STRUCTURE AND ENUMERATION THEOREMS FOR HEREDITARY PROPERTIES IN FINITE RELATIONAL LANGUAGES

CAROLINE TERRY

Abstract. Given a finite relational language $\mathcal{L}$, a hereditary $\mathcal{L}$-property is a class of finite $\mathcal{L}$-structures which is closed under isomorphism and model theoretic substructure. This notion encompasses many objects of study in extremal combinatorics, including (but not limited to) hereditary properties of graphs, hypergraphs, and oriented graphs. In this paper, we generalize certain definitions, tools, and results form the study of hereditary properties in combinatorics to the setting of hereditary $\mathcal{L}$-properties, where $\mathcal{L}$ is any finite relational language. In particular, the goal of this paper is to generalize how extremal results and stability theorems can be combined with well-known techniques and tools to yield approximate enumeration and structure theorems. We accomplish this by generalizing the notions of extremal graphs, asymptotic density, and graph stability theorems using structures in an auxiliary language associated to a hereditary $\mathcal{L}$-property. Given a hereditary $\mathcal{L}$-property $\mathcal{H}$, we prove an approximate asymptotic enumeration theorem for $\mathcal{H}$ in terms of its generalized asymptotic density. Further we prove an approximate structure theorem for $\mathcal{H}$, under the assumption of that $\mathcal{H}$ has a stability theorem. The tools we use include a new application of the hypergraph containers theorem (Balogh-Morris-Samotij [14], Saxton-Thomason [37]) to the setting of $\mathcal{L}$-structures, a general supersaturation theorem for hereditary $\mathcal{L}$-properties (also new), and a general graph removal lemma for $\mathcal{L}$-structures proved by Aroskar and Cummings in [5]. Similar results in the setting of multicolored graphs and hypergraphs were recently proved independently by Falgas-Ravry, O'Connel, Strömberg, and Uzzell [21].

1. Introduction

The study of hereditary properties of combinatorial structures is an important topic within the field of extremal combinatorics. Out of the many results in this line of research has emerged a pattern for how to prove approximate asymptotic enumeration and structure results. The aim of this paper is to provide a general framework in which to view these results and to formalize this pattern of proof.

1.1. **Background.** A nonempty class of graphs $\mathcal{P}$ is called a *hereditary graph property* if it is closed under isomorphism and induced subgraphs. Given a hereditary graph property $\mathcal{P}$, let $\mathcal{P}_n$ denote the set of elements of $\mathcal{P}$ with vertex set $[n]$. There has been extensive investigation into the properties of $\mathcal{P}_n$, where $\mathcal{P}$ is a hereditary property of graphs and n is large, see for instance [1, 2, 9–11, 18, 19, 38]. The main questions addressed in these papers concern *enumeration* (finding an asymptotic formula for $|\mathcal{P}_n|$) and *structure* (understanding what properties elements of $\mathcal{P}_n$ have with high probability). Given a graph H, $\text{Forb}(H)$ (respectively $\text{Forb}_{ind}(H)$) is the class of finite graphs omitting H as a non-induced (respectively induced) subgraph. For any graph H, both $\text{Forb}(H)$ and $\text{Forb}_{ind}(H)$ are hereditary graph properties. Therefore, work on hereditary graph properties can be seen as generalizing the many structure and enumeration results about graph properties of the form $\text{Forb}(H)$ and $\text{Forb}_{ind}(H)$, for instance those appearing in [22, 26, 34–36]. From this perspective, the study of hereditary graph properties has been a central area of research in extremal combinatorics.

There are many results which extend the investigation of hereditary graph properties to other combinatorial structures. Examples of this include [7] for tournaments, [6] for oriented graphs and posets, [20, 25] for k-uniform hypergraphs, and [23] for colored k-uniform hypergraphs. The results in [15, 16, 32, 33] investigate asymptotic enumeration and structure results for specific classes of

H-free hypergraphs, which are examples of hereditary properties of hypergraphs. Similarly, the results in [28] concern specific examples of hereditary properties of digraphs. The results in [31] for metric spaces are similar in flavor, although they have not been studied explicitly as instances of hereditary properties. Thus, extending the investigation of hereditary graph properties to other combinatorial structures has been an active area of research for many years.

From this investigation, patterns have emerged for how to prove these kinds of results, along with a set of standard tools, including extremal results, stability[1] theorems, regularity lemmas, supersaturation results, and the hypergraph containers theorem. In various combinations with extremal results, stability theorems, and supersaturation results, Szemerédi's regularity lemma has played a key role in proving many results in this area, especially those extending results for graphs to other settings. A sampling of these are [3, 8, 35] for graphs, [4] for oriented graphs, [15, 16, 20, 23, 25, 32, 33] for hypergraphs, and [31] for metric spaces. The hypergraph containers theorem, independently developed in [14, 37], has been used in many recent papers in place of the regularity lemma. Examples of this include [12–14, 24, 30, 37] for graphs, [28] for digraphs, and [17] for metric spaces. In these papers, the commonalities in the proofs are especially clear. Given an extremal result, there is clear outline for how to prove an approximate enumeration theorem. If on top of this, one can characterize the extremal structures and prove a corresponding stability theorem, then there is a clear outline for how to prove an approximate structure theorem. The goal of this paper is to make these proof outlines formal using generalizations of tools, definitions, and theorems from these papers to the setting of structures in finite relational languages.

1.2. **Summary of Results.** Given a first-order language $\mathcal{L}$, we say a class $\mathcal{H}$ of $\mathcal{L}$-structures has the *hereditary property* if for all $A \in \mathcal{H}$, if B is a model theoretic substructure of A, then $B \in \mathcal{H}$.

Definition 1. Suppose $\mathcal{L}$ is a finite relational language. A *hereditary $\mathcal{L}$-property* is a nonempty class of finite $\mathcal{L}$-structures which has the hereditary property and which is closed under isomorphism.

This is the natural generalization of existing notions of hereditary properties of various combinatorial structures. Indeed, for appropriately chosen $\mathcal{L}$, almost all of the results cited so far are for hereditary $\mathcal{L}$-properties, including all hereditary properties of graphs, k-uniform hypergraphs, colored k-uniform hypergraphs, directed graphs, and posets, as well as the the metric spaces from [31]. Thus hereditary $\mathcal{L}$-properties are the appropriate objects to study in order to generalize many of the results we are interested in.

We now give a description of our results. The precise statements require extensive preliminaries and appear in Section 4. Fix a finite relational language $\mathcal{L}$ with maximum arity r. Given a hereditary $\mathcal{L}$-property $\mathcal{H}$, we will define an invariant associated to $\mathcal{H}$, called the *asymptotic density of $\mathcal{H}$*, denoted by $\pi(\mathcal{H})$ (see Definition 18). Our first main result, Theorem 3, gives an asymptotic enumeration of $\mathcal{H}_n$ in terms of $\pi(\mathcal{H})$, where $\mathcal{H}_n$ denotes the set of elements in $\mathcal{H}$ with domain $[n]$.

Theorem 3 (Enumeration)**.** *Suppose $\mathcal{H}$ is a hereditary $\mathcal{L}$-property. Then the following hold.*

(1) If $\pi(\mathcal{H}) > 1$, then $|\mathcal{H}_n| = \pi(\mathcal{H})^{\binom{n}{r}+o(n^r)}$.
(2) If $\pi(\mathcal{H}) \leq 1$, then $|\mathcal{H}_n| = 2^{o(n^r)}$.

The tools we use to prove this theorem include a general supersaturation theorem for $\mathcal{L}$-structures (Theorem 7) and a new adaptation of the hypergraph containers theorem to the setting of $\mathcal{L}$-structures (Theorems 6).

Theorem 7 (informal)**.** *A hereditary $\mathcal{L}$-property always has a supersaturation theorem.*

The proof of Theorem 7 uses our hypergraph containers theorem for $\mathcal{L}$-structures (Theorem 6) and a generalization by Aroskar and Cummings [5] of the graph removal removal lemma (Theorem 9).

[1]This use of the word stability refers to a type of theorem from extremal combinatorics and is unrelated to the model theoretic notion of stability.

Our proof strategies for these theorems draw on a series of enumeration results for combinatorial structures which employ the hypergraph containers theorem, namely those in [14, 17, 28, 30, 37].

We will also define generalizations of extremal graphs (Definition 17) and graph stability theorems (Definition 19). We will prove that the existence of a stability theorem, along with an understanding of extremal structure, always yield an approximate structure theorem. This result, Theorem 5, generalizes arguments appearing in many papers, including [15, 16, 28, 31, 33].

Theorem 5 (informal)**.** *Stability theorem + Characterization of extremal structures ⇒ Approximate Structure.*

The main tool used to prove Theorem 5 is a second adaptation of the hypergraph containers theorem to the setting of $\mathcal{L}$-structures, namely Theorem 8. Our adaptations of the hypergraph containers theorem, Theorems 6 and 8, rely on the original hypergraph containers theorem of [14, 37], the general graph removal lemma in [5], and the model theoretic tools developed in this paper.

It is important to note that our results apply to finite languages with relations of arbitrary arities, and to structures with non-symmetric relations. To illustrate this we direct the reader to Section 10 and the appendices of an extended arXiv version of this paper [41] for an explanation of how our results apply to examples in the settings of metric spaces, colored hypergraphs, directed graphs, and triangle-free hypergraphs.

We now clarify what the results in this paper do and what they do not do. Our main theorem, Theorem 3, gives an enumeration theorem for a hereditary $\mathcal{L}$-property in terms of its asymptotic density. However, determining the asymptotic density of a specific hereditary $\mathcal{L}$-property is often a hard combinatorial problem which this paper does not address. Similarly, while Theorem 5 shows that a stability theorem and an understanding of extremal structure implies an approximate structure theorem, proving a specific family $\mathcal{H}$ has a stability theorem and understanding its extremal structures are often difficult problems in practice. These difficulties are not addressed by the results in this paper. The role of this paper is to generalize how extremal results and stability theorems give rise to approximate structure and enumeration theorems.

While our proofs are inspired by and modeled on those appearing in [14, 17, 28, 30, 37], our results are more than just straightforward generalizations of existing combinatorial theorems. We use new tools called $\mathcal{L}_{\mathcal{H}}$-templates (see Section 3) and an application of the hyergraph containers theorem to a hypergraph whose vertices and edges correspond to certain atomic diagrams (see Theorem 10). These technical tools and their appearances in our results are non-obvious and of independent interest from a model theoretic perspective. We also provide a simplified version of the generalized graph removal lemma appearing in [5], by using a simpler notion of the distance between $\mathcal{L}$-structures may be used.

We would like to acknowledge here that similar container, enumeration, and stability results have recently been proved independently by Falgas-Ravry, O'Connel, Strömberg, and Uzzell [21] in the setting of multicolored graphs and hypergraphs. In particular, Theorems 2, 3, 6, and 7 of this paper correspond (under the right choice of language $\mathcal{L}$) to Proposition 2.10, Corollary 2.15, Theorem 2.6, and Lemma 2.11 of [21], respectively. Further, Theorem 5 of this paper is very similar to Theorem 2.19 of [21]. However, the scope and aim of the two papers are different, as are many of the proofs. The results in this paper are more general, for instance applying to unary languages and to languages with symbols of different arities, two settings not encompassed by the results in [21]. On the other hand, [21] explores interesting connections to the theory of graph limits which are not touched on here.

1.3. **Conclusion.** The work in this paper is significant from the perspective of combinatorics for three main reasons. First, problems in finite combinatorics are most often approached one by one, and techniques developed for specific structures often do not translate well into other contexts. While this style of approach is necessary to solve problems, it creates the impression that generalization is not possible. This work serves as an example that searching to generalize results and proofs within finite combinatorics using model theory can be highly fruitful. Second, this work will save researchers time by allowing them to avoid repeating arguments which now appear here in a general context. Third, we believe this paper gives the correct general framework in which to view these questions, which may aid in finding answers to open problems in the area.

This work is also of significance from the model theoretic perspective due to the following connection to logical 0-1 laws. Suppose $\mathcal{L}$ is a finite language, and for each n, $F(n)$ is a set of $\mathcal{L}$-structures with domain $[n]$. We say $F := \bigcup_{n\in\mathbb{N}} F(n)$ has a 0-1 *law* if for every first-order $\mathcal{L}$-sentence ϕ, the limit

$$\mu(\phi) := \lim_{n\to\infty} \frac{|\{G \in F(n) : G \models \phi\}|}{|F(n)|}$$

exists and is equal to 0 or 1. If F has a 0-1 law, then $T_{as}(F) := \{\phi : \mu(\phi) = 1\}$ is a complete consistent first-order theory (see [42] for a survey on 0-1 laws). There are many interesting model theoretic questions related to 0-1 laws and almost sure theories. For instance, it is not well understood in general why some classes of finite structure have 0-1 laws and why others do not. One source of known 0-1 laws are asymptotic structure results from extremal combinatorics. For instance, fix $s \geq 3$ and suppose for each n, $F(n)$ is the set of all graphs with vertex set $[n]$ omitting the complete graph K_s on s vertices. In [26] Kolaitis, Prömel and Rothschild show $F := \bigcup_{n\in\mathbb{N}} F(n)$ has a 0-1 law. Their proof relies crucially on first proving asymptotic structure and enumeration theorems for F. In particular, they show that if $S(n)$ is the set of $(s-1)$-partite graphs on $[n]$, then $S(n) \subseteq F(n)$ for all n and

$$\lim_{n\to\infty} \frac{|S(n)|}{|F(n)|} = 1. \tag{1}$$

They then prove a 0-1 law for $S := \bigcup_{n\in\mathbb{N}} S(n)$, which combines with (1) to imply F has a 0-1 law. Other asymptotic structure results which imply 0-1 laws include [16, 28, 31, 33] (for details on how these structure results imply 0-1 laws, see [27]). In these papers, [16, 28, 31, 33], the precise structure results (which are needed to prove the 0-1 laws) are proven using *approximate* structure and enumeration theorems as stepping stones. This trend suggests that a systematic understanding of precise structure and enumeration will use some version of this "approximate version" stepping stone. Therefore, understanding approximate structure and enumeration results from a model theoretic perspective is a necessary step in gaining a general understanding of precise structure and enumeration results, and consequently of the logical 0-1 laws which rely on them.

