$ have a common realization in that model, because multiplicativity implies that if $\bigwedge_{i \leq n} (t \in d'(\{\varphi_i\}))$ then $t \in d'(\{\varphi_1, \dots \varphi_n\})$. Let $\alpha[t]$ be some such common realization in $M[t]$,

and set $\alpha := \prod_{t<\lambda} \alpha[t]/\mathcal{D}$. Now for each formula $\varphi(x;c) \in p$, we have that $\varphi(\alpha;c)$ by Łoś' theorem, so $\alpha \models p$.

(3) $\implies$ (2) Let d be some accurate distribution. Suppose that p is realized by the element α. For $v \in \mathcal{P}_{\aleph_0}(p)$, $v = \{\varphi_{i_1}, \dots \varphi_{i_k}\}$ set

$$d'(v) := \left\{ t : M[t] \models \bigwedge_{\ell \leq k} \varphi_{i_\ell}(\alpha[t]; a_{i_\ell}[t]) \right\} \cap d(v)$$

Now d' refines d by construction. Suppose $u, v \in \mathcal{P}_{\aleph_0}(p)$. Then $t \in d'(u) \cap d'(v)$ implies $\alpha[t]$ is a common witness, in $M[t]$, to both sets of formulas. A fortiori $t \in d(u) \cap d(v)$, and because there is a common witness and we assumed d was accurate, $t \in d(u \cup v)$. Thus $t \in d'(u \cup v)$ so d' is multiplicative. $\square$

Corollary 3.11. *If $\mathcal{D}$ is λ^+-good and $Th(M)$ is countable then $M^\lambda/\mathcal{D}$ is λ^+-saturated.*

That is to say, we have a way of assigning finitely many of the formulas of a small type to each index model in such a way that the finitely many formulas assigned to $M[t]$ have a common realization in $M[t]$.

Theorem C. *(Keisler [7] Theorem 1.4 p. 27) Let M be the model whose elements are the finite subsets of ω. The language is $\{=, \subseteq\}$, interpreted in the natural way. Let $T = T(\mathcal{P}_{\aleph_0}, \subseteq) := Th(M)$. Let $\varphi(x;y) = x \subset y$. Suppose that the ultrafilter $\mathcal{D}$ on λ is not λ^+-good. Then there is a small φ-type omitted in $N = M^\lambda/\mathcal{D}$.*

Proof. Let $f : \mathcal{P}_{\aleph_0}(\lambda) \to \mathcal{D}$ be a monotonic function with no multiplicative refinement. We would like to find elements $\{a_i : i \in \lambda\} \subset N$ such that f is an accurate distribution of a consistent partial φ-type $p = \{x \subseteq a_i : i < \lambda\}$. It would be enough to define $a_i[t] \in M[t]$ when $\{i\} \in f^{-1}(t)$ so that $M[t] \models \exists x(\bigcap_{j\leq k} x \subset a_{i_j}[t])$ just in case $t \in f(\{i_1, \dots i_k\})$. In other words, in the index model $M[t]$, we choose finitely many sets $a_i[t]$ so that the pattern of incidence is precisely that described by f, and set all other $a_j[t] = \emptyset$. The existence of such a_i is clearly consistent with the theory, by monotonicity of f. Set $a_i := \prod_{i<\lambda} a_i[t]$ to finish; the distribution is accurate by construction, so we are done. $\square$

Corollary 3.12. *A necessary and sufficient condition for maximality in the Keisler order is: for all λ, $M \models T$, $M^\lambda/\mathcal{D}$ is λ^+-saturated iff $\mathcal{D}$ is λ^+-good.*

Proof. Sufficiency is Corollary 3.11. Necessity follows from Theorem C: if the ultrafilter is not λ^+-good, then we have seen that there is a theory whose $\mathcal{D}$-ultrapowers are not λ^+-saturated. $\square$

3.4. **Filters and theories.** The interaction between ultrafilters and theories, in both directions, is both coarse and subtle, as subsequent sections of the paper will show. This section discusses several examples of dimensions in first-order theories to which regular ultrafilters are sensitive, by way of mapping the large territory between the minimum and maximum Keisler class, i.e. between ω^+- and λ^+-goodness.

3.5. **Cardinalities of sets.**

Fact 3.13. *Let M be a model of signature $\mathcal{L}$, $\mathcal{L}_0 \subset \mathcal{L}$ and $\mathcal{D}$ an ultrafilter on λ. Then*

$$\left(M^\lambda/\mathcal{D} \right)|_{\mathcal{L}_0} = (M|_{\mathcal{L}_0})^\lambda /\mathcal{D}$$

Corollary 3.14. *Let $\mathcal{D}$ be any ultrafilter on λ, not necessarily regular, with M countable and $N := M^\lambda/\mathcal{D}$. Let $X[t] \subset M[t]$ be infinite, and set $X := \prod_{t<\lambda} X[t]/\mathcal{D} \subset N$. Then $|X| = |N|$.*

Proof. Let $\mathcal{L}$ be the expansion of the language to include a new function symbol $\{f\}$, interpreted almost everywhere as a bijection $f[t] : X[t] \to M[t]$. Then $f = \prod_t f[t]/\mathcal{D}$ will remain a bijection in N by Łoś' theorem. $\qquad\square$

For pseudofinite subsets of N, the story is different.

Definition 3.15. ([14] Definition III.3.5) *Let $\mathcal{D}$ be a regular ultrapower on λ.*

$$\mu(\mathcal{D}) := \min\left\{ \prod_{t<\lambda} n[t]/\mathcal{D} \ : \ n[t] < \omega, \ \prod_{t<\lambda} n[t]/\mathcal{D} \geq \aleph_0 \right\}$$

be the minimum value of the product of an unbounded sequence of cardinals modulo $\mathcal{D}$.

Theorem D. (Shelah, [14].VI.3.12) *Let $\mu(\mathcal{D})$ be as in Definition 3.15. Then for any infinite λ and $\nu = \nu^{\aleph_0} \leq 2^\lambda$ there exists a regular ultrafilter $\mathcal{D}$ on λ with $\mu(\mathcal{D}) = \nu$.*

This leads to obvious failures of saturation in theories which contain a parametrized family of sets of size n for all n (the finite cover property), because an ultrapower modulo $\mathcal{D}$ will contain nonstandard elements of the family whose size is precisely $\mu(\mathcal{D})$:

Corollary 3.16. (see Keisler [7] Corollary 4.2a p. 40) *Let M be the standard model of the finite cover property, i.e. an equivalence relation E with an equivalence class of size n for each $n < \omega$, and let $\mathcal{D}$ be a regular ultrafilter on λ. Then $M^\lambda/\mathcal{D}$ is λ^+-saturated iff $\mu(\mathcal{D}) \geq \lambda^+$.*

A surprising theorem of Shelah shows that all failures of saturation in ultrapowers of stable theories come from pseudofinite sets which are too small. This will establish the identities of the only two known equivalence classes in the Keisler order: T without the finite cover property, and T stable with the finite cover property (Theorem A). We sketch the proof of this result:

Theorem E. (Shelah, [14] Theorem 5.1(2) p. 379) *If T is a countable stable theory, $M \models T$, $\mathcal{D}$ regular on λ and $\mu(\mathcal{D}) \geq \lambda^+$, then $M^\lambda/\mathcal{D}$ is λ^+-saturated.*

Proof. (Sketch) The proof rests on the following characterization of saturated models in stable theories: when T is stable, $N \models T$ is λ^+-saturated iff N is $\kappa(T)$-saturated and every maximal indiscernible set has size at least λ^+ (see [14]: Theorem 3.10 p. 107). Essentially, this is because any type $p \in S(C)$, $|C| \leq \lambda$ does not fork over a set C_0, $|C_0| < \kappa(T)$: so by $\kappa(T)$-saturation we can find a countable indiscernible set I of realizations of $p|_{C_0}$. Let $J \supset I$ be any indiscernible set extending I. Any element $a \in J$ which is free from C over C_0 will realize the unique nonforking extension of $p|_{C_0}$ to C, which is p. Such an a will exist if $|J| > |C|$.

In our case, T is countable so $\kappa(T) = \aleph_1$; nonprincipal ultrapowers are automatically $\aleph_1$-saturated (Fact 3.8). So it would suffice to show that every countable indiscernible set in a regular ultrapower N of a countable stable theory can be extended to an indiscernible set of size λ^+. Expanding the language slightly to code $\in$, Shelah shows that this is the same as the problem of showing that every pseudofinite set in the ultrapower has cardinality $\geq \lambda^+$. $\qquad\square$

Note the similarity (in a different context) to the Farah-Hart-Sherman proof, Theorem 5.6(1) of [5].

3.6. **Cuts above** ω. A natural way of characterizing ultrafilters is to "sound out" the depth of their multiplicativity, using sufficiently complex theories. Let $M = (\omega, <)$ be a discrete linear order. Define the *lower cofinality of ω modulo $\mathcal{D}$*, written $\mathrm{lcf}(\omega, \mathcal{D})$, to be the reverse cofinality of the set of elements above the image of the diagonal embedding of ω in $(\omega, <)^{\lambda}/\mathcal{D}$, i.e. the coinitiality of ω in the ultrapower.

Remark 3.17. *If* $\mathrm{lcf}(\omega, \mathcal{D}) \leq \lambda$*, then by [14] Theorem 4.8 p. 379 any formula with the order property will omit a type. In combination with the next theorem, this gives a dividing line between stable and unstable theories.*

Finally, the following theorem shows that filters with large $\mu(\mathcal{D})$ in the sense of Definition 3.15 can still have small $\mathrm{lcf}(\omega)$, i.e., that it is possible to construct filters in which the sizes of certain nonstandard integers vary considerably, i.e. in which the size of any pseudofinite set is large but the size of the cut above ω is small. So Keisler's order is able to see a dividing line between stable and unstable theories, because stable theories require pseudofinite sets to be large in order to be saturated, but are indifferent to the sizes of orders, which they cannot code.

Theorem F. *(Shelah, [14] Theorem VI.3.12, pps. 357-367) For any infinite λ, $\nu = \nu^{\aleph_0} \leq 2^{\lambda}$ and $\aleph_0 < \kappa \leq \nu$ there exists a regular ultrafilter $\mathcal{D}$ on λ such that $\mathrm{lcf}(\omega, \mathcal{D}) = \kappa$ and $\mu(\mathcal{D}) = \nu$.*

It is notable that all the work described is in ZFC. The sole result with an additional set-theoretic hypothesis known to the author is [14] Theorem VI.3.10, i.e. assuming MA + not CH there is a filter which, roughly speaking, saturates certain small ultrapowers of models of the random graph but is not good. So it is consistent that there are further dividing lines.

4. The approach via characteristic sequences

The review of prior work points to the following basic issue. Let $N = M^{\lambda}/\mathcal{D}$ be a regular ultrapower, and $Th(M)$ countable. Then, as explained in Remark 3.6, a type $p \in S(A)$, for $A \subset N$ small, can be thought of as a monotonic function $f : \mathcal{P}_{\aleph_0}(\lambda) \to \mathcal{D}$; p will be realized iff f has a multiplicative refinement. By Theorem 2.1 it suffices to consider the case where p is a type in a given formula φ. So the existence of a multiplicative refinement is a question about the strength of the filter, but perhaps more importantly it is a question about what such functions the formula φ admits. How can instances of φ crosscut each other in a given index model? The objects to look at are incidence graphs on the parameter space of φ.

Recall that throughout this paper, we will write $\varphi(x; y)$ when $\ell(x), \ell(y)$ need not be 1.

Definition 4.1. [12] *Fix a formula $\varphi(x; y)$. For each $n < \omega$, let*

$$P_n(y_1, \ldots y_n) = \exists x \bigwedge_{i \leq n} \varphi(x; y_i)$$

Then the characteristic sequence *of φ with respect to the background theory T is $\langle P_n : n < \omega \rangle$.*

Convention 4.2. *We will be interested in consistent partial φ-types, and so would like both positive and negative instances of φ. By convention, if this is not redundant, replace φ by $\varphi(x; y, z) = \varphi(x; y) \wedge \neg\varphi(x; z)$ before computing the characteristic sequence, or by $\varphi(x; y, z, w) = (\varphi(x; y) \wedge z = w) \vee (\neg(\varphi(x; y) \wedge z \neq w)$. Though we will be extracting configurations, this process is usually quite robust, and the particular choice of coding usually will not matter.*

We will often identify the predicates P_n with their interpretation in some monster model; write P_n^M for $\{a_1 \ldots a_n \in M : M \models P_n(a_1, \ldots a_n)\}$ when the identity of the model M is important. Note that the P_n are definable in the background theory T, so e.g. compactness applies. (This is explicitly spelled out in [12] §2.)

In the characteristic sequence, φ-types appear as complete graphs:

Observation 4.3. [12] *Suppose $A \subset P_1$. If $A^n \subset P_n$ for all n, call A a P_∞-complete graph. Then A is a P_∞-complete graph iff $\{\varphi(x; a) : a \in A\}$ is a consistent partial type.*

We will often call P_∞-complete graphs *positive base sets* to emphasize this correspondence.

Definition 4.4. *Let $\langle P_n \rangle$ be a characteristic sequence. By definition, for any $n < k < \omega$,*

$$\forall y_1, \ldots y_k \left(P_k(y_1, \ldots y_k) \to \forall z_1, \ldots z_n \left(\{z_1, \ldots z_n\} \subset \{y_1, \ldots y_k\} \to P_n(z_1, \ldots z_n) \right) \right)$$

Conversely, if for some $k < \omega$ it is true that for all $y_1, \ldots y_m$, $P_m(y_1, \ldots y_m)$ iff P_k holds on all k-element subsets of $\{y_1, \ldots y_m\}$, we say that the sequence has support k, or if k is not specified, that it has finite support.

For any model M and $C \subset P_1^M$, not necessarily definable, say that the sequence $\langle P_n \rangle$ has support k on C (or: finite support) if this holds when the $y_1, \ldots y_m$ are chosen from C. See Remark 4.5.

Remark 4.5. *([12] Remark 2.7) The following are equivalent, for $(T, \varphi) \mapsto \langle P_n \rangle$:*

(1) *There is $k < \omega$ such that the sequence $\langle P_n \rangle$ has support k.*
(2) *φ does not have the finite cover property.*

4.1. **Translation to ultrapowers.** We now explain how this construction can be applied to the analysis of types in regular ultrapowers.

Definition 4.6. *Let $N = M^\lambda/\mathcal{D}$ and $1 \leq k < \omega$. A set $X \subset N^k$ is* induced *if it is equivalent modulo $\mathcal{D}$ to the product of its projections to the index models, i.e. if*

$$\left[\prod_t X[t] \right]_{\mathcal{D}} = X \qquad \text{(modulo } \mathcal{D}\text{)}$$

A predicate X is an induced predicate *if its interpretation in N^k is an induced set. "There exists an induced predicate P such that..." means: we may, assuming it is not redundant, expand $\mathcal{L}$ by adding a new predicate symbol P whose interpretation in N^k is an induced set with the desired property.*

Note that any $\mathcal{L}$-definable set is induced.

The induced sets play a key role in the analysis of types in ultrapowers. Because they come from the index models in a direct way, the induced sets are typically very large (the dimensional invariants of $\mathcal{D}$-pseudofinite sets, such as $\mu(\mathcal{D})$, apply). In particular, they are much larger ($\geq \lambda^+$) than the small ($\leq \lambda$) positive base sets under analysis. At the same time, Fact 3.13 means that induced sets are important carriers of structural information. For instance, if one can show that some first-order property holds on an induced subset of (some power of) N, then Łoś' theorem applies.

Example 4.7. *Let M be the random (Rado) graph from page 5. Let $\varphi(x; y, z) = xRy \wedge \neg xRz$, let $\mathcal{D}$ be a regular ultrafilter on λ and let $N = M^\lambda/\mathcal{D}$. Notice that the characteristic sequence of φ has*

support 2. As before, we may assume that a small $A \subset P_1^N$ will be given by $\{(b,c) : b \in B, c \in C\}$, and A will be a positive base set just in case $B \cap C = \emptyset$. Suppose the type $p(x) = \{\varphi(x; b, c) : (b,c) \in A\}$ corresponding to A is realized by $\alpha \in N$. Then in each index model $M[t]$, the set $X[t] := \{(b,c) : M[t] \models \varphi(\alpha[t]; b, c)\} \subset P_1^{M[t]}$ is a positive base set in $M[t]$. So by Łos' theorem, $X := \prod_t X[t]/\mathcal{D}$ is a positive base set in N, and since we assumed $\alpha \models p$, $A \subset X$. Conversely, suppose there exists an induced set $X \supset A$ which is a P_∞-complete graph, and so a fortiori a P_2-complete graph. Refine the distribution of A so that $t \in f(a)$ iff $a \in X[t]$ and X is a P_2-complete graph in $M[t]$ (the second by Łos). Then a.e. the fragment of A assigned to $M[t]$ will be a P_2, therefore P_∞-complete graph, so will have a common witness $\alpha[t]$. Thus the type corresponding to A will be realized in N by $\alpha := \prod_t \alpha[t]/\mathcal{D}$.

That is, realizing the type corresponding to a positive base set A is a matter of finding an induced set containing A which is itself almost everywhere a P_∞-complete graph: the "interior" world of the first-order language which sees that A is a consistent type in N must align with the "exterior," or induced, world of the ultrapower which can say whether or not A *behaves* a.e. as a consistent type in the index models.

Lemma 4.8. *The following are equivalent for a positive base set $A \subseteq N := M^\lambda/\mathcal{D}$.*

(1) *The type $p := \{\varphi(x; a) : a \in A\}$ corresponding to A is realized.*
(2) *There exists a distribution $d : A \to \mathcal{D}$ such that almost everywhere, $A[t]$ is a $P_{m[t]}$-complete graph, where $m[t] = |\{a[t] : t \in d(a)\}|$, i.e. the cardinality of the image of the finite piece of A assigned by d to index t.*

However, note that these conditions are not *equivalent to finding an induced P_∞-complete graph containing A. One needs instead to find an induced set X containing A which is almost everywhere a P_∞-complete graph.*

Proof. For the crucial last line of the statement, suppose that there is some induced X with $A \subset X \subset P_1^N$ and X a P_∞-complete graph in N. We know that for each n, the set on which X is P_n-complete is large, but this need not coincide with the set on which X has n elements. [For example of a strong failure of alignment, see the section on flexible filters below.]

(2) $\to$ (1): If the $m[t]$-tuple of elements represented in each index model is in $P_{m[t]}$, this says exactly that corresponding instances of φ have a common witness $\alpha[t]$. Then $\alpha := \Pi_t \alpha[t]/\mathcal{D}$ realizes the type p, by definition of distribution.

(1) $\to$ (2): By Observation 3.10, any realized type p has a distribution which is multiplicative. This says precisely that a.e., the $r[t]$ elements $a_{i_1}, \ldots a_{i_{r[t]}}$ assigned to the index model $M[t]$ form a $m[t]$-complete graph, where $m[t]$ is the size of the image $\{ a_j[t] : j \in \{i_1, \ldots i_{r[t]}\} \}$. $\qquad\square$

Remark 4.9. *The project is therefore to analyze the comparative complexity of positive base sets A, A' (in their respective characteristic sequences) by investigating:*

(a) *The relative difficulty of finding, for each $n < \omega$, an induced complete n-graph X_n containing A. This is a question about the complexity of P_n as a hypergraph 'near' A.*
(b) *The relative difficulty of putting these countably many predicates together in order to produce a realization of the type, as the statement of the previous lemma suggests.*

4.2. **A result about instability.** Investigations into the relationship described in Remark 4.9(a), between the complexity of hypergraphs in the characteristic sequence and the classification-theoretic

complexity of the corresponding formula φ, were first carried out in [12]. That analysis did not take place in ultrapowers, so induced sets were not available. Instead, the investigation centered on definable restrictions of P_1, and one of the recurrent themes of those results were dividing lines of the form: either it is possible to find a restriction of P_1 of a certain form containing the positive base set A in which [a certain class of graph-theoretic configurations are absent], or else *every* restriction of that form contains many such configurations (so there is a certain genericity). In particular, we will use a theorem from that paper to show that the random graph is Keisler-minimal among unstable theories. For completeness, we first define "restriction of a certain form." Several motivating examples are given in the Appendix. The idea is to work in the graph language, i.e. by conjoining P_1 with some finite set of positive instances of formulas P_n.