2. Preliminaries

Our goal here is to include enough preliminaries so that anyone with a rudimentary knowledge of first-order logic will be able to read this paper. Definitions we expect the reader to understand include: first-order languages, constant and relation symbols, formulas, variables, structures, substructures, satisfaction, and consistency. We refer the reader to [29] for these definitions.

2.1. Notation and Setup.

In this subsection we fix notational conventions and definitions. The word "collection" denotes either a set or a class. Suppose $\ell \geq 1$ is an integer, X is a set, and $\bar{x} = (x_1, \ldots, x_\ell)$. Then $[\ell] = \{1, \ldots, \ell\}$, $Perm(\ell)$ is the set of permutations of $[\ell]$, $\cup\bar{x} = \{x_1, \ldots, x_\ell\}$, and $|\bar{x}| = \ell$. The power set of X is denoted $\mathcal{P}(X)$ or 2^X. Given $\mu \in Perm(\ell)$, $\mu(\bar{x}) = (x_{\mu(1)}, \ldots, x_{\mu(\ell)})$. An *enumeration of* X is a tuple $\bar{x} = (x_1, \ldots, x_{|X|})$ such that $\cup\bar{x} = X$. Given $x \neq y \in X$, xy is shorthand for the set $\{x, y\}$. Set

$$X^{\underline{\ell}} = \{(x_1, \ldots, x_\ell) \in X^\ell : x_i \neq x_j \text{ for each } i \neq j\} \quad \text{and} \quad \binom{X}{\ell} = \{Y \subseteq X : |Y| = \ell\}.$$

Suppose $\mathcal{L}$ is a finite relational first-order language. Then $r_{\mathcal{L}}$ denotes the maximum arity of any relation symbol in $\mathcal{L}$. Given a formula ϕ and a tuple of variables $\bar{x}$, we write $\phi(\bar{x})$ to denote that the free variables of ϕ are all in the set $\cup\bar{x}$. Similarly, if p is a set of formulas, we write $p(\bar{x})$ to mean every formula in p has free variables in the set $\cup\bar{x}$. We will sometimes abuse notation and write $\bar{x}$ instead of $\cup\bar{x}$ when it is clear from context what is meant.

Suppose M is an $\mathcal{L}$-structure. Then $dom(M)$ denotes the underlying set of M, and the *size* of M is $|dom(M)|$. If $\mathcal{L}' \subseteq \mathcal{L}$, $M \restriction_{\mathcal{L}'}$ is the $\mathcal{L}'$-structure with underlying set $dom(M)$ such that for all $\ell \geq 1$, $\bar{a} \in dom(M)^\ell$, and ℓ-ary relation symbols R from $\mathcal{L}'$, we have $M \restriction_{\mathcal{L}'} \models R(\bar{a})$ if and only if $M \models R(\bar{a})$. We call $M \restriction_{\mathcal{L}'}$ the *reduct of M to $\mathcal{L}'$*. Given $X \subseteq dom(M)$, $M[X]$ is the $\mathcal{L}$-structure with domain X such that for all $\ell \geq 1$, $\bar{a} \in X^\ell$, and ℓ-ary relation symbols R from $\mathcal{L}$, we have $M[X] \models R(\bar{a})$ if and only if $M \models R(\bar{a})$. We call $M[X]$ the *$\mathcal{L}$-structure induced by M on X*. Given a tuple $\bar{a} \in dom(M)^\ell$, the *quantifier-free type* of $\bar{a}$ is

$$qftp^M(\bar{a}) = \{\phi(x_1, \ldots, x_\ell) : \phi(x_1, \ldots, x_\ell) \text{ is a quantifier-free } \mathcal{L}\text{-formula and } M \models \phi(\bar{a})\}.$$

If $\bar{x} = (x_1, \ldots, x_\ell)$ and $p(\bar{x})$ is a set of quantifier-free $\mathcal{L}$-formulas, then p is a *quantifier-free ℓ-type* if there is some $\mathcal{L}$-structure N and a tuple $\bar{a} \in dom(N)^\ell$ such that $N \models \phi(\bar{a})$ for all $\phi(\bar{x}) \in p$. In this case we say $\bar{a}$ *realizes p in N*. If there is some $\bar{a} \in dom(N)^\ell$ realizing p in N, we say *p is realized in* N. A quantifier-free ℓ-type $p(\bar{x})$ is *complete* if for every quantifier-free formula $\phi(\bar{x})$, either $\phi(\bar{x})$ or $\neg\phi(\bar{x})$ is in $p(\bar{x})$. Note that any type of the form $qftp^M(\bar{a})$ is complete. All types and formulas we consider will be quantifier-free, so for the rest of the paper, any use of the words type and formula means quantifier-free type and quantifier-free formula.

If X and Y are both $\mathcal{L}$-structures, $X \subseteq_{\mathcal{L}} Y$ denotes that X is a $\mathcal{L}$-substructure of Y. Given an $\mathcal{L}$-structure H, M is *H-free* if there is no $A \subseteq_{\mathcal{L}} M$ such that $A \cong_{\mathcal{L}} H$. Suppose $\mathcal{H}$ is a collection of $\mathcal{L}$-structures. Then M is *$\mathcal{H}$-free* if M is H-free for all $H \in \mathcal{H}$. For each positive integer n, $\mathcal{H}(n)$ denotes the collection of all elements in $\mathcal{H}$ of size n, and $\mathcal{H}_n$ denotes the set of elements in $\mathcal{H}$ with domain $[n]$. $\mathcal{H}$ is *trivial* if there is N such that $\mathcal{H}(n) = \emptyset$ for all $n \geq N$. Otherwise $\mathcal{H}$ is *non-trivial.*

We now define a modified version of the traditional type space, which is appropriate for working with families of finite structures instead of with complete first-order theories. If $\bar{x} = (x_1, \ldots, x_\ell)$, an ℓ-type $p(\bar{x})$ is *proper* if it contains the formulas $x_i \neq x_j$ for each $i \neq j$.

Definition 2. Suppose $\mathcal{F}$ is a collection of $\mathcal{L}$-structures and $\ell \geq 1$ is an integer. Define $S_\ell(\mathcal{F})$ to be the set of all complete, proper, quantifier-free ℓ-types which are realized in some element of $\mathcal{F}$. Let $S_\ell(\mathcal{L})$ denote the set of all complete, proper, quantifier-free ℓ-types.

We would like to emphasize some important differences between this and the usual type space. First, the elements of these type spaces are proper and contain only quantifier-free formulas. Second, these type spaces are defined relative to families of finite structures instead of complete theories.

It will at times be convenient to expand our languages to contain constant symbols naming elements of the structures under consideration. If V is a set, let C_V denote the set of constant symbols $\{c_v : v \in V\}$. Given $\bar{v} = (v_1, \ldots, v_\ell) \in V^\ell$, let $c_{\bar{v}} = (c_{v_1}, \ldots, c_{v_\ell})$. Suppose M is an $\mathcal{L}$-structure. The *diagram of M*, denoted $Diag(M)$, is the following set of sentences in the language $\mathcal{L} \cup C_{dom(M)}$.

$$Diag(M) = \{\phi(c_{\bar{a}}) : \phi(\bar{x}) \text{ is a quantifier-free } \mathcal{L}\text{-formula, } \cup\bar{a} \subseteq dom(M), \text{ and } M \models \phi(\bar{a})\}.$$

If $A \subseteq dom(M)$, the *diagram of A in M* is the following set of sentences in the language $\mathcal{L} \cup C_A$.

$$Diag^M(A) = \{\phi(c_{\bar{a}}) : \phi(\bar{x}) \text{ is a quantifier-free } \mathcal{L}\text{-formula, } \cup\bar{a} \subseteq A, \text{ and } M \models \phi(\bar{a})\}.$$

Observe that if $A = \{a_1, \ldots, a_\ell\} \subseteq dom(M)$ and $p(\bar{x}) \in S_\ell(\mathcal{L})$ is such that $p(\bar{x}) = qftp^M(a_1, \ldots, a_\ell)$, then $Diag^M(A) = p(c_{a_1}, \ldots, c_{a_\ell})$. Given a set of constants C, a collection of $\mathcal{L}$-structures $\mathcal{F}$, set

$$S_\ell(C) = \{p(\bar{c}) : p(\bar{x}) \in S_\ell(\mathcal{L}) \text{ and } \bar{c} \in C^{\underline{\ell}}\} \quad \text{and} \quad S_\ell(C, \mathcal{F}) = \{p(\bar{c}) : p(\bar{x}) \in S_\ell(\mathcal{F}) \text{ and } \bar{c} \in C^{\underline{\ell}}\}.$$

We would like to emphasize that if $p(\bar{c}) \in S_\ell(C)$, then $\bar{c} \in C^{\underline{\ell}}$ is a tuple of ℓ *distinct* constants. Note that by this definition, if $A \in \binom{dom(M)}{\ell}$, then $Diag^M(A) \in S_\ell(C_{dom(M)})$.

2.2. **Hypergraph Containers Theorem.** In this subsection we state a version of the hypergraph containers theorem, which was independently developed by Balogh-Morris-Samotij in [14] and by Saxton-Thomason in [37]. The particular statement we use, Theorem 1 below, is a simplified version of Corollary 3.6 in [37]. We will use Theorem 1 directly in Section 9. We also think it will be useful for the reader to compare it to the versions for $\mathcal{L}$-structures stated in Section 4 (Theorem 6 and Corollary 8).

We begin with some definitions. Recall that if $s \geq 2$ is an integer, an *s-uniform hypergraph* is a pair (V, E) where V is a set of *vertices* and $E \subseteq \binom{V}{s}$ is a set of *edges*. Suppose H is an s-uniform hypergraph. Then $V(H)$ and $E(H)$ denote the vertex and edge sets of G respectively. We set $v(H) = |V(H)|$ and $e(H) = |E(H)|$. Given $X \subseteq V(H)$, $H[X]$ is the hypergraph $(X, E \cap \binom{V(H)}{s})$. If $v(H)$ is finite and $e(H) > 0$, then the *average degree of* H is $d = e(H)s/v(H)$. If $e(H) = 0$ the average degree of H is $d = 0$.

Definition 3. Suppose $s \geq 2$, H is a finite s-uniform hypergraph with n vertices and $\tau > 0$.

- For every $\sigma \subseteq V(H)$, the *degree of* σ *in* H is $d(\sigma) = |\{e \in E(H) : \sigma \subseteq e\}|$.
- Given $v \in V(H)$ and $j \in [s]$, set $d^{(j)}(v) = \max\{d(\sigma) : v \in \sigma \subseteq V(H), |\sigma| = j\}$.
- If $d > 0$, then for each $j \in [s]$, define $\delta_j = \delta_j(\tau)$ to satisfy the equation
$$\delta_j \tau^{j-1} n d = \sum_{v \in V(H)} d^{(j)}(v), \qquad \text{and set} \qquad \delta(H,\tau) = 2^{\binom{s}{2}-1} \sum_{j=2}^{s} 2^{-\binom{j-1}{2}} \delta_j.$$
If $d = 0$, set $\delta(H, \tau) = 0$. $\delta(H, \tau)$ is called the *co-degree function*.

Unless otherwise stated, n is always a positive integer.

Theorem 1 (**Saxton-Thomason [37]**)**.** *Let H be an ℓ-uniform hypergraph with a vertex set V of size n. Suppose $0 < \epsilon, \tau < \frac{1}{2}$ and τ satisfies $\delta(H, \tau) \leq \epsilon/12s!$. Then there exists a constant $c = c(s)$ and a collection $\mathcal{C} \subseteq \mathcal{P}(V)$ such that the following hold.*

(i) For every independent set I in H, there exists $C \in \mathcal{C}$ such that $I \subseteq C$.
(ii) For all $C \in \mathcal{C}$, we have $e(H[C]) \leq \epsilon e(G)$, and
(iii) $\log |\mathcal{C}| \leq c \log(1/\epsilon) n\tau \log(1/\tau)$.

2.3. **Distance between first-order structures.** In this subsection we define a notion of distance between finite first-order structures (Definition 5 below). It is a simplified version of the distance notion appearing in [5].

Definition 4. Suppose $\mathcal{L}$ is a first-order language, B is a finite $\mathcal{L}$-structure of size ℓ, and M is a finite $\mathcal{L}$-structure of size L.

- The *set of copies of* B *in* M is $\mathrm{cop}(B, M) = \{A : A \subseteq_{\mathcal{L}} M \text{ and } A \cong_{\mathcal{L}} B\}$.
- The *induced structure density of* B *in* M is $\mathrm{prob}(B, M) = |\mathrm{cop}(B, M)|/\binom{L}{\ell}$
- If $\mathcal{B}$ is a set of finite $\mathcal{L}$-structures, let
$$\mathrm{cop}(\mathcal{B}, M) = \bigcup_{B \in \mathcal{B}} \mathrm{cop}(B, M) \qquad \text{and} \qquad \mathrm{prob}(\mathcal{B}, M) = \max\{p(B, M) : B \in \mathcal{B}\}.$$

If $\mathcal{B}$ is a *class* of finite $\mathcal{L}$-structures, define $\mathrm{cop}(\mathcal{B}, M) = \mathrm{cop}(\mathcal{B}', M)$ and $\mathrm{prob}(\mathcal{B}, M) = \mathrm{prob}(\mathcal{B}', M)$, where $\mathcal{B}'$ is any set containing one representative of each isomorphism type in B.