Definition G. ([12], [13]) *Given a characteristic sequence $\langle P_n \rangle$, a* localization *is a definable subset of P_1 given by a function $f : m \to \omega \times \mathcal{P}_{\aleph_0}(B)$ where $m < \omega$ and writing $f(i) = (r_i, \beta_i)$, where $\beta_i = b_1^i, \ldots b_{r_i}^i$, we have:*

$$P_1^f(y) := \bigwedge_{i \leq m} P_{r_i+1}(y, b_1^i, \ldots b_{r_i}^i)$$

We will always assume that a localization is nontrivial, meaning that it contains at least one infinite positive base set. Given a positive base set A, a localization around A *is simply a localization containing A.*

"After localization, [X holds]" means "for any positive base set A in the parameter space of φ, there exists a localization of P_1 which contains A in which [X holds]".

In the language of the gloss given above, the next theorem answers the question "when, for any given n, is it possible to localize around a given positive base set so that all missing P_n-edges disappear?" Answer: precisely when φ does not have the independence property.

Theorem H. (Theorem 6.17 of [12], as reworded in [13]) *Let φ be a formula of T and $\langle P_n \rangle$ its characteristic sequence.*

(1) *Suppose $X \subseteq P_1$ is a localization and that φ does not have the independence property on parameters in X. Then for each positive base set $A \subset X$ and each $n < \omega$, there is a further localization $Y \subset X$ such that $A \subset Y$ and $Y^n \subset P_n$, i.e. Y is a P_n-complete graph.*

(2) *Suppose $X \subseteq P_1$ is a localization and that φ has the independence property on parameters in X. Then for all $n < \omega$, there are elements $z_1, \ldots z_n \in X$ such that $\neg P_n(z_1, \ldots z_n)$, i.e. X is not a P_n-complete graph for any n.*

Remark 4.10. *As stated, this is not a theorem about ultrapowers, and so we cannot* a priori *assume that the localizations mentioned there are defined using parameters which are present in an ultrapower at hand. For instance, this will not be the case when the background theory is "strictly NIP," i.e., not stable and without the independence property, unless the ultrafilter is good. However, when $\mu(\mathcal{D})$ is large, we can always define the required localizations within the ultrapower if φ is stable. The examples in the Appendix show that this statement is sharp. In other words, in the context of ultrapowers, this theorem reveals a dividing line at stability. By Remark 4.5, the other known Keisler-dividing line is also visible in the characteristic sequence. However, in this paper we will focus on unstable theories.*

In our context, Theorem H implies:

Claim 4.11. *Suppose that the background theory T does not have the strict order property. Let φ be unstable with characteristic sequence $\langle P_n : n < \omega \rangle$, computed with respect to $\varphi(x; y) \wedge \neg\varphi(x; z)$.*

If $M \models T$ is any $\aleph_1$-saturated model, then there is an array $I = \{c_i^t : i < \omega, t = 0, 1\} \subset P_1^M$ such that $A \subset I$ is a positive base set iff it contains no more than one element from each column, i.e. $c_i^t, c_j^s \in A \implies (i = j \to s = t)$. In particular, this will be true in any regular ultrapower.

Proof. The key idea of the proof of Theorem H is that either the localization is successful at some finite stage or else one is able to extract such a configuration, in which the many missing edges line up to form the array described. (One then has to correct for P_n versus P_2, which takes some work.) For us, this means that the countably infinite partial type describing the existence of such an array is consistent with the characteristic sequence, and thus there will be an array I in any $\aleph_1$-saturated model. $\square$

Remark 4.12. *Since we are free to work with any formula or sequence of the theory for the purposes of establishing complexity, it is often useful to choose one which gives the structure most clearly.*

Notice that the sequence restricted to the given array has support 2.

The result does not appeal to classification theory, but it is not difficult to see that if the characteristic sequence has such an array then φ must have the independence property, and conversely that if φ has the independence property then the sequence described has such an array ([12] Claim 3.7 and Remark 5.2). The background point is that certain key dividing lines, familiar as existential statements from classification theory, turn out to coincide with what might be called genericity in the graphs, that is, with the phenomenon of certain configurations (e.g. missing edges) persisting no matter how (definably) close one gets to a positive base set. The results of this article suggest, among other things, that this is the kind of unstable structure to which ultrafilters are sensitive. In the argument just given, one *could have* simply extracted the independence property from the hypothesis on the theory and cooked up an array, but this is an a posteriori argument: it works because we know a fair amount of classification theory (e.g. any unstable formula has either the independence property or some boolean combination of its instances has strict order [14] Theorem 4.7, p. 70). If one looks at the structure which arises organically from characteristic sequences, some familiar model-theoretic lines emerge and others do not (in keeping with our expectations of coarseness), but fortuitously, those which do arise, arise along with an alternate characterization in terms of genericity, or persistence, in the characteristic sequence.

For the remainder of the paper, we analyze unstable theories without SOP. Two recurring themes will be, first, the interaction of positive base sets (as in the array just described), and second, the mutual influence of types and regularizing sets.

5. A MINIMAL UNSTABLE THEORY

Suppose then that for some positive base set A and some $n < \omega$, it is not possible to localize around A to obtain a P_n-complete graph. Then every localization will contain many missing P_n-edges, and so, roughly speaking, generic configurations begin to appear.

Following [13], call the array I of Claim 4.11 above (the avatar of the independence property) an $(\omega, 2)$-*array*. It may be thought of as representing countably many independent choices. The content of the following lemma is that in order to saturate models of the random graph, it suffices to realize types whose elements come, almost everywhere, from an $(\omega, 2)$-array. We first spell out why this common configuration is enough:

Observation 5.1. *Suppose that for some countable T, $\mathcal{D}$ regular on λ, and φ a formula of T whose characteristic sequence $\langle P_n \rangle$ contains an $(\omega, 2)$-array in some $M \models T$, we have that $M^\lambda / \mathcal{D}$*

is λ^+-saturated. Let T', φ', M' be any other triple (a countable theory, a formula and a model of T') whose corresponding characteristic sequence also contains an $(\omega, 2)$-array in M'. Then if $A' \subset N' := (M')^\lambda / \mathcal{D}$ is a small positive base set which comes a.e. from the $(\omega, 2)$-array, the type corresponding to A' is realized in N'.

Proof. Let $d' : A' \to \mathcal{D}$ be a distribution whose image is a.e. contained in the $(\omega, 2)$-array. Write $C'[t]$ for the image $\{a_i'[t] : t \in d'(a_j')\} \cup \{b_j'[t] : t \in d'(b_j')\}$. Index model by index model, define a bijection f_t from $C'[t]$ into some isomorphic sub-hypergraph $C[t]$ of the $(\omega, 2)$-array of $M[t]$. By Łos' theorem, the induced bijection $f := \prod_t f_t / \mathcal{D}$ will send A' to some small positive base set $A \subset N$. By hypothesis, the type corresponding to A is realized. By Lemma 4.8, there is a distribution of A which is a.e. a P_∞-complete graph. Use the bijection f_t in each index model, and the hypothesis of isomorphism, to construct a distribution for A' which also has this property. $\qquad\square$

Remark 5.2. *Let M contain an $(\omega, 2)$-array W and let N be its regular ultrapower modulo some given $\mathcal{D}$. Suppose we had expanded each index model M by a predicate W interpreted on the set W. Then the sequence $\langle P_n \rangle$ restricted to W has support 2, and this remains true on the image of W in the ultrapower by Łos' theorem. In particular, this will be true of the sequence on any positive base set which comes a.e. from W, and will remain true even without a name for W, by Fact 3.13.*

Lemma 5.3. *Suppose $\mathcal{D}$ is a regular ultrafilter on λ, M is a model of the random graph, and M_1 is a model of any other unstable theory. If $M_1^\lambda / \mathcal{D}$ is λ^+-saturated, then $M^\lambda / \mathcal{D}$ is also λ^+-saturated. In other words, the random graph is minimal among unstable theories in Keisler's order.*

Proof. By Theorem A and Claim 4.11, any unstable theory is either known to be maximal (i.e. to require goodness) or known to have an $(\omega, 2)$-array in one of its associated characteristic sequences. So to establish minimality of the random graph among unstable theories, we will show that realizing types which come a.e. from an $(\omega, 2)$-array is sufficient for its saturation. Recall from Example 2 p. 5 and Example 4.7 that, by quantifier elimination, we need simply to show that given any two small disjoint sets A, B in $M^\lambda / \mathcal{D}$ and some fixed enumeration $\langle a_i : i < \lambda \rangle, \langle b_j : j < \lambda \rangle$ there is a distribution $d : A \cup B \to \mathcal{D}$ such that for almost all t, $A[t] \cap B[t] = \emptyset$. Write A_t for the set $\{a \in A : t \in d(a)\}$, as distinguished from the image $A[t] = \{a[t] : a \in A_t\}$, and likewise for B_t.

By Claim 4.11, there is a model $M_2 \equiv M_1$ which contains a countable $(\omega, 2)$-array. By Theorem B, $M_2^\lambda / \mathcal{D}$ is λ^+-saturated if and only if $M_1^\lambda / \mathcal{D}$ is. Thus, without loss of generality, we may assume M_1 contains a countable $(\omega, 2)$-array W. Let $N_1 := M_1^\lambda / \mathcal{D}$.