Definition 5. Let $\mathcal{L}$ be a finite relational first-order language with $r_{\mathcal{L}} = r$. Suppose M and N are two finite $\mathcal{L}$-structures with the same underlying set V of size n. Let
$$\mathrm{diff}(M, N) = \{A \in \binom{V}{r} : \text{ for some enumeration } \bar{a} \text{ of } A,\ qftp^M(\bar{a}) \neq qftp^N(\bar{a})\}.$$

Then $\mathrm{dist}(M, N) = |\mathrm{diff}(M, N)|/\binom{n}{r}$, and given $\delta > 0$, M and N are *δ-close* if $\mathrm{dist}(M, N) \leq \delta$.

Observe that in the notation of Definition 5, $\text{diff}(M,N) = \{A \in \binom{V}{r} : Diag^M(A) \neq Diag^N(A)\}$. Clearly M and N are the same $\mathcal{L}$-structure if and only if $\text{diff}(M,N) = \emptyset$.

2.4. **A fact about hereditary properties.** Suppose $\mathcal{L}$ is a finite relational language. In this subsection we state a well known fact about hereditary $\mathcal{L}$-properties, Observation 1 below.

Definition 6. If $\mathcal{F}$ is a collection of finite $\mathcal{L}$-structures, let $\text{Forb}(\mathcal{F})$ be the class of all finite $\mathcal{L}$-structures which are $\mathcal{F}$-free.

It is easy to check that for any collection $\mathcal{F}$ of finite $\mathcal{L}$-structures, $\text{Forb}(\mathcal{F})$ is a hereditary $\mathcal{L}$-property. The converse to this statement is also true in the sense of Observation 1 below, which will be used throughout the paper.

Ovservation 1. *If $\mathcal{H}$ is a hereditary $\mathcal{L}$-property, then there is a class of finite $\mathcal{L}$-structures $\mathcal{F}$ which is closed under isomorphism and such that $\mathcal{H} = \text{Forb}(\mathcal{F})$.*

Observation 1 is straightforward to prove by taking $\mathcal{F}$ to be the class of all finite $\mathcal{L}$-structures A which have the property that $A \not\subseteq_{\mathcal{L}} M$ for all M in $\mathcal{H}$.

3. $\mathcal{L}_{\mathcal{H}}$-STRUCTURES

From now on, $\mathcal{L}$ is a fixed finite relational language and $r := r_{\mathcal{L}}$. Note this means $r \geq 1$. For this section, $\mathcal{H}$ is a nonempty collection of finite $\mathcal{L}$-structures. In this section we introduce a language $\mathcal{L}_{\mathcal{H}}$ associated to $\mathcal{L}$ and $\mathcal{H}$. Structures in this new language play key roles in our main theorems.

Definition 7. Define $\mathcal{L}_{\mathcal{H}} = \{R_p(\bar{x}) : p(\bar{x}) \in S_r(\mathcal{H})\}$ to be the relational language with one r-ary relation for each $p(\bar{x})$ in $S_r(\mathcal{H})$.

The goal of this section is to formalize how an $\mathcal{L}_{\mathcal{H}}$-structure M with the right properties can serve as a "template" for building $\mathcal{L}$-structures with the same underlying set as M. We now give an example of a hereditary property and its corresponding auxiliary language as in Definition 7.

Example 1. To avoid confusion, we will use $\mathcal{P}$ to refer to specific hereditary properties in example settings. Suppose $\mathcal{L} = \{d_1(x,y), d_2(x,y), d_3(x,y)\}$. Let $\mathcal{P}$ be the class of all finite metric spaces with distances in $\{1,2,3\}$, considered as $\mathcal{L}$-structures in the natural way (i.e. $d_i(x,y)$ if and only if $d(x,y) = i$). It is easy to see that $\mathcal{P}$ is a hereditary $\mathcal{L}$-property. Observe that since $r_{\mathcal{L}} = 2$, $\mathcal{L}_{\mathcal{P}} = \{R_p(x,y) : p(x,y) \in S_2(\mathcal{P})\}$. For each $i \in [3]$, set

$$q_i(x,y) := \{x \neq y\} \cup \{R_i(x,y), R_i(y,x)\} \cup \{\neg R_j(x,y), \neg R_j(y,x) : j \neq i\},$$

and let $p_i(x,y)$ be the unique quantifier-free 2-type containing $q_i(x,y)$. Informally, the type $p_i(x,y)$ says "the distance between x and y is equal to i." We leave it as an exercise to the reader that $S_2(\mathcal{P}) = \{p_i(x,y) : i \in [3]\}$ (recall $S_2(\mathcal{P})$ consists of *proper* types). Thus $\mathcal{L}_{\mathcal{P}} = \{R_{p_i}(x,y) : i \in [3]\}$.

In an arbitrary $\mathcal{L}_{\mathcal{H}}$-structure, the relation symbols in $\mathcal{L}_{\mathcal{H}}$ may have nothing to do with the properties of the type space $S_r(\mathcal{H})$. For instance, in the notation of Example 1, we can easily build an $\mathcal{L}_{\mathcal{P}}$-structure M so that for some $a,b \in dom(M)$, $M \models R_{p_1}(a,b) \wedge \neg R_{p_1}(b,a)$, even though $p_1(x,y) = p_1(y,x)$ in $S_2(\mathcal{P})$. This kind of behavior will be undesirable for various technical reasons. We now define the class of $\mathcal{L}_{\mathcal{H}}$-structures which are most nicely behaved for our purposes, and where in particular, this bad behavior does not happen.

Definition 8. An $\mathcal{L}_{\mathcal{H}}$-structure M with domain V is *complete* if for all $A \in \binom{V}{r}$ there is an enumeration $\bar{a}$ of A and $R_p \in \mathcal{L}_{\mathcal{H}}$ such that $M \models R_p(\bar{a})$.

Definition 9. An $\mathcal{L}_{\mathcal{H}}$-structure M with domain V is an *$\mathcal{L}_{\mathcal{H}}$-template* if it is complete and the following hold.

(1) If $p(\bar{x}) \in S_r(\mathcal{H})$ and $\bar{a} \in V^r \setminus V^{\underline{r}}$, then $M \models \neg R_p(\bar{a})$.

(2) If $p(\bar{x}), p'(\bar{x}) \in S_r(\mathcal{H})$ and $\mu \in Perm(r)$ are such that $p(\bar{x}) = p'(\mu(\bar{x}))$, then for every $\bar{a} \in V^{\underline{r}}$, $M \models R_p(\bar{a})$ if and only if $M \models R_{p'}(\mu(\bar{a}))$.

The idea is that $\mathcal{L}_{\mathcal{H}}$-templates most accurately reflect the properties of $S_r(\mathcal{H})$.

Example 2. Let $\mathcal{L}$ and $\mathcal{P}$ be as in Example 1. Let M be the $\mathcal{L}_{\mathcal{P}}$-structure with underlying set $V = \{u, v, w\}$ such that $M \models \bigwedge_{i=1}^{3}(R_{p_i}(u, v) \wedge R_{p_i}(v, u))$,

$$M \models R_{p_1}(w, v) \wedge R_{p_1}(v, w) \wedge R_{p_2}(w, v) \wedge R_{p_2}(v, w) \wedge \neg R_{p_3}(w, v) \wedge \neg R_{p_3}(v, w),$$
$$M \models R_{p_1}(w, u) \wedge R_{p_1}(w, u) \wedge \neg R_{p_2}(w, u) \wedge \neg R_{p_2}(u, w) \wedge \neg R_{p_3}(w, u) \wedge \neg R_{p_3}(u, w),$$

and for all $x \in V$, $M \models \bigwedge_{i=1}^{3} \neg R_{p_i}(x, x)$. We leave it to the reader to verify M is a $\mathcal{L}_{\mathcal{P}}$-template.

While $\mathcal{L}_{\mathcal{H}}$-templates are important for the main results of this paper, many of the definitions and facts in the rest of this section will be presented for $\mathcal{L}_{\mathcal{H}}$-structures with weaker assumptions.

3.1. **Choice functions and subpatterns.** In this subsection, we give crucial definitions for how we can use $\mathcal{L}_{\mathcal{H}}$-structures to build $\mathcal{L}$-structures.

Definition 10. Suppose M is an $\mathcal{L}_{\mathcal{H}}$-structure with domain V.

(1) Given $A \in \binom{V}{r}$, the *set of choices for A in M* is

$$Ch_M(A) = \{p(c_{a_1}, \ldots, c_{a_r}) \in S_r(C_V, \mathcal{H}) : \{a_1, \ldots, a_r\} = A \text{ and } M \models R_p(a_1, \ldots, a_r)\}.$$

(2) A *choice function for M* is a function $\chi : \binom{V}{r} \to S_r(C_V, \mathcal{H})$ such that for each $A \in \binom{V}{r}$, $\chi(A) \in Ch_M(A)$. Let $Ch(M)$ denote the set of all choice functions for M.

Observe that in the notation of Definition 10, $Ch(M) \neq \emptyset$ if and only if M is complete.

Example 3. Recall that if x and y are distinct elements of a set, then xy is shorthand for the set $\{x, y\}$. Let $\mathcal{L}$, $\mathcal{P}$, V, and M be as in Example 2. Then $C_V = \{c_u, c_v, c_w\}$, and

$$S_2(C_V, \mathcal{P}) = \{p_i(c_u, c_v) : i \in [3]\} \cup \{p_i(c_v, c_w) : i \in [3]\} \cup \{p_i(c_u, c_w) : i \in [3]\}.$$

By definition of M, $Ch_M(uv) = \{p_1(c_u, c_v), p_2(c_u, c_v), p_3(c_u, c_v)\}$, $Ch_M(vw) = \{p_1(c_v, c_w), p_2(c_v, c_w)\}$, and $Ch_M(uw) = \{p_1(c_u, c_w)\}$. Therefore $Ch(M)$ is the set of functions

$$\chi : \{uv, vw, uw\} \to \{p_i(c_u, c_v) : i \in [3]\} \cup \{p_i(c_v, c_w) : i \in [3]\} \cup \{p_i(c_u, c_w) : i \in [3]\}$$

with the properties that $\chi(uv) \in \{p_1(c_u, c_v), p_2(c_u, c_v), p_3(c_u, c_v)\}$, $\chi(vw) \in \{p_1(c_v, c_w), p_2(c_v, c_w)\}$ and $\chi(uw) = p_1(c_u, c_w)$. Clearly this shows $|Ch(M)| = |Ch_M(uv)||Ch_M(vw)||Ch_M(uw)| = 6$.

The following observation is immediate from the definition of $\mathcal{L}_{\mathcal{H}}$-template.

Ovservation 2. *If M is an $\mathcal{L}_{\mathcal{H}}$-template with domain V, then for all $\bar{a} \in V^r$ and $R_p \in \mathcal{L}_{\mathcal{H}}$, $M \models R_p(\bar{a})$ if and only if $|\cup \bar{a}| = r$ and $p(c_{\bar{a}}) \in Ch_M(\cup \bar{a})$.*

Proposition 1 below is one reason why $\mathcal{L}_{\mathcal{H}}$-templates are convenient.

Proposition 1. *Suppose M_1 and M_2 are $\mathcal{L}_{\mathcal{H}}$-templates with domain V. Then*

$$\text{diff}(M_1, M_2) = \{A \in \binom{V}{2} : Ch_{M_1}(A) \neq Ch_{M_2}(A)\}. \tag{2}$$

Consequently, if M_1 and M_2 satisfy $Ch_{M_1}(A) = Ch_{M_2}(A)$ for all for all $A \in \binom{V}{r}$, then M_1 and M_2 are the same $\mathcal{L}_{\mathcal{H}}$-structure.

Proof. Fix $A \in \binom{V}{r}$ and an enumeration $\bar{a}$ of A. By Observation 2, for all $R_p \in \mathcal{L}_{\mathcal{H}}$, $M_1 \models R_p(\bar{a})$ if and only if $p(c_{\bar{a}}) \in Ch_{M_1}(\cup \bar{a})$, and $M_2 \models R_p(\bar{a})$ if and only if $p(c_{\bar{a}}) \in Ch_{M_2}(\cup \bar{a})$. Therefore $qftp^{M_1}(\bar{a}) = qftp^{M_2}(\bar{a})$ if and only if $Ch_{M_1}(\cup \bar{a}) = Ch_{M_2}(\cup \bar{a})$. By definition, this means that $A \notin \text{diff}(M_1, M_2)$ if and only if $Ch_{M_1}(\cup \bar{a}) = Ch_{M_2}(\cup \bar{a})$, which implies (2). Consequently, if M_1 and M_2 satisfy $Ch_{M_1}(A) = Ch_{M_2}(A)$ for all for all $A \in \binom{V}{r}$, then $\text{diff}(M_1, M_2) = \emptyset$ implies M_1 and M_2 are the same $\mathcal{L}_{\mathcal{H}}$-structure. □

The next example shows Proposition 1 can fail when we are not dealing with $\mathcal{L}_{\mathcal{H}}$-templates.

Example 4. Let $\mathcal{L}$, $\mathcal{P}$, V, and M be as in Example 5. Let M' be the $\mathcal{L}_{\mathcal{P}}$-structure with domain V which agrees with M on $V^2 \setminus \{(v,u),(w,v),(w,u)\}$ and where

$$M' \models \bigwedge_{i=1}^{3} (\neg R_{p_i}(v,u) \wedge \neg R_{p_i}(w,v) \wedge \neg R_{p_i}(w,u)).$$

We leave it to the reader to check that for all $xy \in \binom{V}{2}$, $Ch_{M'}(xy) = Ch_M(xy)$. However, M and M' are not the same $\mathcal{L}_{\mathcal{H}}$-structure. In particular, $uv \in \text{diff}(M,M')$ since $M \models R_{p_1}(v,u)$ and $M' \models \neg R_{p_1}(v,u)$. Observe that M' is not an $\mathcal{L}_{\mathcal{P}}$-template because $M' \models R_{p_1}(u,v) \wedge \neg R_{p_1}(v,u)$ while $p_1(x,y) = p_1(y,x)$.

The next definition shows how choice functions give rise to $\mathcal{L}$-structures.

Definition 11. Suppose M is a complete $\mathcal{L}_{\mathcal{H}}$-structure with domain V, N is an $\mathcal{L}$-structure such that $dom(N) \subseteq V$, and $\chi \in Ch(M)$.

(1) N is a *χ-subpattern* of M, denoted $N \leq_\chi M$, if for every $A \in \binom{dom(N)}{r}$, $\chi(A) = Diag^N(A)$.
(2) N is a *full χ-subpattern of* M, denoted $N \trianglelefteq_\chi M$, if $N \leq_\chi M$ and $dom(N) = V$.

When $N \trianglelefteq_\chi M$, we say χ *chooses* N. We say N is a *subpattern of* M, denoted $N \leq_p M$, if $N \leq_\chi M$ for some choice function χ for M. We say N is a *full subpattern of* M, denoted $N \trianglelefteq_p M$, if $N \trianglelefteq_\chi M$ for some choice function χ for M. The subscript in $\leq_p$ and $\trianglelefteq_p$ is for "pattern."

Ovservation 3. *Suppose M is a complete $\mathcal{L}_{\mathcal{H}}$-structure, $\chi \in Ch(M)$, and G is an $\mathcal{L}$-structure such that $G \trianglelefteq_\chi M$. If G' is another $\mathcal{L}$-structure such that $G' \trianglelefteq_\chi M$, then G and G' are the same $\mathcal{L}$-structure. If $\chi' \in Ch(M)$ satisfies $G \trianglelefteq_{\chi'} M$, then $\chi = \chi'$.*

Proof. By definition, $G \trianglelefteq_\chi M$ and $G' \trianglelefteq_\chi M$ imply that $Diag^G(A) = \chi(A) = Diag^{G'}(A)$ for all $A \in \binom{V}{r}$. This implies $Diag(G) = Diag(G')$, and thus G and G' are the same $\mathcal{L}$-structure. If $G \trianglelefteq_{\chi'} M$, then for all $A \in \binom{V}{r}$, $\chi(A) = Diag^G(A) = \chi'(A)$. Thus $\chi = \chi'$. □

Example 5. Let $\mathcal{L}$, $\mathcal{P}$, V and M be as in Example 2. We give two examples of subpatterns of M. Let χ be the function from $\binom{V}{2} \to S_2(C_V, \mathcal{P})$ defined by $\chi(uv) = p_1(c_u,c_v)$, $\chi(vw) = p_2(c_v,c_w)$, and $\chi(uw) = p_1(c_u,c_w)$. Clearly $\chi \in Ch(M)$. Let H be the $\mathcal{L}$-structure with domain V which satisfies $H \models p_1(u,v) \cup p_2(v,w) \cup p_1(u,w)$. By definition, $Diag^H(uv) = p_1(c_u,c_v)$, $Diag^H(vw) = p_2(v,w)$, and $Diag^H(uw) = p_1(c_u,c_w)$. In other words, $H \leq_\chi M$. Since $dom(H) = dom(M) = V$, $H \trianglelefteq_\chi M$. Note that H is a metric space, that is, $H \in \mathcal{P}$.

Let χ' be the function from $\binom{V}{2} \to S_2(C_V, \mathcal{P})$ defined by $\chi'(uv) = p_3(c_u,c_v)$, $\chi'(vw) = p_1(c_v,c_w)$, and $\chi'(uw) = p_1(c_u,c_w)$. Clearly $\chi' \in Ch(M)$. Let H' be the $\mathcal{L}$-structure with domain V such that $H' \models p_3(u,v) \cup p_1(v,w) \cup p_1(u,w)$. Then as above, it is easy to see that $H' \leq_{\chi'} M$, and since $dom(H') = V$, $H' \trianglelefteq_{\chi'} M$. However, H' is *not* a metric space, that is, $H' \notin \mathcal{P}$.

Example 5 demonstrates that although $\mathcal{L}_{\mathcal{H}}$-templates are well behaved in certain ways, an $\mathcal{L}_{\mathcal{H}}$-template may have full subpatterns that are not in $\mathcal{H}$. We will give further definitions to address this in Section 3.3.

3.2. **Errors and counting subpatterns.**

In this subsection we characterize when an $\mathcal{L}_{\mathcal{H}}$-structure has the property that every choice function gives rise to a subpattern. This will be important for counting subpatterns of $\mathcal{L}_{\mathcal{H}}$-structures.

Definition 12. Given $r < \ell < 2r$, an *error of size* ℓ is a complete $\mathcal{L}_{\mathcal{H}}$-structure M of size ℓ with the following properties. There are $\bar{a}_1, \bar{a}_2 \in dom(M)^{\underline{r}}$ such that $dom(M) = \cup\bar{a}_1 \bigcup \cup\bar{a}_2$ and for some $p_1(\bar{x}), p_2(\bar{x}) \in S_r(\mathcal{H})$, $M \models R_{p_1}(\bar{a}_1) \wedge R_{p_2}(\bar{a}_2)$ but $p_1(c_{\bar{a}_1}) \cup p_2(c_{\bar{a}_2})$ is unsatisfiable.

Definition 13. Let $\mathcal{E}$ be the class of $\mathcal{L}_{\mathcal{H}}$-structures which are errors of size ℓ for some $r < \ell < 2r$. An $\mathcal{L}_{\mathcal{H}}$-structure M is *error-free* if it is $\mathcal{E}$-free.

We leave it to the reader to verify that if all the relation symbols in $\mathcal{L}$ have the same arity, then $\mathcal{E} = \emptyset$ (this includes the case where $r = 1$). However the next example demonstrates $\mathcal{E}$ may be nonempty when $\mathcal{L}$ contains relations of different arities.

Example 6. Let $\mathcal{L} = \{E(x,y,z), d_1(x,y), d_2(x,y), d_3(x,y)\}$. Suppose $\mathcal{P}$ is the class of all finite $\mathcal{L}$-structures G such that the reduct of G to $\{d_1,d_2,d_3\}$ is a metric space with distances in [3] (we put no restrictions on E). For $i \in [3]$, let $p_i(x,y)$ be the quantifier-free $\{d_1,d_2,d_3\}$-type from Examples 1-5. Let $\bar{x} = (x_1,x_2,x_3)$ and define $q_1(\bar{x})$ and $q_1(\bar{x})$ to be the elements of $S_3(\mathcal{P})$ determined by the following information.

$$\{E(x_i,x_j,x_k) : 1 \leq i,j,k \leq 3\} \cup p_1(x_1,x_2) \cup p_1(x_1,x_3) \cup p_1(x_2,x_3) \subseteq q_1(\bar{x}) \text{ and}$$
$$\{E(x_i,x_j,x_k) : 1 \leq i,j,k \leq 3\} \cup p_2(x_1,x_2) \cup p_1(x_1,x_3) \cup p_1(x_2,x_3) \subseteq q_2(\bar{x}).$$

Observe that q_1 and q_2 agree about how E behaves, but disagree on how the binary relations in $\mathcal{L}$ behave. Let $V = \{t,u,v,w\}$ be a set of size 4. Choose M to be the $\mathcal{L}_{\mathcal{P}}$-structure which satisfies $M \models R_{q_1}(x,y,z)$ if and only if x, y and z are distinct, $M \models R_{q_2}(x,y,z)$ if and only if $(x,y,z) = (t,u,v)$, and $M \models \neg R_q(x,y,z)$ for all $q \in S_3(\mathcal{P}) \setminus \{p_1,p_2\}$. By construction, M is a complete $\mathcal{L}_{\mathcal{P}}$-structure. Let $\bar{a}_1 = (u,v,w)$ and $\bar{a}_2 = (u,v,t)$. Then $dom(M) = \cup\bar{a}_1 \bigcup \cup\bar{a}_2$ and $M \models R_{q_1}(\bar{a}_1) \wedge R_{q_2}(\bar{a}_2)$. However, $d_1(c_u,c_v) \in p_1(c_u,c_v) \subseteq q_1(c_u,c_v,c_w) = q_1(c_{\bar{a}_1})$ while $\neg d_1(c_u,c_v) \in p_2(c_u,c_v) \subseteq q_2(c_u,c_v,c_w) = q_2(c_{\bar{a}_2})$. Therefore $q_1(c_{\bar{a}_1}) \cup q_2(c_{\bar{a}_2})$ is unsatisfiable, and M is an error of size 4.

Error-free $\mathcal{L}_{\mathcal{H}}$-structures will be important for the following reason.

Proposition 2. *Suppose M is a complete $\mathcal{L}_{\mathcal{H}}$-structure with domain V. Then M is error-free if and only if for every $\chi \in Ch(M)$, there is an $\mathcal{L}$-structure N such that $N \trianglelefteq_{\chi} M$.*

Proof. Suppose first that there exists a choice function $\chi : \binom{V}{r} \rightarrow S_r(C_V,\mathcal{H})$ such that there are no χ-subpatterns of M. This means $\Gamma := \bigcup_{A \in \binom{V}{r}} \chi(A)$ is not satisfiable. Consequently, there is an atomic formula $\psi(\bar{x})$ and a tuple $\bar{c} \subseteq C_V^r$ such that $\{\psi(\bar{c}), \neg\psi(\bar{c})\} \subseteq \Gamma$. For each $A \in \binom{V}{r}$, because $\chi(A) \in S_r(C_A,\mathcal{H})$, exactly one of $\psi(\bar{c})$ or $\neg\psi(\bar{c})$ is in $\chi(A)$. This implies there must be distinct A_1, $A_2 \in \binom{V}{r}$ such that $\cup\bar{c} \subseteq A_1 \cap A_2$, $\psi(\bar{c}) \in \chi(A_1)$, and $\neg\psi(\bar{c}) \in \chi(A_2)$. Note $A_1 \neq A_2$ and $A_1 \cap A_2 \neq \emptyset$ imply that $r < \ell := |A_1 \cup A_2| < 2r$. Let N be the $\mathcal{L}_{\mathcal{H}}$-structure $M[A_1 \cup A_2]$. We show N is an error of size ℓ. By definition of χ being a choice function, there are $p_1,p_2 \in S_r(C_V,\mathcal{H})$ and $\bar{a}_1$, $\bar{a}_2$ such that $\cup\bar{a}_1 = A_1$, $\cup\bar{a}_2 = A_2$, $p_1(c_{\bar{a}_1}) = \chi(A_1)$, $p_2(c_{\bar{a}_2}) = \chi(A_2)$, $M \models R_{p_1}(\bar{a}_1)$, and $M \models R_{p_2}(\bar{a}_2)$. By definition, $N \subseteq_{\mathcal{L}_{\mathcal{H}}} M$, thus $N \models R_{p_1}(\bar{a}_1) \wedge R_{p_2}(\bar{a}_2)$. Then

$$\{\psi(\bar{c}), \neg\psi(\bar{c})\} \subseteq \chi(A_1) \cup \chi(A_2) = p_1(c_{\bar{a}_1}) \cup p_2(c_{\bar{a}_2})$$

implies $p_1(c_{\bar{a}_1}) \cup p_2(c_{\bar{a}_2})$ is unsatisfiable. Thus $N \in \mathcal{E}$ and $N \subseteq_{\mathcal{L}_{\mathcal{H}}} M$ implies M is not error-free.

Suppose on the other hand that M is not error-free. Say $r < \ell < 2r$ and N is an error of size ℓ in M. Then $N \subseteq_{\mathcal{L}_{\mathcal{H}}} M$ and there are $\bar{a}_1, \bar{a}_2 \in dom(N)^{\underline{r}}$ and types $p_1(\bar{x}), p_2(\bar{x}) \in S_r(\mathcal{H})$ such that $dom(N) = \cup\bar{a}_1 \bigcup \cup\bar{a}_2$, $N \models R_{p_1}(\bar{a}_1) \wedge R_{p_2}(\bar{a}_2)$, and $p_1(c_{\bar{a}_1}) \cup p_2(c_{\bar{a}_2})$ is unsatisfiable. Define a function $\chi : \binom{V}{r} \rightarrow S_r(C_V,\mathcal{H})$ as follows. Set $\chi(\cup\bar{a}_1) = p_1(c_{\bar{a}_1})$ and $\chi(\cup\bar{a}_2) = p_2(c_{\bar{a}_2})$. For every $A' \in \binom{V}{r} \setminus \{A_1,A_2\}$, choose $\chi(A')$ to be any element of $Ch_M(A')$ (note $Ch_M(A')$ is nonempty since M is complete). By construction, $\chi \in Ch(M)$. Suppose there is $G \trianglelefteq_{\chi} M$. Then $p_1(c_{\bar{a}_1}) = Diag^G(\cup\bar{a}_1)$ and $p_2(c_{\bar{a}_2}) = Diag^G(\cup\bar{a}_2)$ implies $G \models p_1(\bar{a}_1) \cup p_2(\bar{a}_2)$, contradicting that $p_1(c_{\bar{a}_1}) \cup p_2(c_{\bar{a}_2})$ is unsatisfiable. Thus $\chi \in Ch(M)$ and there are no χ-subpatterns of M. $\square$

Definition 14. Given a finite $\mathcal{L}_{\mathcal{H}}$-structure M, let $sub(M) = |\{G : G \trianglelefteq_p M\}|$ be the number of full subpatterns of M.

Definition 14 and the following observation will be crucial to our enumeration theorem.