For each index t, let us define a map f_t from the labeled set of indices $\{(0, i) : i < \lambda, a_i \in A_t\} \cup \{(1, j) : j < \lambda, b_j \in B_t\}$ into W such that:

(1) For all $a_i, a_j \in A_t$,
    $P_2(f_t((0, i)), f_t((0, j)))$ and $f_t((0, i)) = f_t((0, j))$ iff $a_i[t] = a_j[t]$
(2) For all $b_i, b_j \in A_t$,
    $P_2(f_t((1, i)), f_t((1, j)))$ and $f_t((1, i)) = f_t((1, j))$ iff $b_i[t] = b_j[t]$
(3) For all $a_i \in A_t, b_j \in B_t$, $P_2(f_t((0, i)), f_t((1, j)))$ iff $a[t] \neq b[t]$

Since equality is an equivalence relation, this is easy to do. Partition $A_t \cup B_t$ into disjoint maximal sets S_ℓ such that each S_ℓ consists of elements whose images in $M[t]$ are equal. On each such S_ℓ map the indices from $S_\ell \cap A_t$ and $S_\ell \cap B_t$, respectively, to distinct elements c_s^0, c_s^1 in some given column of the array W. If $k \neq \ell$, let the images of the indices in S_ℓ and S_k be disjoint. Since W is infinite, there is always enough room.

Finally, let $f = \prod_t f_t/\mathcal{D}$ be the function induced on N_1 by the f_t. Then $\mathcal{A} := f(\{(0,i) : i < \lambda\} \cup \{(1,j) : j < \lambda\}$ will be contained in the product of the sets W. By construction, for any $a_i \in A, b_j \in B$ we have that $P_2^{N_1}(f((0,i)), f((1,j)))$ since $a_i \neq b_j$ almost everywhere. By Remark 5.2, $\mathcal{A}$ is therefore a (small) positive base set. By hypothesis, N_1 realizes the type corresponding to $\mathcal{A}$ and so, by Lemma 4.8, there is a distribution δ of $\mathcal{A}$ which is a.e. a complete graph. Since this distribution is a.e. in W by construction, it amounts to the choice, in each index model, of precisely one side of each of the conflicting pairs (that is, either the elements from A_t or those from B_t). In other words, the distribution of $A \cup B$ given by $a_i \mapsto \delta(f((0,i))), b_j \mapsto \delta(f((1,j)))$ will have the property that $A[t] \cap B[t] = \emptyset$ almost everywhere. $\qquad\square$

Remark 5.4. *Any unstable theory will contain a formula with the finite cover property ([14] Theorem II.4.2 p. 62). However, they may not play a crucial role. In order to realize types in the random graph, it suffices to consider the characteristic sequence of $\varphi(x; y, z) = xRy \wedge \neg xRz$, which has support 2.*

Moreover, realizing types in this sequence guarantees that stable types with the finite cover property will also be realized. If $\mathcal{D}$ is a regular ultrafilter on λ and $M^\lambda/\mathcal{D}$ is λ^+-saturated, where M is a model of the random graph, then $\mu(\mathcal{D}) \geq \lambda^+$ as stable theories are strictly above unstable theories in Keisler's order, Remark 3.17.

We now investigate the effect on regularizing sets.

Definition 5.5. (Sizes and true sets)

(1) *Let $X = \{X_i : i < \lambda\}$ be a regularizing set. For each index $t < \lambda$, we define $\sigma[t] := |\{i : t \in X_i\}|$. Define the size $\sigma(X) := \prod_t \sigma[t]/\mathcal{D}$.*

(2) *Let $S \subset M^\lambda/\mathcal{D}$ be a small set. S is* true *if there exists an assignment $f : S \to \mathcal{D}$ such that for any $t \in \lambda$, $t \in f(s) \cap f(s')$ implies $M[t] \models s[t] \neq s'[t]$.*

True sets may be thought of as a kind of image of a regularizing set (as the proof of Lemma 5.7 will illustrate):

Observation 5.6. *The following are equivalent for a regular ultrafilter $\mathcal{D}$ on λ and an unbounded sequence of finite cardinals $\{n_t\}_{t<\lambda}$: (1) There is a true set S of size λ and a distribution $d : S \to \mathcal{D}$ such that for almost all $t < \lambda$, $|\{s \in S : t \in d(s)\}| \leq n_t$. (2) $\mathcal{D}$ contains a regularizing set X with $\sigma(X)[t] \leq n_t$ a.e. In particular, true sets exist by the regularity of $\mathcal{D}$.*

Proof. Let $\langle X_i : i < \lambda \rangle$ be a regularizing set in $\mathcal{D}$, M any infinite model, and $M^\lambda/\mathcal{D}$ an ultrapower. We define a corresponding set $A = \{a_i : i < \lambda\} \subset N$ as follows. In the index model $M[t]$, define the elements $a_i[t]$ (for $t \in X_i$) to be pairwise distinct elements of $M[t]$. Let $a_i := \prod_t a_i[t]/\mathcal{D}$. Then the distribution $d(a_i) = X_i$ makes A true. In the other direction, reverse the construction. $\qquad\square$

We now show that when an ultrafilter saturates the random graph, we can always split the size of a regularizing set in half. Note that it is always possible to map $\kappa \times \lambda$ into λ and redistribute, which might be called halving (or dividing by κ) from the point of view of the ultrapower. However, what is described here is a priori stronger: the size is halved index model by index model.

Lemma 5.7. *Suppose $\mathcal{D}$ is a regular ultrafilter on λ such that for some countable unstable theory T, and for any $M \models T$, the ultrapower $M^\lambda/\mathcal{D}$ is λ^+-saturated. Let $\{n_t\}_{t<\lambda}$ be an unbounded sequence of finite cardinals, and suppose that $\mathcal{D}$ contains a regularizing set X of size $n^* := \prod_t n_t/\mathcal{D}$. Then $\mathcal{D}$ also contains a regularizing set Y of size $m^* := \prod_t \frac{n_t}{2}/\mathcal{D}$.*

Proof. As usual, i will subscript elements of sets of size λ and t will refer to elements of the index set. Let $X = \{X_i : i < \lambda\}$ and the sequence $\{n_t\}$ be given, along with any infinite index model M and its ultrapower $N := M^\lambda/\mathcal{D}$. Choose any true set $A = \{a_i : i < \lambda\}$ corresponding to X, that is, any set of elements of N which is true under the distribution $a_i \mapsto X_i$. Now add to the language a new binary relation symbol $<$. We will expand each index model $M[t]$ to a model in which $<$ is a total discrete linear order as follows. Writing $A[t]$ for the set of $M[t]$-images of A under d, i.e. $\{a_i[t] : t \in d(a_i)\}$, we have by assumption of trueness that if $t \in d(a_i) \cap d(a_j)$, $a_i[t] = a_j[t]$ iff $i = j$. Since these elements are distinct, we can ensure that the linear order on the indices is respected by $<$, i.e. $a_i[t] <^{M[t]} a_j[t]$ if and only if $i < j$. On the other elements of $M[t]$, choose $<^{M[t]}$ however you like but so that it is a discrete linear order on the model. If we do this in each index model, it induces a total discrete linear order on N by Fact 3.13.

Notice that $\langle a_i : i < \lambda \rangle$ will be a $<^N$-increasing λ-indexed set in N. In particular, it will be well-ordered (this is certainly not true of the order $<^N$ on N).

Now we define a second true set $B = \{b_i : i < \lambda\}$ as follows, a kind of inverse of A. In each index model $M[t]$, suppose $A[t]$ contains the images of $a_{i_1}, \ldots a_{i_n}$, with $i_1 < \cdots < i_n$ and thus $a_{i_1}[t] <^{M[t]} \cdots <^{M[t]} a_{i_n}[t]$. We will define $b_{i_1}[t], \ldots b_{i_n}[t]$ on the same set but in reverse order, i.e. $b_{i_{k+1}}[t] := a_{i_{n-k}}[t]$ for $0 \le k \le n-1$. Finally, for each $i < \lambda$, set $b_i := \prod_t b_i[t]/\mathcal{D}$. Because we have simply reversed the sequence at each index, the set B in the ultrapower is a λ-indexed strictly $<^N$-descending sequence, and moreover B is true under the same distribution $d(b_i) := d(a_i)$.

In particular, since A is $<^N$-well ordered, this means that the intersection $A \cap B$ must be finite. We can therefore assume (by throwing away finitely many elements, if necessary) that A, B are both true sets of size λ and are disjoint.

Now we invoke saturation. By hypothesis, $\mathcal{D}$ saturates some unstable theory, and so by Lemma 5.3 we may assume it saturates models of the random graph. In particular, given any two disjoint small sets $A, B \subset N$, there is a distribution $d_0 : A \cup B \to \mathcal{D}$ such that $A[t] \cap B[t] = \emptyset$ almost everywhere. From this, define a distribution $d_h : A \cup B \to \mathcal{D}$ by $d_h(a_i) = d_h(b_i) := d_0(a_i) \cap d_0(b_i)$, and likewise for b_i. Therefore a.e. $|A[t]| = |B[t]| \le \frac{n_t}{2}$. By the previous Observation, this suffices for the existence of the desired Y: that is, let $Y = \{X_i \cap d_h(a_i) : i < \lambda\}$. $\qquad\square$

6. A minimal TP_2 theory

We recall the definitions of simple, TP_1, and TP_2. For simplicity as seen by the characteristic sequence, see [12] Theorem 5.22; analogously to the unstable case, the relevant hypergraph configuration is in Claim 6.3 below.