Ovservation 4. *If M is a complete $\mathcal{L}_{\mathcal{H}}$-structure with finite domain V, then*

$$sub(M) \leq \prod_{A \in \binom{V}{r}} |Ch_M(A)|,$$

and equality holds if and only if M is error-free.

Proof. By definition of a choice function, $|Ch(M)| = \prod_{A\in\binom{V}{r}} |Ch_M(A)|$. By definition of subpattern, for each $G \trianglelefteq_p M$, there is $\chi_G \in Ch(M)$ which chooses G. Observation 3 implies the map $f : G \mapsto \chi_G$ is a well-defined injection from $\{G : G \trianglelefteq_p M\}$ to $Ch(M)$. Thus

$$sub(M) = |\{G : G \trianglelefteq_p M\}| \leq |Ch(M)| = \prod_{A\in\binom{V}{r}} |Ch_M(A)|.$$

We now show equality holds if and only if M is error-free. Suppose first M is error-free. We claim f is surjective. Fix $\chi \in Ch(M)$. Since M is error-free, Proposition 2 implies that there is an $\mathcal{L}$-structure G_χ such that $G_\chi \trianglelefteq_\chi M$. So $G_\chi \in \{G : G \trianglelefteq_p M\}$ implies $f(G_\chi)$ exists. By Observation 3, we must have $f(G_\chi) = \chi$. Thus f is surjective, and consequently $sub(M) = |Ch(M)|$. Conversely, suppose equality holds. Then f is an injective map from a finite set to another finite set of the same size, thus it must be surjective. Therefore, for all $\chi \in Ch(M)$, there is an $\mathcal{L}$-structure G such that $G \trianglelefteq_\chi M$. By Proposition 2, this implies M is error-free. □

Remark 1. *Suppose $\mathcal{L}$ contains no relations of arity less than r. If $\mathcal{M}$ is a complete $\mathcal{L}_{\mathcal{H}}$-structure with finite domain V, then $sub(M) = \prod_{A\in\binom{V}{r}} |Ch_M(A)|$.*

Proof. Since $\mathcal{L}$ consists only of relations of arity r, $\mathcal{E} = \emptyset$. Thus $\mathcal{M}$ is error-free, so Observation 4 implies $sub(M) = \prod_{A\in\binom{V}{r}} |Ch_M(A)|$. □

Remark 1 applies to most examples we are interested in, including graphs, (colored) k-uniform hypergraphs for any $k \geq 2$, directed graphs, and discrete metric spaces.

3.3. **$\mathcal{H}$-random $\mathcal{L}_{\mathcal{H}}$-structures and $\mathcal{L}_{\mathcal{H}}$-templates.** In this subsection we consider $\mathcal{L}_{\mathcal{H}}$-structures with the property that all choice functions give rise to subpatterns in $\mathcal{H}$.

Definition 15. An $\mathcal{L}_{\mathcal{H}}$-structure M is *$\mathcal{H}$-random* if it is complete and for every $\chi \in Ch(M)$, there is an $\mathcal{L}$-structure $N \in \mathcal{H}$ such that $N \trianglelefteq_\chi M$.

Observe that by Proposition 2, any $\mathcal{H}$-random $\mathcal{L}_{\mathcal{H}}$-structure is error-free. The difference between being error-free and being $\mathcal{H}$-random is as follows. If an $\mathcal{L}_{\mathcal{H}}$-structure is error-free, then it must have at least one full subpattern, however some or all its subpatterns may not be in $\mathcal{H}$. On the other hand, if an $\mathcal{L}_{\mathcal{H}}$-structure is $\mathcal{H}$-random, then it must have at least one full subpattern, and further, all its full subpatterns must also be in $\mathcal{H}$.

Example 7. Let $\mathcal{L}$, $\mathcal{P}$, V, H' and M be as in Example 5. Observe that M is *not* $\mathcal{P}$-random, since $H' \trianglelefteq_p M$, but $H' \notin \mathcal{P}$. We now define an $\mathcal{L}_{\mathcal{P}}$-structure M'' which is $\mathcal{P}$-random. Let M'' be any $\mathcal{L}_{\mathcal{P}}$-structure with domain V such that for all $(x, y) \in V^{\underline{2}}$, $M'' \models R_{p_1}(x, y) \wedge R_{p_2}(x, y) \wedge \neg R_{p_3}(x, y)$ and $M \models \bigwedge_{i=1}^3 \neg R_{p_i}(x, x)$ for all $x \in V$. We leave it to the reader to check that M'' is a $\mathcal{L}_{\mathcal{P}}$-template and $Ch_{M''}(xy) = \{p_1(c_x, c_y), p_2(c_x, c_y)\}$ for all $xy \in \binom{V}{2}$. It is then straightforward to see that M'' is $\mathcal{P}$-random (because there is no way to violate the triangle inequality using distances in $\{1, 2\}$).

The most important $\mathcal{L}_{\mathcal{H}}$-structures for the rest of the paper are $\mathcal{H}$-random $\mathcal{L}_{\mathcal{H}}$-templates. We now fix notation for these special $\mathcal{L}_{\mathcal{H}}$-structures.

Definition 16. Suppose V is a set, and n is an integer. Then

- $\mathcal{R}(V,\mathcal{H})$ is the set of all $\mathcal{H}$-random $\mathcal{L}_{\mathcal{H}}$-templates with domain V and
- $\mathcal{R}(n,\mathcal{H})$ is the class of all $\mathcal{H}$-random $\mathcal{L}_{\mathcal{H}}$-templates of size n.

In the above notation, $\mathcal{R}$ is for "random." Note that if $\mathcal{H}(n)=\emptyset$ for some n, then $\mathcal{R}(n,\mathcal{H})=\emptyset$.

4. Main Results

In this section we state the main results of this paper. Recall that $\mathcal{L}$ is a fixed finite relational language of maximum arity $r\geq 1$. We now define our generalization of extremal graphs. By convention, $\max\emptyset=0$.

Definition 17. Suppose $\mathcal{H}$ is a collection of finite $\mathcal{L}$-structures. Given n, set

$$ex(n,\mathcal{H})=\max\{sub(M):M\in\mathcal{R}(n,\mathcal{H})\}.$$

We say $M\in\mathcal{R}(n,\mathcal{H})$ is *extremal* if $sub(M)=ex(n,\mathcal{H})$. If V is a set and $n\in\mathbb{N}$, then

- $\mathcal{R}_{ex}(V,\mathcal{H})$ is the set of extremal elements of $\mathcal{R}(V,\mathcal{H})$ and
- $\mathcal{R}_{ex}(n,\mathcal{H})$ is the class of extremal elements of $\mathcal{R}(n,\mathcal{H})$.

The main idea is that when $\mathcal{H}$ is a hereditary $\mathcal{L}$-property, $\mathrm{ex}(n,\mathcal{H})$ is the correct generalization of the extremal number of a graph, and elements of $\mathcal{R}_{ex}(n,\mathcal{H})$ are the correct generalizations extremal graphs of size n.

Definition 18. Suppose $\mathcal{H}$ is a nonempty collection of finite $\mathcal{L}$-structures. When it exists, set

$$\pi(\mathcal{H})=\lim_{n\to\infty}\mathrm{ex}(n,\mathcal{H})^{1/\binom{n}{r}}$$

Using techniques similar to those in [18] we will show the following.

Theorem 2. *If $\mathcal{H}$ is hereditary $\mathcal{L}$-property, then $\pi(\mathcal{H})$ exists.*

We now state our approximate enumeration theorem in terms of the asymptotic density.

Theorem 3 (Enumeration)**.** *Suppose $\mathcal{H}$ is a hereditary $\mathcal{L}$-property. Then the following hold.*

(1) If $\pi(\mathcal{H})>1$, then $|\mathcal{H}_n|=\pi(\mathcal{H})^{\binom{n}{r}+o(n^r)}$.
(2) If $\pi(\mathcal{H})\leq 1$, then $|\mathcal{H}_n|=2^{o(n^r)}$.

The notion $\pi(\mathcal{H})$ is related to many existing notions of asymptotic density for various combinatorial structures, and Theorem 3 can be seen as generalizing many existing enumeration theorems. We refer the reader to Section 10 and the appendices of [41] for examples. We say a hereditary $\mathcal{L}$-property $\mathcal{H}$ is *fast-growing* if $\pi(\mathcal{H})>1$. In this case, we informally say $M\in\mathcal{R}(n,\mathcal{H})$ is *almost extremal* if $sub(M)\geq \mathrm{ex}(n,\mathcal{H})^{1-\epsilon}$ for some small ϵ. Our next theorem shows that almost all elements in a fast-growing hereditary $\mathcal{L}$-property $\mathcal{H}$ are close to subpatterns of almost extremal elements of $\mathcal{R}(n,\mathcal{H})$. Given $\epsilon>0$, n, and a collection $\mathcal{H}$ of $\mathcal{L}$-structures, let

$$E(n,\mathcal{H})=\{G\in\mathcal{H}_n:G\trianglelefteq_p M\text{ for some }M\in\mathcal{R}_{ex}(n,\mathcal{H})\}\text{ and}$$
$$E(\epsilon,n,\mathcal{H})=\{G\in\mathcal{H}_n:G\trianglelefteq_p M\text{ for some }M\in\mathcal{R}(n,\mathcal{H})\text{ with }sub(M)\geq\mathrm{ex}(n,\mathcal{H})^{1-\epsilon}\}.$$

Given $\delta>0$, let $E^\delta(n,\mathcal{H})$ and $E^\delta(\epsilon,n,\mathcal{H})$ denote the set of $G\in\mathcal{H}_n$ which are δ-close to any element of $E(n,\mathcal{H})$ and $E(\epsilon,n,\mathcal{H})$, respectively.

Theorem 4. *Suppose $\mathcal{H}$ is a fast-growing hereditary $\mathcal{L}$-property. For all $\epsilon,\delta>0$ there is $\beta>0$ such that for sufficiently large n,*

$$\frac{|\mathcal{H}_n\setminus E^\delta(\epsilon,n,\mathcal{H})|}{|\mathcal{H}_n|}\leq 2^{-\beta\binom{n}{r}}.$$

We now define our generalization of a graph stability theorem.

Definition 19. Suppose $\mathcal{H}$ is a nontrivial collection of $\mathcal{L}$-structures. We say $\mathcal{H}$ *has a stability theorem* if for all $\delta > 0$ there is $\epsilon > 0$ and N such that $n > N$ implies the following. If $M \in \mathcal{R}(n, \mathcal{H})$ satisfies $sub(M) \geq \text{ex}(n, \mathcal{H})^{1-\epsilon}$, then M is δ-close to some $M' \in \mathcal{R}_{ex}(n, \mathcal{H})$.

Our next result, Theorem 5 below, shows that if a fast-growing hereditary $\mathcal{L}$-property $\mathcal{H}$ has a stability theorem, we can strengthen Theorem 4 to say that that almost all elements in $\mathcal{H}_n$ are approximately subpatterns of elements of $\mathcal{R}_{ex}(n, \mathcal{H})$.

Theorem 5. *Suppose $\mathcal{H}$ is a fast growing hereditary $\mathcal{L}$-property with a stability theorem. Then for all $\delta > 0$, there is a $\beta > 0$ such that for sufficiently large n,*

$$\frac{|\mathcal{H}_n \setminus E^{\delta}(n, \mathcal{H})|}{|\mathcal{H}_n|} \leq 2^{-\beta\binom{n}{r}}.$$

When one has a good understanding of the structure of elements in $\mathcal{R}_{ex}(n, \mathcal{H})$, Theorem 5 gives us a good description of the approximate structure of most elements in $\mathcal{H}_n$, when n is large. The main new tool we will use to prove our main theorems is Theorem 6 below, which is an adaptation of the hypergraph containers theorem to the setting of $\mathcal{L}$-structures.

Definition 20. If F is an $\mathcal{L}$-structure, let $\tilde{F}$ be the set of $\mathcal{L}_{\mathcal{H}}$-structures M such that $F \trianglelefteq_p M$. If $\mathcal{F}$ is a collection of $\mathcal{L}$-structures, let $\tilde{\mathcal{F}} = \bigcup_{F \in \mathcal{F}} \tilde{F}$.

Theorem 6. *Suppose $0 < \epsilon < 1$ and $k \geq r$ is an integer. Then there exist positive constants $c = c(k, r, \mathcal{L}, \epsilon)$ and $m = m(k, r) \geq 1$ such that for all sufficiently large n the following holds. Assume $\mathcal{F}$ is a collection of finite $\mathcal{L}$-structures each of size at most k and $\mathcal{B} := \text{Forb}(\mathcal{F}) \neq \emptyset$. For any n-element set W, there is a collection $\mathcal{C}$ of $\mathcal{L}_{\mathcal{B}}$-templates with domain W such that*

(1) For all $\mathcal{F}$-free $\mathcal{L}$-structures M with domain W, there is $C \in \mathcal{C}$ such that $M \trianglelefteq_p C$,
(2) For all $C \in \mathcal{C}$, $\text{prob}(\tilde{\mathcal{F}}, C) \leq \epsilon$ and $\text{prob}(\mathcal{E}, C) \leq \epsilon$.
(3) $\log |\mathcal{C}| \leq cn^{r-\frac{1}{m}} \log n$.

We will combine Theorem 6 with a general version of the graph removal lemma proved by Aroskar and Cummings in [5] to prove a supersaturation theorem for hereditary $\mathcal{L}$-properties (Theorem 7 below), and a version of the hypergraph containers theorem for hereditary $\mathcal{L}$-properties (Theorem 8 below).