Definition 6.1. *(Shelah) The formula $\varphi(x; y)$ has the (2-) tree property if, in the monster model, there is a tree of instances of φ, $\{\varphi(x; a_\eta) : \eta \in {}^{<\omega}\omega\}$ such that:*

- *the set of instances which correspond to any given branch on the tree form a consistent partial φ-type;*
- *any two instances which are siblings (i.e. have a common immediate predecessor) are inconsistent*

Call two instances $\varphi(x; a_\eta), \varphi(x; a_\nu)$ incomparable if η, ν do not lie along the same branch. If, moreover, we can assume that:

- *any two incomparable instances are inconsistent, then φ has* the tree property of the first kind, TP_1
- *any two incomparable instances (which are not siblings) are consistent, then φ has* the tree property of the second kind, TP_2

T is simple *if no formula φ has the tree property.*

The analogue of the result that any unstable theory either contains a formula with the independence property or a formula with the strict order property is, in this context:

Theorem I. (Shelah, [16]) *If T is not simple, then it contains either a formula with TP_2 or a formula with TP_1.*

Definition 6.2. *Say that the characteristic sequence $\langle P_n \rangle$ has an (ω, ω)-array if (it is consistent that) there exist elements $\{ c_i^t : i < \omega, t < \omega \} \subset P_1^M$ such that $A \subset I$ is a positive base set iff it contains no more than one element from each column, i.e. $c_i^t, c_j^s \in A \implies (i = j \rightarrow s = t)$. (Compare Claim 4.11.)*

Claim 6.3. *([12] Claim 3.8) The following are equivalent:*

(1) *φ has TP_2.*
(2) *$\langle P_n \rangle$ has an (ω, ω)-array.*

We will show that there is a minimal theory among those whose associated sequences contain such an array. But there is a subtlety involved in the "obvious guess" of building a theory from an (ω, ω)-array. Consider what appears to be a simplification of the theory T^*_{feq} from [15] Definition 2.1. T^*_{feq} is the model completion of a parametrized family of independent equivalence relations (which eliminates quantifiers and is not SOP_3, in fact not SOP_1 [17]). Here we just want a single equivalence relation, so call it T^s:

Definition 6.4. *T^s is the model completion of the following theory. The language contains two unary predicates X, Y, a binary relation E, a binary relation R and a partial binary function f. The theory states that:*

- *X, Y are both infinite and disjoint, and partition the universe*
- *E is an equivalence relation on Y with infinitely many infinite classes*
- *Each element of X is connected to no more than one element of each E-equivalence class*
- *$f : X \times Y \rightarrow Y$ satisfies $R(a, f(a, b)) \wedge E(b, f(a, b))$*

However, as noted in Džamonja and Shelah [4] Remark 1.8 (where this simplification is also called T^*_{feq}), the preimages of the function $f(x, a)$ give an equivalence relation with infinitely many infinite classes on X. (Without this function, the theory does not eliminate quantifers but the preimage remains definable.) So saturating models of this theory will, on some level, involve making decisions about infinitely many independent equivalence relations, and our argument will need to take this into account. We will continue with the notation T^s.

Definition 6.5. *Let $\mathcal{D}$ be a regular ultrafilter on λ. Say that $\mathcal{D}$ solves (ω, ω) if it realizes all types corresponding to positive base sets which come, a.e., from an (ω, ω) array. More precisely, for any characteristic sequence $\langle P_n \rangle$ and any model M containing a set $C \subset P_1^M$ which is an (ω, ω) array for $\langle P_n^M \rangle$, let $N := M^\lambda / \mathcal{D}$. Write $C[t]$ for the copy of C in the index model $M[t]$ and set $C^N := \prod_t C[t] / \mathcal{D}$ (this is an induced set). Then for any small positive base set $A \subset C^N$, the type corresponding to A is realized.*

Recall Remark 5.2.

Remark 6.6. *If $\mathcal{D}$ solves (ω, ω), then it is straightforward to see that it must solve $(\omega, 2)$-arrays (see the previous section) as well, and therefore saturate models of the random graph.*

Throughout this section "induced function" (resp. induced equivalence relation) means an induced predicate whose projection to almost every index model is a function (resp. equivalence relation).

Lemma 6.7. *Let $\mathcal{D}$ be a regular ultrafilter on $\lambda \geq \aleph_0$ and $N_0 = M_0^\lambda / \mathcal{D}$ any infinite ultrapower. Suppose that*

(1) *$\mathcal{D}$ solves (ω, ω), and*
(2) *for any family $\langle A_i : i < \lambda \rangle$ of disjoint small subsets of any (infinite) regular ultrapower $N_0 = M_0^\lambda / \mathcal{D}$, there exists an induced equivalence relation E such that the A_i are contained in distinct E-equivalence classes, i.e. $a_t^i, a_s^i \in A_i$ implies $E(a_t^i, a_s^i)$, and $a^i \in A_i, a^j \in A_j$ implies $E(a^i, a^j)$ if and only if $i = j$.*

If these two conditions are satisfied, then for any model M of T^s, $N := M^\lambda / \mathcal{D}$ is λ^+-saturated.

Proof. By Remark 6.6, we may assume that $\mathcal{D}$ saturates models of the random graph and all pseudofinite sets are large. By quantifier elimination, the picture of the key types is thus straightforward. The [positive] atomic formulas are $Y(z), X(z)$ [we will use x, y to denote this membership and z, w for indeterminates], $E(y, y')$, $z = w$, and $f(x, y) = y'$. The key type of an element $x \in X$ in the ultrapower N is not affected by other parameters from X: the issue is which element from each E-equivalence class it will choose to relate to (since it chooses precisely one from each class, we may replace the negative statements with other positive ones). This is an (ω, ω)-problem, so these types will be realized whenever $\mathcal{D}$ solves (ω, ω).

By the first sentence of the proof, the key type of an element $y \in Y$ is the one involving f. In particular, given any family $\langle C_i : i < \lambda \rangle$ of disjoint small subsets of X^N, the partial type $\{f(c, y) = f(c', y) : c, c' \in C_i, i < \lambda\} \cup \{\neg f(c, y) = f(c', y) : c \in C_i, c' \in C_j, i < j < \lambda\}$ is finitely consistent because the theory is model complete. Suppose we are given a distribution $d : \bigcup_i C_i \to \mathcal{D}$ assigning finitely many elements $c_{i_1}, \ldots c_{i_n}$ to each index model. Now invoke hypothesis (2) for the family $\langle C_i \rangle$. In each index model $M[t]$, choose an element y which has the property that for the finitely many relevant c_i, c_j, we have $f^{M[t]}(c_i[t], y) = f^{M[t]}(c_j[t], y)$ if and only if $E(c_i[t], c_j[t])$. By Łoś' theorem, Fact 3.13 and the hypothesis on E, f behaves as desired. $\square$

Lemma 6.8. *Suppose that $\mathcal{D}$ solves (ω, ω). Let C be any small set in an ultrapower $N := M^\lambda / \mathcal{D}$. Then there is a distribution $d : C \to \mathcal{D}$ such that for all $t \in \lambda$ and any two distinct elements $c, c' \in C$, if $t \in d(c) \cap d(c')$ then $c[t] \neq c'[t]$. In other words, there is a distribution under which C is true.*

Proof. Begin with C, M, N as described and a distribution $e : C \to \mathcal{D}$ which assigns finitely many elements to each index model. Let M' be a model containing an (ω, ω)-array W for some background characteristic sequence $\langle P_n \rangle$, and let $N' := (M')^\lambda / \mathcal{D}$. We will define, index model by index model, a map $\rho : \lambda \to C' \subset N'$ whose domain is the subscripts (indices) of elements of C under some fixed enumeration and whose image is a positive base set.

In $M[t]$, let $C_t = \{c_{i_1}, \ldots c_{i_n}\}$ be the finitely many elements whose image under e includes t, as distinguished from $C[t] := \{c_{i_1}[t], \ldots c_{i_n}[t]\}$. Similarly to the proof of Lemma 5.3, we define a map

$\rho_t : \{i < \lambda : i \in C_t\} \to W$ from this set of indices into W so that

$$(\dagger) \qquad M' \models P_2(\rho_t(i), \rho_t(j)) \text{ if and only if } M \models c_i[t] \neq c_j[t]$$

In other words, since equality is an equivalence relation, we map equivalence classes of $=^{M[t]}$-equal elements into columns of the (ω, ω)-array, with distinct equivalence classes mapping to distinct columns. Within each equivalence class, each element has a distinct image. Since the rows and columns of W are infinite, there is always enough room. For each index t and subscript i, let $c'_i[t] := \rho_t(i)$. Finally, set $c'_i := \prod_t c'_i[t]/\mathcal{D}$ for each $i < \lambda$, and let $\rho := \prod_t \rho_t/\mathcal{D}$.

By $(\dagger)$ and Łos' theorem, if c_i, c_j are distinct elements of C then c'_i, c'_j are P_2-consistent elements of C' in N', so C' is a $P_2^{N'}$-complete graph, and it is small since it is in bijection with C. Once again, by Remark 5.2, it is therefore a positive base set, so corresponds to a consistent partial type. By hypothesis, $\mathcal{D}$ realizes all such types. By Lemma 4.8, this means that there is a distribution $d' : C' \to \mathcal{D}$ which is a.e. a P_∞-complete graph. Since C' comes a.e. from W, this means that for almost all t the indices corresponding to the finitely many elements of C' assigned to the model $M[t]$ are in distinct columns of W. Since ρ was defined to be a bijection, the map $c_i \mapsto d'(c'_i)$ gives us a corresponding distribution of C under which, a.e., distinct elements are distinct, as desired. $\qquad\square$

Of course, we need not be able to find distributions which are true with respect to other relations which may be in the language, e.g. order.

Theorem 6.9. *Suppose $\mathcal{D}$ is a regular ultrafilter which solves (ω, ω) and $N = M^\lambda/\mathcal{D}$ an ultrapower. Then for any family $\langle C_i : i < \lambda \rangle$ of disjoint small sets and any sequence $\langle b_i : i < \lambda \rangle$, all from N, (1) there exists an induced many-to-one function f which satisfies: for each $i < \omega$ and all $c \in C_i$, $f(c_i) = b_i$. (2) There exists an induced equivalence relation with infinitely many infinite classes in which each C_i is contained in a single equivalence class, but C_i, C_j are inequivalent for $i \neq j$. (3) If we assume each C_i contains a single element c_i, there exists an induced bijection which satisfies $f(c_i) = b_i$ for all $i < \lambda$.*

Proof. Apply the previous lemma to obtain a distribution of $\bigcup_i C_i \cup \{b_i : i < \lambda\}$ which is true. Given such a distribution, it is straightforward to define the desired many-to-one function, equivalence relation or bijection, index model by index model. Since these are all definable properties, they will also hold of the induced predicate in the ultrapower, again by Łos' theorem and Fact 3.13. $\qquad\square$

Corollary 6.10. *T^*_{feq}, or equivalently T^s, is minimal among theories with TP_2 in Keisler's order.*

Proof. By Claim 6.3, all TP_2 theories have an associated characteristic sequence which contains an (ω, ω)-array; therefore, analogously to Observation 5.1, one can show that if $\mathcal{D}$ saturates models of some TP_2 theory then it will solve (ω, ω). The result now follows by Lemma 6.7 and Theorem 6.9. The equivalence of T^*_{feq} and T^s is [4] Remark 1.8. $\qquad\square$

Remark 6.11. *It is likely that this theory is in fact minimal among all theories which are not simple (i.e. theories with the tree property). By Theorem I above, it would suffice to show that T^s it is smaller than $(\trianglelefteq)$ any theory with TP_1.*

7. Density and maximality

So far, considering the characteristic sequences of unstable theories, an underlying theme has been the interaction of P_∞-complete graphs, i.e., positive base sets:

- The configuration which characterized the minimal unstable theory, an $(\omega, 2)$-array, says essentially that there are two infinite positive base sets which are only slightly inconsistent: the number of missing edges is linear in the size of a finite subgraph.
- The configuration which characterized the minimal TP_2-theory, an (ω, ω)-array, can be thought of as a generalization of the $(\omega, 2)$-configuration: if we choose any two distinct elements from each column, we produce an $(\omega, 2)$-array. The inconsistency is still uniform, but informally speaking, relatively sparse.

What happens when the interaction of positive base sets becomes more complex? In this section we apply the analysis of [13] to give a new sufficient condition for maximality, Theorem 7.11 below. In this section, we will regularly assume that $\langle P_n \rangle$ has support 2. As the examples indicate, this assumption may not be as restrictive as one might think.

Warning 7.1. *Uniquely in this section we will need the word "induced" in the graph-theoretic sense of isomorphic subgraph, which has nothing to do with Definition 4.6. We write <u>induced</u> for this graph-theoretic meaning.*

7.1. **Background: The need for randomness.** It has been observed by various people that formulas which "can code anything" must be maximal (starting with Keisler's notion of a *versatile formula*, p. 33 of [7]). Here is the statement in our language; the picture of graphs will, eventually, suggests not just when this occurs but when it may not.

Fact 7.2. *Suppose that we have two characteristic sequences, $\langle P_n \rangle$ for φ and $\langle P_n' \rangle$ for φ'. Suppose that every finite <u>induced</u> sub-hypergraph of $\langle P_n' \rangle$ can be isomorphically embedded into $\langle P_n \rangle$. Then any regular ultrafilter which realizes all small φ-types must also realize all small φ'-types.*

Proof. Let $A' \subset P_1'$ be a positive base set for φ', given with a distribution $d' : A' \to \mathcal{D}$. Essentially, the hypothesis allows us to transfer the blueprint of A' (index model by index model) over into the characteristic sequence of φ where, by Łoś' theorem, it will induce a positive base set A for φ in the ultrapower. A realization of this φ-type can then be pulled back to realize the type corresponding to A'. $\square$

Corollary 7.3. *If the characteristic sequence $\langle P_n \rangle$ is universal for the finite <u>induced</u> subhypergraphs of every consistent characteristic sequence, then T is maximal in the Keisler order.*

7.2. **The random (Rado) graph.** A stronger fact is that universality in certain initial segments of the sequence may suffice.

Remark 7.4. *There is a theory which is maximal in the Keisler order and whose saturation depends on a characteristic sequence of support 2.*

Proof. Shelah proved (Theorem A above) that any theory with the strict order property is maximal in Keisler's order. Consider $\langle \mathbb{Q}, < \rangle$ and let $\varphi(x; y, z) = y \geq x > z$. $\square$

Definition 7.5. *Suppose that $(T, \varphi) \mapsto \langle P_n \rangle$, and let $M \models T$ be $\aleph_1$-saturated. Say that the sequence $\langle P_n \rangle$ embeds a 2-random graph if there exists an infinite set $A \subset P_1^M$ such that:*

- *The structure whose base set is A and whose relation R is given by P_2 on A is a model for the Rado graph.*
- *The sequence $\langle P_n \rangle$ restricted to A has support 2.*

Claim 7.6. *Suppose that $(T, \varphi) \mapsto \langle P_n \rangle$ and that the sequence $\langle P_n \rangle$ embeds a 2-random graph. Then T is maximal in Keisler's order.*

Proof. Apply Fact 7.2 to the example of Remark 7.4. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

7.3. **Another reason for randomness.** We recall two definitions from graph theory.

Definition 7.7. *Let (X, Y) be a finite 2-partite graph, i.e. a graph whose vertex set is partitioned into two sets X, Y (we avoid the term bipartite as we do not assume that either X or Y contains no edges). The* density *$\delta(X, Y) \in [0, 1]$ is given by $e(X, Y)/|X||Y|$, where the numerator signifies the number of edges between elements of X and elements of Y.*

The following definition gives a measure of how uniformly the edges of the graph are distributed:

Definition 7.8. *For any real $\epsilon \in (0, 1)$, the finite 2-partite graph (X, Y) is said to be ϵ-regular if the following holds: for any $X' \subset X$, $Y' \subset Y$ with $|X'| > \epsilon|X|, |Y'| > \epsilon|Y|$, we have $|\delta(X, Y) - \delta(X', Y')| < \epsilon$.*

In the context of characteristic sequences (and this paper), we will compute density using the graph edge P_2.

Definition 7.9. *Let $\langle P_n \rangle$ be a characteristic sequence (usually – we will clearly say when – assumed to be of support 2). Say that* positive base sets generically interrelate *in $\langle P_n \rangle$ if there exists $\epsilon_0 \in (0, \frac{1}{2})$ such that for all $\epsilon > 0$, $N \in \mathbb{N}$ there exist $X, Y \subset P_1$ such that:*

- *$|X| = |Y| \geq N$*
- *X, Y are positive base sets*
- *(X, Y) is ϵ-regular of density δ considered as a P_2-graph, where $\epsilon_0 < \delta < 1 - \epsilon_0$.*

In other words, the density of (some such sequence of) arbitrary large ϵ-regular pairs of positive base sets stays bounded away from 0 and 1 as $N \to \infty$, $\epsilon \to 0$.

In any characteristic sequence, the existence of arbitrarily large ϵ-regular pairs of positive base sets is given by Szemerédi's celebrated regularity lemma (see [10]; the model-theoretic implications were developed and discussed in Malliaris [13]). For instance, one can apply the regularity lemma for bipartite graphs to sufficiently large pairs of disjoint positive base sets. The important criterion in this definition is that the density stay bounded away from $0, 1$, and the interesting question in general, studied abstractly in [13], is what the density means for model theory. For a discussion of a relevant result from §6 of that paper, see Remark 7.12 below. The property studied here seems to get at the more essential structure, but it would be useful to tie the two results together.

The following classic result was an early and powerful application of Szemerédi regularity (see the survey [10]). So as not to burden this paper with theory, we will leave it (as well as the Regularity Lemma) as a dark grey box, if not a black one. An extensive discussion for model theorists can be found in [13] §4, in particular Corollary 4.4 and Theorem D following. Recall Warning 7.1 above.

Theorem J. *Let G be an infinite graph. Suppose that for arbitrarily large N and arbitrarily small ϵ, there are arbitrarily large disjoint subsets X, Y of the vertex set of G such that (X, Y) is ϵ-regular and these densities remain bounded away from 0 and 1 [and each of X, Y is a complete graph]. Then G is universal for finite 2-partite graphs (X', Y') [for which each of X', Y' is a complete graph].*

Remark 7.10. *"Universal for" means "contains a copy of any such graph as an <u>induced</u> subgraph".*
In the case where we make no assumptions about X, Y the <u>induced</u> copy is only required to align on
edges between elements $x \in X', y \in Y'$.

Proof. (Sketch – for more details, see e.g. [10] and for a justification of "<u>induced</u> subgraph," [13]
Corollary 4.4.) Very informally, if we are given an ϵ-regular pair (X, Y) where $|X| = |Y|$ is large
enough and ϵ small enough, the vertex sets can be decomposed (by the Regularity Lemma) into
disjoint, relatively large sets $X = \bigcup X_i, Y = \bigcup_j Y_j$, all of which are approximately the same size,
and such that most of the pairs (X_i, Y_j) are ϵ-regular. Since each pair (X_i, Y_j) is not too small,
regularity means that their density δ_{ij} is close to that of the original pair (X, Y), so also bounded
away from $0, 1$. One then shows that we may approximately think of each set X_i (resp. Y_j) as being
represented by a new point x_i (resp: y_j) in a new graph (the so-called reduced graph) where an
edge between x_i, y_j occurs with probability δ_{ij} if the pair (X_i, Y_j) is indeed ϵ-regular. Let H be a
2-partite configuration which occurs with positive probability in the reduced 2-partite graph. If ϵ
is sufficently small and N sufficiently large relative to the degree of H, then the theorem basically
states that H will in fact occur as an <u>induced</u> subgraph of the original graph (X, Y). As $N \to \infty$
and $\epsilon \to 0$, we are thus able to obtain any finite 2-partite configuration. Moreover, to justify the
square brackets: if we may assume that the sets X and Y are always themselves complete (or empty)
graphs, then one can modify this argument to obtain any finite 2-partite configuration where the
vertices on each side of the partition form a complete (or empty) graph, essentially by observing
that the pairs (X_i, X_j) and (Y_i, Y_j) are ϵ-regular and building this into the reduced graph (though
they are of density 0 or 1, so there is no choice about these edges in the computation of possible
H). □

Theorem 7.11. *Let $\langle P_n \rangle$ be the characteristic sequence for φ with respect to a background theory*
T. If $\langle P_n \rangle$ has support 2 and positive base sets generically interrelate in $\langle P_n \rangle$, then T is maximal
in Keisler's order.