Theorem 7 (Supersaturation)**.** *Suppose $\mathcal{H}$ is a non-trivial hereditary $\mathcal{L}$-property and $\mathcal{F}$ is as in Observation 1 so that $\mathcal{H} = \text{Forb}(\mathcal{F})$. Then for all $\delta > 0$ there are $\epsilon > 0$ and K such that for all sufficiently large n, if M is an $\mathcal{L}_{\mathcal{H}}$-template of size n such that $\text{prob}(\tilde{\mathcal{F}}(K) \cup \mathcal{E}(K), M) < \epsilon$, then*

(1) If $\pi(\mathcal{H}) > 1$, then $sub(M) \leq \text{ex}(n, \mathcal{H})^{1+\delta}$.
(2) If $\pi(\mathcal{H}) \leq 1$, then $sub(M) \leq 2^{\delta\binom{n}{r}}$.

Theorem 8. *Suppose $\mathcal{H}$ is a hereditary $\mathcal{L}$-property. Then there is $m = m(\mathcal{H}, r_{\mathcal{L}}) \geq 1$ such that the following holds. For every $\delta > 0$ there is a constant $c = c(\mathcal{H}, \mathcal{L}, \delta)$ such that for all sufficiently large n there is a set of $\mathcal{L}_{\mathcal{H}}$-templates $\mathcal{C}$ with domain $[n]$ satisfying the following properties.*

(1) For every $H \in \mathcal{H}_n$, there is $C \in \mathcal{C}$ such that $H \trianglelefteq_p C$.
(2) For every $C \in \mathcal{C}$, there is $C' \in \mathcal{R}([n], \mathcal{H})$ such that $dist(C, C') \leq \delta$.
(3) $\log |\mathcal{C}| \leq cn^{r-\frac{1}{m}} \log n$.

5. Proofs of Main Theorems

In this section we prove our main results using Theorems 6, 7, and 8. For the rest of the section, $\mathcal{H}$ is a fixed hereditary $\mathcal{L}$-property.

Lemma 1. *Suppose N is an $\mathcal{L}$-structure and $\tilde{N}$ is the $\mathcal{L}_{\mathcal{H}}$-structure such that $dom(\tilde{N}) = dom(N)$ and for each $\bar{a} \in dom(\tilde{N})^r$ and $p(\bar{x}) \in S_r(\mathcal{H})$, $\tilde{N} \models R_p(\bar{a})$ if and only if $N \models p(\bar{a})$. Then $\tilde{N}$ is an $\mathcal{L}_{\mathcal{H}}$-template and N is the unique full subpattern of $\tilde{N}$.*

Proof. Let $V = dom(N) = dom(\tilde{N})$. We leave it to the reader to verify that $\tilde{N}$ is an $\mathcal{L}_{\mathcal{H}}$-template (this is straightforward from the definitions). Define $\chi : \binom{V}{r} \rightarrow S_r(C_V, \mathcal{H})$ by $\chi(A) = Diag^N(A)$ for each $A \in \binom{V}{r}$. It is clear that $\chi \in Ch(\tilde{N})$ and $N \trianglelefteq_{\chi} \tilde{N}$. By definition of $\tilde{N}$, χ is the *only* choice function for $\tilde{N}$, so any full subpattern of $\tilde{N}$ must be chosen by χ. By Observation 3, χ chooses at most one $\mathcal{L}$-structure, so N is the unique full subpattern of $\tilde{N}$. $\square$

We now prove Theorem 3. The proof is based on the method of proof in [18].

Proof of Theorem 2. Let $b_n = ex(n, \mathcal{H})^{1/\binom{n}{r}}$. If $\mathcal{H}$ is trivial, then for sufficiently large n, $\mathcal{R}(n, \mathcal{H}) = \emptyset$ so by convention, $\mathrm{ex}(n, \mathcal{H}) = 0$. Thus, for sufficiently large n, $b_n = 0$ and $\pi(\mathcal{H})$ exists and is equal to zero.

Assume now $\mathcal{H}$ is nontrivial. We show that the sequence b_n is bounded below and non-increasing. Since $\mathcal{H}$ is non-trivial and has the hereditary property, $\mathcal{H}_n \neq \emptyset$ for all n. Fix $n \geq 1$ and choose any $N \in \mathcal{H}_n$. Let $\tilde{N}$ be the $\mathcal{L}_{\mathcal{H}}$-structure defined as in Lemma 1 for N. Then $\tilde{N}$ is an $\mathcal{L}_{\mathcal{H}}$-template, and its only full subpattern is N. Since $N \in \mathcal{H}$, this implies $\tilde{N} \in \mathcal{R}(n, \mathcal{H})$ and $sub(\tilde{N}) = 1$. So we have shown $b_n \geq 1$ for all $n \geq 1$.

We now show the b_n are eventually non-increasing. Fix $n \geq r + 1$. Let $M \in \mathcal{R}(n, \mathcal{H})$ be such that $sub(M) \geq 1$ and let $V = dom(M)$. Fix $a \in V$ and set $V_a = V \setminus \{a\}$ and $M_a = M[V_a]$. We claim $M_a \in \mathcal{R}(n-1, \mathcal{H})$. Because M is an $\mathcal{L}_{\mathcal{H}}$-template, the definition of M_a implies M_a is also an $\mathcal{L}_{\mathcal{H}}$-template. Suppose $\chi \in Ch(M_a)$. We want to show there exists $N_a \in \mathcal{H}$ with $N_a \trianglelefteq_{\chi} M_a$. We define a function $\chi' : \binom{V}{r} \rightarrow S_r(C_V, \mathcal{H})$ as follows. For $A \in \binom{V_a}{r}$, set $\chi'(A) = \chi(A)$, and for $A \in \binom{V}{r} \setminus A \in \binom{V_a}{r}$, choose $\chi'(A)$ to be any element of $Ch_{M_a}(A) = Ch_M(A)$ (this is possible since M is complete). Note that for each $A \in \binom{V_a}{r}$, $\chi(A) \in Ch_M(A)$, so $\chi' \in Ch(M)$. Because M is $\mathcal{H}$-random, there is $N \in \mathcal{H}$ such that $N \trianglelefteq_{\chi'} M$. Let $N_a = N[V_a]$. Because $\mathcal{H}$ has the hereditary property and $N_a \subseteq_{\mathcal{L}} N$, $N_a \in \mathcal{H}$. For each $A \in \binom{V_a}{r}$, $Diag^{N_a}(A) = Diag^N(A) = \chi'(A) = \chi(A)$, so $N_a \trianglelefteq_{\chi} M_a$. Thus we have verified that $M_a \in \mathcal{R}(n-1, \mathcal{H})$. By definition of b_{n-1}, this implies $sub(M_a)^{1/\binom{n-1}{r}} \leq b_{n-1}$. Because M_a is $\mathcal{H}$-random, Proposition 2 implies it is error-fre. Therefore Observation 4 implies $sub(M_a) = \prod_{A \in \binom{V_a}{r}} |Ch_{M_a}(A)|$. Then observe that

$$sub(M) = \Big(\prod_{a \in V} \prod_{A \in \binom{V_a}{r}} |Ch_{M_a}(A)| \Big)^{1/(n-r)} = \Big(\prod_{a \in V} sub(M_a) \Big)^{1/(n-r)}.$$

Since $sub(M_a) \leq b_{n-1}^{\binom{n-1}{r}}$, this implies

$$sub(M) \leq \Big(\prod_{a \in V} b_{n-1}^{\binom{n-1}{r}} \Big)^{1/(n-r)} = b_{n-1}^{n\binom{n-1}{r}/(n-r)} = b_{n-1}^{\binom{n}{r}}.$$

Thus for all $M \in \mathcal{R}(n, \mathcal{H})$, $sub(M)^{1/\binom{n}{r}} \leq b_{n-1}$. So by definition, $b_n \leq b_{n-1}$. $\square$

The following observations follow from the proof of Theorem 2.

Ovservation 5. *Assume $\mathcal{H}$ is a hereditary $\mathcal{L}$-property.*

(a) For all n, $\mathrm{ex}(n, \mathcal{H})^{1/\binom{n}{r}} \geq \pi(\mathcal{H})$ (since $(b_n)_{n \in \mathbb{N}}$ is non-increasing and converges to $\pi(\mathcal{H})$).
(b) Either $\mathcal{H}$ is trivial and $\pi(\mathcal{H}) = 0$ or $\mathcal{H}$ is non-trivial and $\pi(\mathcal{H}) \geq 1$.

Proof of Theorem 3. Assume $\mathcal{H}$ is a hereditary $\mathcal{L}$-property. Recall we want to show the following.

(1) If $\pi(\mathcal{H})>1$, then $|\mathcal{H}_n|=\pi(\mathcal{H})^{\binom{n}{r}+o(n^r)}$.
(2) If $\pi(\mathcal{H})\leq 1$, then $|\mathcal{H}_n|=2^{o(n^r)}$.

Assume first that $\mathcal{H}$ is trivial. Then by Observation 5(b), $\pi(\mathcal{H})=0\leq 1$, so we are in case 2. Since $|\mathcal{H}_n|=0$ for all sufficiently large n, $|\mathcal{H}_n|=2^{o(n^r)}$ holds, as desired. Assume now $\mathcal{H}$ is non-trivial, so $\pi(\mathcal{H})\geq 1$ by Observation 5(b). We show that for all $0<\eta<1$, either $\pi(\mathcal{H})=1$ and $|\mathcal{H}_n|\leq 2^{\eta n^r}$ or $\pi(\mathcal{H})>1$ and $\pi(\mathcal{H})^{\binom{n}{r}}\leq|\mathcal{H}_n|\leq\pi(\mathcal{H})^{\binom{n}{r}+\eta n^r}$. Fix $0<\eta<1$. Let $\mathcal{F}$ be as in Observation 1 for $\mathcal{H}$ so that $\mathcal{H}=\mathrm{Forb}(\mathcal{F})$. Choose $\epsilon>0$ and K as in Theorem 7 for $\delta=\eta/4$. Replacing K if necessary, assume $K\geq r$. Apply Theorem 6 to ϵ and $\mathcal{F}(K)$ to obtain $m=m(K,r)\geq 1$ and $c=c(K,r,\mathcal{L},\epsilon)$. Assume n is sufficiently large. Theorem 6 with $W=[n]$ and $\mathcal{B}:=\mathrm{Forb}(\mathcal{F}(K))$ implies there is a collection $\mathcal{C}$ of $\mathcal{L}_{\mathcal{B}}$-templates with domain $[n]$ such that the following hold.

(i) For all $\mathcal{F}(K)$-free $\mathcal{L}$-structures M with domain $[n]$, there is $C\in\mathcal{C}$ such that $M\trianglelefteq_p C$,
(ii) For all $C\in\mathcal{C}$, $\mathrm{prob}(\widetilde{\mathcal{F}(K)},C)\leq\epsilon$ and $\mathrm{prob}(\mathcal{E},C)\leq\epsilon$.
(iii) $\log|\mathcal{C}|\leq cn^{r-\frac{1}{m}}\log n$.

Note that because $K\geq r$, $\mathcal{H}=\mathrm{Forb}(\mathcal{F})$ and $\mathcal{B}=\mathrm{Forb}(\mathcal{F}(K))$ imply we must have $S_r(\mathcal{H})=S_r(\mathcal{B})$. Consequently all $\mathcal{L}_{\mathcal{B}}$-templates are also $\mathcal{L}_{\mathcal{H}}$-templates. In particular the elements in $\mathcal{C}$ are all $\mathcal{L}_{\mathcal{H}}$-templates. Therefore, (ii) and Theorem 7 imply that for all $C\in\mathcal{C}$, either $sub(C)\leq\mathrm{ex}(n,\mathcal{H})^{1+\eta/4}$ (case $\pi(\mathcal{H})>1$) or $sub(C)\leq 2^{\eta\binom{n}{r}/4}$ (case $\pi(\mathcal{H})=1$). Note every element in $\mathcal{H}_n$ is $\mathcal{F}$-free, so is also $\mathcal{F}(K)$-free. This implies by (i) that every element of $\mathcal{H}_n$ is a full subpattern of some $C\in\mathcal{C}$. Therefore we can construct every element in $\mathcal{H}_n$ as follows.

- Choose a $C\in\mathcal{C}$. There are at most $|\mathcal{C}|\leq 2^{cn^{r-\frac{1}{m}}\log n}$ choices.
- Choose a full subpattern of C. There are at most $sub(C)\leq\mathrm{ex}(n,\mathcal{H})^{1+\eta/4}$ choices if $\pi(\mathcal{H})>1$ and at most $sub(C)\leq 2^{\eta\binom{n}{r}/4}$ choices if $\pi(\mathcal{H})=1$.

This implies

$$|\mathcal{H}_n|\leq\begin{cases}2^{cn^{r-\frac{1}{m}}\log n}\mathrm{ex}(n,\mathcal{H})^{1+\eta/4} & \text{if } \pi(\mathcal{H})>1\\ 2^{cn^{r-\frac{1}{m}}\log n}2^{\eta\binom{n}{r}/4} & \text{if } \pi(\mathcal{H})=1.\end{cases}\tag{3}$$

If $\pi(\mathcal{H})>1$, then we may assume n is sufficiently large so that $\mathrm{ex}(n,\mathcal{H})\leq\pi(\mathcal{H})^{(1+\eta/4)\binom{n}{r}}$ (see Observation 5(a)). Combining this with (3), we have that when $\pi(\mathcal{H})>1$,

$$|\mathcal{H}_n|\leq 2^{cn^{r-\frac{1}{m}}\log n}\pi(\mathcal{H})^{(1+\eta/4)^2\binom{n}{r}}\leq\pi(\mathcal{H})^{\binom{n}{r}+\eta\binom{n}{r}},$$

where the last inequality is because $\pi(\mathcal{H})>1$, $(1+\eta/4)^2<1+\eta$, and n is sufficiently large. If $\pi(\mathcal{H})=1$, then (3) implies

$$|\mathcal{H}_n|\leq 2^{cn^{r-\frac{1}{m}}\log n}2^{\eta\binom{n}{r}/4}\leq 2^{\eta\binom{n}{r}},$$

where the last inequality is because n is sufficiently large. Thus, we have shown $|\mathcal{H}_n|\leq 2^{\eta n^r}$ when $\pi(\mathcal{H})=1$ and $|\mathcal{H}_n|\leq\pi(\mathcal{H})^{\binom{n}{r}+\eta n^r}$ when $\pi(\mathcal{H})>1$. We have left to show that when $\pi(\mathcal{H})>1$, then $|\mathcal{H}_n|\geq\pi(\mathcal{H})^{\binom{n}{r}}$. This holds because for any $M\in\mathcal{R}_{ex}([n],\mathcal{H})$, all $\mathrm{ex}(n,\mathcal{H})$ many full subpatterns of M are in $\mathcal{H}_n$. Thus $|\mathcal{H}_n|\geq\mathrm{ex}(n,\mathcal{H})\geq\pi(\mathcal{H})^{\binom{n}{r}}$, where the second inequality is by Observation 5(a). This finishes the proof. □

We now prove two lemmas needed for Theorems 4 and 5.

Lemma 2. *Suppose $\mathcal{H}$ is a non-trivial hereditary $\mathcal{L}$-property. Then there is $\gamma=\gamma(\mathcal{H})>0$ such that for all $\delta>0$ and $n\geq r$, if C and C' are $\mathcal{L}_{\mathcal{H}}$-templates with domain $[n]$ such that C' is error-free and $\mathrm{dist}(C,C')\leq\delta$, then the following holds.*

(1) If $\pi(\mathcal{H})>1$, then $sub(C)\leq sub(C')\mathrm{ex}(n,\mathcal{H})^{\gamma\delta}$.

(2) If $\pi(\mathcal{H}) = 1$*, then* $sub(C) \leq sub(C')2^{\gamma\delta\binom{n}{r}}$.

Proof. Fix $n \geq r$ and assume C and C' are $\mathcal{L}_{\mathcal{H}}$-templates with domain $[n]$ such that C' is error-free and $\text{dist}(C, C') \leq \delta$. Then by definition of $\text{dist}(C, C')$, $|\text{diff}(C, C')| \leq \delta\binom{n}{r}$. By Proposition 1,

$$\text{diff}(C, C') = \{A \in \binom{V}{r} : Ch_C(A) \neq Ch_{C'}(A)\}. \tag{4}$$

Note that for every $A \in \binom{V}{r}$, $|Ch_C(A)| \leq |S_r(\mathcal{H})|$ (by definition of $Ch_C(A)$) and $1 \leq |Ch_{C'}(A)|$ (since C' is complete). Thus $\frac{|Ch_C(A)|}{|Ch_{C'}(A)|} \leq |S_r(\mathcal{H})|$. By Observation 4 and (4),

$$\begin{aligned} sub(C) \leq \prod_{A \in \binom{V}{r}} |Ch_C(A)| &= \Big(\prod_{A \notin \text{diff}(C,C')} |Ch_{C'}(A)|\Big)\Big(\prod_{A \in \text{diff}(C,C')} |Ch_C(A)|\Big) \\ &= \Big(\prod_{A \in \binom{V}{r}} |Ch_{C'}(A)|\Big)\Big(\prod_{A \in \text{diff}(C,C')} \frac{|Ch_C(A)|}{|Ch_{C'}(A)|}\Big). \end{aligned}$$

Combining this with $\frac{|Ch_C(A)|}{|Ch_{C'}(A)|} \leq |S_r(\mathcal{H})|$ and $|\text{diff}(C, C')| \leq \delta\binom{n}{r}$ yields

$$sub(C) \leq \Big(\prod_{A \in \binom{V}{r}} |Ch_{C'}(A)|\Big)|S_r(\mathcal{H})|^{\delta\binom{n}{r}} = sub(C')|S_r(\mathcal{H})|^{\delta\binom{n}{r}}, \tag{5}$$

where the equality is by Observation 4 and because C' is error-free. If $\pi(\mathcal{H}) > 1$, choose $\gamma > 0$ such that $|S_r(\mathcal{H})| = \pi(\mathcal{H})^{\gamma}$ (this is possible since $\pi(\mathcal{H}) > 1$ implies $S_r(\mathcal{H}) > 1$). Recall from Observation 5(a) that for all n, $\text{ex}(n, \mathcal{H}) \geq \pi(\mathcal{H})^{\binom{n}{r}}$. Combining this with our choice of γ and (5), we have

$$sub(C) \leq sub(C')|S_r(\mathcal{H})|^{\delta\binom{n}{r}} = sub(C')\pi(\mathcal{H})^{\gamma\delta\binom{n}{r}} \leq sub(C')\text{ex}(n, \mathcal{H})^{\gamma\delta}.$$

If $\pi(\mathcal{H}) = 1$, choose $\gamma > 0$ such that $|S_r(\mathcal{H})| \leq 2^{\gamma}$ (this is possible since $\mathcal{H}$ nontrivial implies $|S_r(\mathcal{H})| \geq 1$). Combining our choice of γ with (5) implies

$$sub(C) \leq sub(C')|S_r(\mathcal{H})|^{\delta\binom{n}{r}} \leq sub(C')2^{\gamma\delta\binom{n}{r}}.$$

□

Lemma 3. *Suppose* C *is an* $\mathcal{L}_{\mathcal{H}}$*-template with domain* W *of size* $n \geq r$ *and* $G \trianglelefteq_p C$*. If* $D \in \mathcal{R}(W, \mathcal{H})$ *is such that* $\text{dist}(C, D) \leq \delta$*, then there is* $G' \in \mathcal{H}$ *such that* $G' \trianglelefteq_p D$ *and* $\text{dist}(G, G') \leq \delta$*.*

Proof. Fix C and D satisfying the hypotheses Because $\text{dist}(C, D) \leq \delta$, we have $|\text{diff}(C, D)| \leq \delta\binom{n}{r}$. By Proposition 1,

$$\text{diff}(C, D) = \{A \in \binom{W}{r} : Ch_C(A) \neq Ch_D(A)\}. \tag{6}$$

Define a function $\chi : \binom{W}{r} \to S_r(C_W)$ as follows. For $A \in \binom{W}{r} \setminus \text{diff}(C, D)$, set $\chi(A) = Diag^G(A)$. For each $A \in \text{diff}(C, D)$, choose $\chi(A)$ to be any element of $Ch_D(A)$ (which is nonempty because D is an $\mathcal{L}_{\mathcal{H}}$-template). Since $G \trianglelefteq_p C$, for all $A \in \binom{W}{r}$, $Diag^G(A) \in Ch_C(A)$. Thus, by definition of χ and (6), for all $A \in \binom{W}{r} \setminus \text{diff}(C, D)$, $\chi(A) = Diag^G(A) \in Ch_C(A) = Ch_D(A)$. For $A \in \text{diff}(C, D)$, $\chi(A) \in Ch_D(A)$ by assumption. Thus $\chi \in Ch(D)$. Because D is $\mathcal{H}$-random, there is $G' \in \mathcal{H}$ such that $G' \trianglelefteq_{\chi} D$. We show $\text{dist}(G, G') \leq \delta$. By definition of χ and since $G' \trianglelefteq_{\chi} D$, we have that for all $A \in \binom{W}{r}$, if $A \notin \text{diff}(C, D)$, then $Diag^{G'}(A) = \chi(A) = Diag^G(A)$, which implies $A \notin \text{diff}(G, G')$. Thus $\text{diff}(G, G') \subseteq \text{diff}(C, D)$ so $|\text{diff}(G, G')| \leq \delta\binom{n}{r}$ and $\text{dist}(G, G') \leq \delta$ by definition. □

Proof of Theorem 4. Let $\mathcal{H}$ be a fast-growing hereditary $\mathcal{L}$-property. Fix ϵ and $\delta > 0$. Given n, let $A(n, \epsilon, \delta) = \mathcal{H}_n \setminus E^{\delta}(\epsilon, n, \mathcal{H})$. Recall, we want to show there is $\beta > 0$ such that for sufficiently large n,

$$\frac{|A(n,\epsilon,\delta)|}{|\mathcal{H}_n|} \leq 2^{-\beta\binom{n}{r}}. \tag{7}$$

Let $\gamma > 0$ be as in Lemma 2 for $\mathcal{H}$. Choose $K > 2r$ sufficiently large so that $1-\epsilon+\gamma\delta/K < 1-\epsilon/2$. Apply Theorem 8 to $\frac{\delta}{K}$ to obtain constants c and $m \geq 1$. Assume n is sufficiently large. Then Theorem 8 implies there is a collection $\mathcal{C}$ of $\mathcal{L}_{\mathcal{H}}$-templates with domain $[n]$ such that the following hold.

(i) For every $H \in \mathcal{H}_n$, there is $C \in \mathcal{C}$ such that $H \trianglelefteq_p C$.
(ii) For every $C \in \mathcal{C}$, there is $C' \in \mathcal{R}([n], \mathcal{H})$ such that $dist(C, C') \leq \delta$.
(iii) $\log|\mathcal{C}| \leq cn^{r-\frac{1}{m}}\log n$.

Suppose $G \in A(n,\epsilon,\delta)$. By (i), there is $C \in \mathcal{C}_n$ such that $G \trianglelefteq_p C$. By (ii), there is $M_C \in \mathcal{R}([n], \mathcal{H})$ such that $\text{dist}(C, M_C) \leq \frac{\delta}{K}$. By Lemma 3, there is $G' \trianglelefteq_p M_C$ with $\text{dist}(G, G') \leq \frac{\delta}{K} \leq \delta$. If $sub(M_C) \geq \text{ex}(n,\mathcal{H})^{1-\epsilon}$, then by definition of $E^{\delta}(\epsilon, n, \mathcal{H})$, $\text{dist}(G, G') \leq \delta$ and $G' \trianglelefteq_p M_C$ would imply $G \in E^{\delta}(\epsilon, n, \mathcal{H})$, contradicting our assumption that $G \in A(n,\epsilon,\delta) = \mathcal{H}_n \setminus E^{\delta}(\epsilon, n, \mathcal{H})$. Therefore, we must have $sub(M_C) < \text{ex}(n,\mathcal{H})^{1-\epsilon}$. Note $M_C \in \mathcal{R}([n], \mathcal{H})$ implies M_C is error-free, so Lemma 2 and the fact that $\text{dist}(C, M_C) \leq \delta/K$ imply $sub(C) \leq sub(M_C)\text{ex}(n,\mathcal{H})^{\gamma\delta/K}$. Combining this with the fact that $sub(M_C) < \text{ex}(n,\mathcal{H})^{1-\epsilon}$ we have that

$$sub(C) < \text{ex}(n,\mathcal{H})^{1-\epsilon}\text{ex}(n,\mathcal{H})^{\gamma\delta/K} = \text{ex}(n,\mathcal{H})^{1-\epsilon+\gamma\delta/K} \leq \text{ex}(n,\mathcal{H})^{1-\epsilon/2},$$

where the second inequality is by assumption on K. Therefore every $G \in A(n,\epsilon,\delta)$ can be constructed as follows.

- Choose $C \in \mathcal{C}_n$ with $sub(C) < \text{ex}(n,\mathcal{H})^{1-\epsilon/2}$. There are at most $|\mathcal{C}_n| \leq 2^{cn^{r-\frac{1}{m}}\log n}$ ways to do this, where the bound is by (iii). Since n is large and $\pi(\mathcal{H}) > 1$, we may assume $2^{cn^{r-\frac{1}{m}}\log n} \leq \pi(\mathcal{H})^{\epsilon\binom{n}{r}/4}$.
- Choose a full subpattern of C. There are at most $sub(C) < \text{ex}(n,\mathcal{H})^{1-\epsilon/2}$ ways to do this.

Combining these bounds yields $|A(n,\epsilon,\delta)| \leq \pi(\mathcal{H})^{\epsilon\binom{n}{r}/4}\text{ex}(n,\mathcal{H})^{1-\epsilon/2}$. Recall that $|\mathcal{H}_n| \geq \text{ex}(n,\mathcal{H})$ holds, since for any $M \in \mathcal{R}_{ex}([n],\mathcal{H})$, all $\text{ex}(n,\mathcal{H})$-many full subpatterns of M are all in $\mathcal{H}_n$. Therefore

$$\frac{|A(n,\epsilon,\delta)|}{|\mathcal{H}_n|} \leq \frac{\pi(\mathcal{H})^{\epsilon\binom{n}{r}/4}\text{ex}(n,\mathcal{H})^{1-\epsilon/2}}{\text{ex}(n,\mathcal{H})} = \pi(\mathcal{H})^{\epsilon\binom{n}{r}/4}\text{ex}(n,\mathcal{H})^{-\epsilon/2} \leq \pi(\mathcal{H})^{-\epsilon\binom{n}{r}/4}, \tag{8}$$

where the last inequality is because $\pi(\mathcal{H})^{\binom{n}{r}} \leq \text{ex}(n,\mathcal{H})$. Therefore we have $\frac{|A(n,\epsilon,\delta)|}{|\mathcal{H}_n|} \leq 2^{-\beta\binom{n}{r}}$, where $\beta = \frac{\epsilon\log\pi(\mathcal{H})}{4\log 2}$. Note $\beta > 0$ since $\pi(\mathcal{H}) > 1$. □

Proof of Theorem 5. Suppose $\mathcal{H}$ is a fast growing hereditary $\mathcal{L}$-property with a stability theorem. Fix $\delta > 0$. Recall we want to show there is $\beta > 0$ such that for sufficiently large n,

$$\frac{|\mathcal{H}_n \setminus E^{\delta}(n,\mathcal{H})|}{|\mathcal{H}_n|} \leq 2^{-\beta\binom{n}{r}}$$

By Theorem 4, it suffices to show that there are $\epsilon_1, \delta_1 > 0$ such that for all sufficiently large n, $E^{\delta_1}(\epsilon_1, n, \mathcal{H}) \subseteq E^{\delta}(n,\mathcal{H})$.