Proof. Now all the gears line up, and the proof is short. Let $\langle P_n' \rangle$ be the characteristic sequence of
$\varphi_2(x; y, z) := \varphi(x; y) \wedge \varphi(x; z)$. We will show that $\langle P_n' \rangle$ contains an <u>induced</u> copy of the Rado graph,
i.e. a 2-random graph. It suffices to show that $\langle P_n' \rangle$ universal for finite graphs. More precisely,
given a template graph $x_1, \ldots x_n$, we look for elements $a_1, \ldots a_n, b_1, \ldots b_n \subset P_1$ such that:

- $P_2(a_i, a_j)$ for $i \leq j \leq n$
- $P_2(b_i, b_j)$ for $i \leq j \leq n$
- $P_2(b_j, a_i)$ for $i \leq j \leq n$
- For $i > j$, $P_2(b_j, a_i)$ if and only if $x_i R x_j$ in the template graph

Clearly we may then set $c_i := (a_i, b_i)$ and then $P_2'(c_i, c_j)$ iff $x_i R x_j$. By Theorem J, our hypothesis
on $\langle P_n \rangle$ allows us to extract any such configuration of as and bs, and this completes the proof. □

Remark 7.12. *While this formulation seems to be the stronger one, it would be nice to establish*
an equivalence with the results of [13]§6. There we considered the compatible order property, *in*
general slightly stronger than the statement that P_2 witnesses the order property between sequences
$\langle a_i \rangle$ and $\langle b_i \rangle$ which are themselves positive base sets; in the case of support 2, this is exactly the
definition. We proved that the compatible order property is, on the level of theories, equivalent to
SOP_3, and so sufficient for maximality in Keisler's order. (SOP_3 is a reasonable candidate for
sufficiency and thus for a dividing line, but it was not clear what structure theory its absence would
allow for.)

That result ties in here as follows. If positive base sets generically interrelate in $\langle P_n \rangle$, Theorem J allows us to conclude that $\langle P_n \rangle$ has the order property between positive base sets, i.e. the "compatible order property." The contrapositive then says that if $\langle P_n \rangle$ does not have the compatible order property, positive base sets do not generically interrelate, which can be more suggestively restated as "base sets for types in generic position are either essentially consistent or essentially inconsistent." In this light, "not SOP_3" has interesting consequences for the structure theory. It would be especially interesting to know the converse, i.e. whether under reasonable assumptions (e.g. NSOP) the compatible order property, and therefore SOP_3, is equivalent to the statement that positive base sets generically interrelate.

Even assuming support 2, classification theory shows that any theory with the order property but without strict order contains a copy of a 2-partite random graph (essentially [14] Theorem II.4.10, suitably modified in [13] Lemma 5.4); but that argument does not generally preserve relations between pairs of elements contained on either side.

8. Flexibility

Given a positive base set $A \subset N := M^\lambda/\mathcal{D}$, suppose we could find, for each $n < \omega$, an induced set X such that $A \subset X \subset P_1^N$ is a P_n-complete graph. How far are we from proving that the type corresponding to A is realized?

We continue to focus on unstable formulas and theories. Recall that the condition on lcf in the next lemma will hold of any ultrafilter $\mathcal{D}$ which saturates any unstable theory.

Lemma 8.1. ([11] Lemma 9) *Let $\mathcal{D}$ be a regular ultrafilter on λ, $M \models T$ a model of some countable theory, and $A \subset N := M^\lambda/\mathcal{D}$ is any small set. Suppose that $\mathrm{lcf}(\omega, \mathcal{D}) \geq \lambda^+$. Let X_i $(i < \omega)$ be a sequence of induced predicates in N satisfying $X_i \supset X_{i+1} \supset A$, for all $i < \omega$. Then there exists an induced predicate X_∞ such that for each i, $X_i \supset X_\infty \supset A$.*

For those familiar with the paper, a remark on the proof: In that Lemma, using our language, the limit predicates X_∞ sit inside a concentric sequence X_i of definable P_∞-complete graphs extending A. In other words, they are consistent types all the way down; the refinements are to control the interaction of these X_i with partial types in other formulas. Thus the fact that X_∞ is interpreted to be X_k in the index model $M[t]$ still allows us to realize the type X_k in $M[t]$. In the general case considered here, where the predicates X_i are only assumed to be P_i-complete graphs, things need not be so simple, as the next definition suggests.

8.1. Flexible filters.
We will use the definitions of true set and of the size $\sigma(X)$ of a regularizing set X, given in Definition 5.5 above.

Definition 8.2. *(Flexibility) Let $\mathcal{D}$ be a regular ultrafilter. Say that $\mathcal{D}$ is flexible if for every $\mathcal{D}$-nonstandard integer n^*, $\mathcal{D}$ contains a regularizing set X such that $\{t : \sigma[t] \leq n[t]\} \in \mathcal{D}$, where $\sigma = \sigma(X)$.*

Kunen suggested to me that flexibility implies λ-OK, and so we verify this here.

Definition 8.3. [1], [3] *An ultrafilter $\mathcal{D}$ (on λ) is λ-OK if for each monotonic function $g : \mathcal{P}_{\aleph_0}(\lambda) \to \mathcal{D}$ which satisfies: $g(w) = g(v)$ whenever $|w| = |v|$, g has a multiplicative refinement.*

Claim 8.4. *If $\mathcal{D}$ is a regular ultrafilter on λ which is flexible, then it is λ-OK.*

Proof. Suppose $\mathcal{D}$ is λ-OK, and let $g : \mathcal{P}_{\aleph_0}(\lambda) \to \mathcal{D}$ satisfies $g(w) = g(v)$ whenever $|w| = |v|$. This says that for each index model M_t there is a largest integer n_t such that some (equivalently all, by definition) elements v of the domain of g of cardinality n_t, $t \in g(v)$. By flexibility, there exists a regularizing set $\{X_i : i < \lambda\}$ whose size is less than $n^* := \Pi_{t<\lambda} n_t / \mathcal{D}$. Define g' by $g'(\{i\}) := g(\{i\}) \cap X_i$ for each $i \in \lambda$, and for $|w| > 1$ $g'(w) := g(w) \cap \bigcap_{v \subsetneq w} g'(v)$. Let us check that f is multiplicative. At each index t,

$$\left| i < \lambda : t \in g'(\{i\}) \right| \leq n_t$$

and by hypothesis, $t \in g(v)$ for all $m \leq n_t$ and all $v \subset \lambda$, $|v| = m$. In particular, $t \in g'(\{i < \lambda : t \in g'(\{i\})\})$, as desired. $\square$

As in the case of goodness, in order to establish flexibility as a useful property for saturation of ultrapowers, we show that it is captured by types in some countable first-order theory. The following class of theories was first studied by Buechler:

Definition 8.5. *The theory T is* low *if for every formula φ there exists $k < \omega$ such that for every instance $\varphi(x; a)$ of φ, $\varphi(x; a)$ divides iff it $\leq k$-divides. The theory is* simple *if for every formula φ and every $k < \omega$ there exists n_k such that $D(x = x, \varphi, k) < n_k$, that is, φ cannot sequentially k-divide more than n_k times.*

Remark 8.6. *Any stable theory is low (in fact, any theory without the independence property is low: [12] Observation 4.6).*

Lemma 8.7. *Let φ be a formula of T which is not low, $M \models T$ $\aleph_1$-saturated, $\mathcal{D}$ regular on λ. If $M^\lambda/\mathcal{D}$ is λ^+-saturated, then $\mathcal{D}$ must be flexible.*

Proof. Suppose we are given some nonstandard integer $n^* = \prod_t n[t]/\mathcal{D}$. Let us identify a small set $A \subset N$ and a consistent type $p \in S_\varphi(A)$ which is realized iff $\mathcal{D}$ has a regularizing set with size $\sigma \leq n^*$ modulo $\mathcal{D}$.

By choosing the index model M to be sufficiently saturated, we may assume by the hypothesis of non-lowness that M contains indiscernible sequences I_k, for each $k < \omega$, such that $\{\varphi(x; c) : c \in I_k\}$ is k-consistent but $(k + 1)$-inconsistent. Write $I_k[t]$ for this sequence in the tth copy of the index model, $M[t]$.

Fix a regularizing set $X = \{X_i : i < \lambda\}$. To build our set A, it suffices to define each element a_i on X_i. As before, we will think of the element a_i as a representative of the set X_i, in a sense that will be clear from the construction: a realization of the type $p \in S(A)$ will give the desired refinement of X. Let $Y[t] = \{t : t \in X_i\}$ be the indices of elements to be defined in $M[t]$, and let $m[t] = |Y[t]|$. We define the $a_j[t]$ (for $j \in Y[t]$) to be distinct elements of $I_{n[t]}[t]$, where recall that $n[t]$ is the size we are aiming for at index t. More precisely, we choose the a_j such that:

(1) $j, k \in Y[t] \implies a_j[t] \neq a_k[t]$
(2) for all $\rho \subseteq Y[t]$, because the elements are chosen along the indiscernible sequence $I_{n[t]}$,

$$M[t] \models \exists x \left(\bigwedge_{j \in \rho} \varphi(x; a_j[t]) \right) \iff |\rho| \leq n[t]$$

i.e., we have chosen our $m[t]$-many elements of $M[t]$ so that precisely the subsets of size $\leq n[t]$ are consistent. To finish, for each $i < \lambda$, set $a_i := \prod_t a_i[t]/\mathcal{D}$, and $p := \{\varphi(x; a_i) : i < \lambda\}$. To see that this is a consistent type, let $p_0 = \{\varphi(x; a_{i_1}), \dots \varphi(x; a_{i_k})\}$ be any finite subset. Then $X_{i_1} \cap \cdots \cap X_{i_k} \cap \{t :$

$n[t] > k\} \in \mathcal{D}$, and at every index in this large set the formula $\varphi(x; a_{i_1}[t]), \ldots \varphi(x; a_{i_k}[t])$ have a common realization in $M[t]$, by condition (2).

On the other hand, suppose that $\alpha \models p$ for some $\alpha \in N$. Now the distinctness of the elements of A allows us to push down the size of the original regularizing set X. Namely, let $Z = \{Z_i : i < \lambda\}$ be given by $Z_i = \{t \; : \; t \in X_i, \; M[t] \models \varphi(\alpha[t]; a_i[t])\}$. By Łoś' theorem, $Z_i \in \mathcal{D}$. Because it refines X, Z remains a regularizing set. Because the a_i were chosen to be distinct,

$$\left| \{i : t \in X_i, \; M[t] \models \varphi(\alpha[t]; a_i[t])\} \right| \leq n[t]$$

so $\sigma(Z) \leq n^*$ modulo $\mathcal{D}$. $\qquad\qquad\square$

Recall from above that the dividing line within stable theories, namely the finite cover property, corresponds to sequences which do not have finite support; however, this need not remain true for more complex theories. There is a maximal theory whose saturation depends on a sequence of support 2, but we can be more precise. In the case of fcp, we have seen that there is a minimal unstable theory and its saturation depends on a sequence which has support 2. Parallel to this result, we now show that any filter which saturates the minimal TP_2 theory (which itself has finite support) must be flexible.

Lemma 8.8. *If $\mathcal{D}$ saturates any theory with TP_2 then $\mathcal{D}$ is flexible.*

Proof. By Remark 6.6, $\mathcal{D}$ will saturate ultrapowers of the random graph and by the second half of Remark 5.4, we may therefore assume $\mu(\mathcal{D})$ is large. In particular, given any nonstandard integer $n^* := \Pi_{t<\lambda} n_t / \mathcal{D}$ and an induced set X such that $|X[t]| \leq n_t$ almost everywhere under some given distribution $d : X \to \mathcal{D}$, we have that $|X| \geq \lambda^+$. Let $C \subset X$ be any small set. Apply Lemma 6.8 to obtain a distribution, refining the given one d, under which C is true. This translates directly into a regularizing set of size n^* by Observation 5.6, which completes the proof. $\qquad\square$

Remark 8.9. *Thus there are low theories without the strict order property whose saturation nonetheless requires that $\mathcal{D}$ be flexible.*

What seems to be at stake is the following (compare the definition of the finite cover property).

Definition 8.10. (Approximations to independence) *Say that the formula $\varphi(x; y)$ has* finite approximations to independence *if for arbitrarily large $n < \omega$ there is an indiscernible sequence $I_n = \langle a_i^n : i < \omega \rangle$ such that for any $\sigma \subset I_n$, $\{\varphi(x; a_i) : i \in \sigma\}$ is consistent if and only if $|\sigma| \leq n$.*

Conclusion 8.11. *Whenever we may assume the filter $\mathcal{D}$ is flexible, the distance between the size n_t of the projection $A[t]$ under some distribution and its degree of uniform completeness k_t in the index model $M[t]$ is immaterial. This means we may focus on finding predicates for each of the complete P_n-graphs X_n, without worrying about how to put them together. Otherwise, the possible gaps between k_t and n_t which the theory is able to represent will play a role.*

9. Appendix: Some stable examples

The following basic examples illustrate the process of finding induced predicates which are P_n-complete graphs: an ultrapower may fail to realize a type either because for some n, no induced superset of A is a P_n-complete graph, or because these countably many predicates (or conditions) do not align. In the stable case, it is essentially a process of sequentially dividing finitely many

times followed by localizing to deal with finite inconsistencies (i.e. finite D-rank plus the finite equivalence relation theorem).

9.0.1. *Example 0.* Suppose $\mathcal{L}$ contains equality and a binary relation $E(x, y)$ which is interpreted in M to be an equivalence relation with infinitely many infinite classes. For the following definition, use the shorthand $z = ...$ for a formula with more variables in which we encode cases by whether or not certain inputs coincide. Let $\langle P_n \rangle$ be the characteristic sequence of

$$\varphi(x; y, z) = \begin{cases} E(x; y) \text{ if } z = 0 \\ \neg E(x; y) \text{ if } z = 1 \\ x = y \text{ if } z = 2 \\ \neg x = y \text{ otherwise} \end{cases}$$

Suppose that in the ultrapower $N := M^\lambda / \mathcal{D}$ we look at P_1^N. There are many different positive base sets A corresponding to different consistent partial φ-types. If A consists of pairs $(a, 0)$ whose first elements are all in some given equivalence class, a is an element of that class, then the set $P_1' := \{y \in P_1 : P_2(y, (a, 0))$ is definable (therefore induced) and moreover is a positive base set in N. It will in fact be a positive base set almost everywhere: in each index model, we need only avoid finitely many classes, and there are infinitely many available. In other words, the type corresponding to p includes a dividing formula, and it suffices to restrict the set P_1 of all possible parameters to those consistent with this formula.

9.0.2. *Example 1.* Consider a case where p does not divide, e.g. when the positive base set A is of the form $(a, 1)$ for some small infinite set of pairwise inequivalent elements a. There are in N infinitely many pairwise incompatible extensions of p, visible as infinitely many positive base sets A_i each of which extend A but which are not compatible, i.e. $A_i \cup A_j$ is not P_∞ complete (in fact, not P_2-complete) for any i, j. In this case, it would suffice to find $c \in N$ such that $\neg E(c, a)$ for all $(a, 1) \in A$. Then $P_1' := \{y \in P_1 : P_2((c, 0), y)\}$ will be a definable complete graph containing A. In other words, we may need to take a dividing extension of P_1 to obtain the correct induced set, even if p itself does not divide.

9.0.3. *Example 2.* Suppose that we modify the previous example as follows. Replace the equivalence relation $E(x, y)$ with a parametrized family of disjoint equivalence relations $E(x, y, z)$ and suppose that for each $n < \omega$, T says that there exists a_n such that $E(x, y, a)$ is an equivalence relation with finitely many, but more than a_n, classes all of which are infinite. In the modified Example 1, we will have a positive base set corresponding to the type of an $E(x, y, a)$-generic element (i.e. in the domain of this equivalence relation but not equivalent to any of the parameters mentioned). However, finding the element c is no longer straightforward: it will not exist unless $\mu(D) \geq \lambda^+$.

9.0.4. *Example 3.* Given a positive base set A in the ultrapower, write $A[t]$ for the image of A in index model $M[t]$ as assigned by some distribution. In the previous example, or in the case of the standard model of the finite cover property (page 5) using $\varphi(x; y, z) = E(x, y) \wedge x \neq z$, we will have the discrepancy that $n_t := |A[t]|$ is only guaranteed to be a P_{k_t}-complete graph in $M[t]$ for some $k_t < n_t$. However we may not necessarily, as in the case of non-low theories, assume that any discrepancy is possible. For instance, one can show that if it happens a.e. that $A[t]$ is uniformly P_{k_t}-complete but uniformly P_{k_t+1}-inconsistent, then either there is some constant c such that $n_t - k_t = c$ a.e. or else φ has the independence property.

References

[1] J. Baker and K. Kunen, "Limits in the Uniform Ultrafilters." Trans. AMS Vol. 353, No. 10 (Oct., 2001), pp. 4083-4093.

[2] W. Comfort and S. Negrepontis, *The Theory of Ultrafilters.* Springer-Verlag, 1974.

[3] A. Dow, "Good and OK ultrafilters." Trans. AMS Vol. 290, No. 1 (Jul., 1985), pp. 145-160.

[4] M. Džamonja and S. Shelah, "On $\lhd^*$-maximality." APAL 125 (2004) 119–158.

[5] I. Farah, B. Hart and D. Sherman. "Model theory of operator algebras II: model theory." (2010) Arxiv: 1004.0741v1.

[6] V. Guingona, "On uniform definability of types over finite sets." (2010) Arxiv: 1005.4924

[7] H. J. Keisler, "Ultraproducts which are not saturated." JSL, 32 (1967) 23–46.

[8] S. Kochen, "Ultraproducts in the Theory of Models." Ann. of Math. (2) Vol. 74, No. 2 (Sep., 1961), pp. 221-261

[9] K. Kunen, "Ultrafilters and independent sets." Trans. AMS 172 (1972), 299-306.

[10] J. Komlós and M. Simonovitz, "Szemerédi's Regularity Lemma and its Applications in Graph Theory." Combinatorics, Paul Erdös is Eighty, Vol. 2, Budapest (1996) 295–352. Bolyai Society Mathematical Studies, Proceedings of the meeting in honor of P. Erdös, Keszthely, Hungary.

[11] M. Malliaris, "Realization of φ-types and Keisler's order." Annals of Pure and Applied Logic, 157 (2009) 220–224.

[12] M. Malliaris, "The characteristic sequence of a first-order formula." In press, Journal of Symbolic Logic (2010).

[13] M. Malliaris, "Edge distribution and density in the characteristic sequence." In press, Annals of Pure and Applied Logic (2010).

[14] S. Shelah, *Classification Theory and the number of non-isomorphic models*, rev. ed. North-Holland, 1990.

[15] S. Shelah, "The Universality Spectrum : Consistency for more Classes." In *Combinatorics: Paul Erdös is Eighty*, Vol. 1, Bolyai Society Mathematical Studies (1993).

[16] S. Shelah, "Toward classifying unstable theories." Annals of Pure and Applied Logic 80 (1996) 229–255.

[17] S. Shelah and A. Usvyatsov, "More on SOP$_1$ and SOP$_2$." Annals of Pure and Applied Logic 155 (2008), no. 1, 16–31.

Department of Mathematics, University of Chicago, 5734 S. University Ave, Chicago, IL 60637

E-mail address: mem@math.uchicago.edu