Since $\mathcal{H}$ has a stability theorem, there is ϵ such that for all sufficiently large n, if $H \in \mathcal{R}([n],\mathcal{H})$ satisfies $sub(H) \geq \text{ex}(n,\mathcal{H})^{1-\epsilon}$, then there is $H' \in \mathcal{R}_{ex}([n],\mathcal{H})$ with $\text{dist}(H,H') \leq \frac{\delta}{2}$. Fix n sufficiently large. We claim $E^{\delta/2}(\epsilon, n, \mathcal{H}) \subseteq E^{\delta}(n,\mathcal{H})$. Suppose $G \in E^{\delta/2}(\epsilon, n, \mathcal{H})$. Then by definition, G is $\delta/2$-close to some G' such that $G' \trianglelefteq_p H$, for some $H \in \mathcal{R}([n],\mathcal{H})$ which satisfies

$sub(H) \geq \mathrm{ex}(n,\mathcal{H})^{1-\epsilon}$. By choice of ϵ and because n is sufficiently large, there is $H' \in \mathcal{R}_{ex}([n],\mathcal{H})$ such that $\mathrm{dist}(H,H') \leq \frac{\delta}{2}$. Lemma 3 implies there is some $G'' \trianglelefteq_p H'$ such that $\mathrm{dist}(G',G'') \leq \frac{\delta}{2}$. Observe that $G'' \in E(n,\mathcal{H})$ and

$$\mathrm{dist}(G,G'') \leq \mathrm{dist}(G,G') + \mathrm{dist}(G',G'') \leq \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

This implies that $G \in E^{\delta}(n,\mathcal{H})$, as desired. □

6. Characterization of $\mathcal{H}$-random $\mathcal{L}_{\mathcal{H}}$-templates

In this section we give an equivalent characterization for when an $\mathcal{L}_{\mathcal{H}}$-structure is an $\mathcal{H}$-random $\mathcal{L}_{\mathcal{H}}$-template, where $\mathcal{H}$ is a hereditary $\mathcal{L}$-property. The results in this section will be used in the proofs of our remaining results, Theorems 6, 7, and 8. For the rest of this section, $\mathcal{H}$ is a fixed nonempty collection of finite $\mathcal{L}$-structures.

Definition 21. Define FLAW to be the class of all $\mathcal{L}_{\mathcal{H}}$-structures of size r which are not $\mathcal{L}_{\mathcal{H}}$-templates. Elements of FLAW are called flaws.

Lemma 4. *An $\mathcal{L}_{\mathcal{H}}$-structure M is an $\mathcal{L}_{\mathcal{H}}$-template if and only if it is* FLAW*-free.*

Proof. Let $dom(M) = V$. It is straightforward from Definition 9 that M is an $\mathcal{L}_{\mathcal{H}}$-template if and only if for all $A \in \binom{V}{r}$, $M[A]$ is an $\mathcal{L}_{\mathcal{H}}$-template. By definition of FLAW, M is FLAW-free if and only if for all $A \in \binom{V}{r}$, $M[A]$ is an $\mathcal{L}_{\mathcal{H}}$-template. This finishes the proof. □

We are now ready to prove the main result of this section.

Proposition 3. *Suppose $\mathcal{H}$ is a hereditary $\mathcal{L}$-property, and $\mathcal{F}$ is the class of finite $\mathcal{L}$-structures from Observation 1 such that* $\mathrm{Forb}(\mathcal{F}) = \mathcal{H}$. *Then a complete $\mathcal{L}_{\mathcal{H}}$-structure M is $\mathcal{H}$-random if and only if M is $\tilde{\mathcal{F}}$-free and error-free.*

Proof. By Observation 1, $\mathcal{F}$ is closed under isomorphism. Fix a complete $\mathcal{L}_{\mathcal{H}}$-structure M and let $V = dom(M)$. Suppose first that M is $\mathcal{H}$-random. Then M is complete and for every $\chi \in Ch(M)$, there is $N \in \mathcal{H}$ such that $N \trianglelefteq_{\chi} M$. This implies by Proposition 2 that M is error-free. Suppose by contradiction M is not $\tilde{\mathcal{F}}$-free. Combining the assumption that $\mathcal{F}$ is closed under isomorphism and the definition of $\tilde{\mathcal{F}}$, this implies there is $B \subseteq V$ and $F \in \mathcal{F}$ such that $M[B] \in \tilde{F}$. By definition of $\tilde{F}$, there is $\chi_B \in Ch(M[B])$ such that $F \trianglelefteq_{\chi_B} M[B]$. Define a function $\chi : \binom{V}{r} \rightarrow S_r(C_V,\mathcal{H})$ as follows. For each $A \in \binom{B}{r}$, set $\chi(A) = \chi_B(A)$. Clearly, $\chi_B \in Ch(M[B])$ implies that for all $A \in \binom{B}{r}$, $\chi_B(A) \in Ch_M(A)$. For each $A \in \binom{V}{r} \setminus \binom{B}{r}$, define $\chi(A)$ to be any element of $Ch_M(A)$ (this is possible since M is complete by assumption). By construction, $\chi \in Ch(M)$. Because M is $\mathcal{H}$-random, there is $D \in \mathcal{H}$ such that $D \trianglelefteq_{\chi} M$. By choice of $\mathcal{F}$, since $D \in \mathcal{H}$, we have that D is $\mathcal{F}$-free, which implies D is F-free since $F \in \mathcal{F}$. We claim $D[B] \cong_{\mathcal{L}} F$, a contradiction. For each $A \in \binom{B}{r}$, $D \trianglelefteq_{\chi} M$, $F \trianglelefteq_{\chi_B} M[B]$, and the definition of χ imply

$$Diag^D(A) = \chi(A) = \chi_B(A) = Diag^F(A).$$

Thus $Diag(D[B]) = \bigcup_{A\in\binom{B}{r}} Diag^D(A) = \bigcup_{A\in\binom{B}{r}} Diag^F(A) = Diag(F)$ implies $D[B] \cong_{\mathcal{L}} F$.

For the converse, suppose M is a complete $\mathcal{L}_{\mathcal{H}}$-structure which is $\tilde{\mathcal{F}}$-free and error-free. Suppose by contradiction M is not $\mathcal{H}$-random. Then there is $\chi \in Ch(M)$ such that there is no $N \in \mathcal{H}$ with $N \trianglelefteq_{\chi} M$. Since M is error-free, Proposition 2 implies there is some $\mathcal{L}$-structure N such that $N \trianglelefteq_{\chi} M$. Thus we must have $N \notin \mathcal{H}$. By choice of $\mathcal{F}$ from Observation 1, N is not $\mathcal{F}$-free. This along with the fact that $\mathcal{F}$ is closed under isomorphism implies there is $B \subseteq V$ such that $N[B] \in \mathcal{F}$. But $N \trianglelefteq_p M$ implies $N[B] \trianglelefteq_p M[B]$ (this is straightforward to check). Since $N[B] \in \mathcal{F}$, this implies $M[B] \in \tilde{\mathcal{F}}$ by definition of $\tilde{\mathcal{F}}$, contradicting that M is $\tilde{\mathcal{F}}$-free. □

Corollary 1. *Suppose $\mathcal{H}$ is a hereditary $\mathcal{L}$-property, and $\mathcal{F}$ is the class of finite $\mathcal{L}$-structures from Observation 1 such that* $\mathrm{Forb}(\mathcal{F}) = \mathcal{H}$. *Let M be an $\mathcal{L}_{\mathcal{H}}$-structure. Then $M \in \mathcal{R}(dom(M), \mathcal{H})$ if and only if M is $\tilde{\mathcal{F}}$-free, error-free, and* FLAW*-free.*

Proof. By definition, $M \in \mathcal{R}(dom(M), \mathcal{H})$ if and only if M is an $\mathcal{H}$-random $\mathcal{L}_{\mathcal{H}}$-template. By Lemma 4 and Proposition 3, this holds if and only if M is $\tilde{\mathcal{F}}$-free, error-free, and $FLAW$-free. □

7. Graph Removal and Proofs of Theorems 7 and 8.

In this section we will use a generalized version of the graph removal lemma to prove Theorem 7 and to prove Theorem 8 from Theorem 6. We begin by stating this generalized version of the graph removal lemma, Theorem 9, which is a reformulation of Theorem 2 in [5]. We refer the reader to Appendix A for the proof of Theorem 9 from the original version in [5].

Theorem 9 (Aroskar-Cummings). *Suppose $\mathcal{L}_0$ is a fixed finite relational language with $r_{\mathcal{L}_0} = r$. Suppose $\mathcal{A}$ is a collection of finite $\mathcal{L}_0$-structures. For every $\delta > 0$ there exists $\epsilon > 0$ and K such that for all sufficiently large finite $\mathcal{L}_0$-structures M, if* $\mathrm{prob}(\mathcal{A}(K), M) < \epsilon$, *then there is an $\mathcal{L}_0$-structure M' with $dom(M') = dom(M)$ such that* $\mathrm{dist}(M', M) < \delta$ *and* $\mathrm{prob}(\mathcal{A}, M') = 0$.

Proof of Theorem 7. Let $\mathcal{H}$ be a nontrivial hereditary $\mathcal{L}$-property and let $\mathcal{F}$ be as in Observation 1 so that $\mathcal{H} = \mathrm{Forb}(F)$. Recall we want to show that for all $\delta > 0$, there are $\epsilon > 0$ and K such that for sufficiently large n, for any $\mathcal{L}_{\mathcal{H}}$-template M of size n, if $\mathrm{prob}(\tilde{\mathcal{F}}(K) \cup \mathcal{E}(K), M) \leq \epsilon$ then

(1) If $\pi(\mathcal{H}) > 1$, then $sub(M) \leq \mathrm{ex}(n, \mathcal{H})^{1+\delta}$.
(2) If $\pi(\mathcal{H}) \leq 1$, then $sub(M) \leq 2^{\delta\binom{n}{r}}$.

Fix $\delta > 0$. Apply Lemma 2 to $\mathcal{H}$ to obtain $\gamma > 0$. Let $\mathcal{A} = \tilde{\mathcal{F}} \cup \mathcal{E} \cup \mathrm{FLAW}$. Apply Theorem 9 to obtain K and ϵ for $\delta/2\gamma$ and $\mathcal{A}$. Suppose n is sufficiently large and M is an $\mathcal{L}_{\mathcal{H}}$-template of size n satsifying $\mathrm{prob}(\tilde{\mathcal{F}}(K) \cup \mathcal{E}(K), M) < \epsilon$. Because M is an $\mathcal{L}_{\mathcal{H}}$-template, Lemmas 4 implies for all $B \in \mathrm{FLAW}$, $\mathrm{prob}(B, M) = 0$. Therefore $\mathrm{prob}(\mathcal{A}(K), M) < \epsilon$, so by Theorem 9, there is an $\mathcal{L}_{\mathcal{H}}$-structure M' with $dom(M) = dom(M')$ such that $\mathrm{prob}(\mathcal{A}, M') = 0$ and $\mathrm{dist}(M, M') \leq \delta/2\gamma$. Since $\mathrm{prob}(\mathcal{A}, M') = 0$, Corollary 1 implies $M' \in \mathcal{R}(n, \mathcal{H})$. Thus $sub(M') \leq \mathrm{ex}(n, \mathcal{H})$ holds by definition of $\mathrm{ex}(n, \mathcal{H})$. Combining this with Lemma 2 (note $M' \in \mathcal{R}(n, \mathcal{H})$ implies M' is error-free), we have the following.

(1) If $\pi(\mathcal{H}) > 1$, then $sub(M) \leq sub(M')\mathrm{ex}(n, \mathcal{H})^{\gamma(\delta/2\gamma)} = sub(M')\mathrm{ex}(n, \mathcal{H})^{\delta/2} \leq \mathrm{ex}(n, \mathcal{H})^{1+\delta/2}$.
(2) If $\pi(\mathcal{H}) = 1$, then $sub(M) \leq sub(M')2^{\gamma(\delta/2\gamma)\binom{n}{r}} = sub(M')2^{\delta\binom{n}{r}/2} \leq \mathrm{ex}(n, \mathcal{H})2^{\delta/2\binom{n}{r}}$.

We are done in the case where $\pi(\mathcal{H}) > 1$. If $\pi(\mathcal{H}) = 1$, assume n is sufficiently large so that $\mathrm{ex}(n, \mathcal{H}) \leq 2^{\delta/2\binom{n}{r}}$. Then (2) implies $sub(M) \leq 2^{\delta\binom{n}{r}}$, as desired. □

Proof of Theorem 8 from Theorem 6. Suppose $\mathcal{H}$ is a hereditary $\mathcal{L}$-property. Let $\mathcal{F}$ be the class of finite $\mathcal{L}$-structures from Observation 1 so that $\mathcal{H} = \mathrm{Forb}(\mathcal{F})$. Then for each n, $\mathcal{H}_n$ is the set of all $\mathcal{F}$-free $\mathcal{L}$-structures with domain $[n]$. Let $\mathcal{A} = \tilde{\mathcal{F}} \cup \mathcal{E} \cup \mathrm{FLAW}$. Fix $\delta > 0$ and choose K and ϵ as in Theorem 9 for δ and the family $\mathcal{A}$. By replacing K if necessary, assume $K \geq r$. Apply Theorem 6 to $\mathcal{B} := \mathcal{F}(K)$ to obtain $c = c(K, r, \mathcal{L}, \epsilon)$, $m = m(K, r)$. Observe the choice of K depended only on $\mathcal{H}$ and $r = r_{\mathcal{L}}$, so $m = m(\mathcal{H}, r_{\mathcal{L}})$. Since $r_{\mathcal{L}}$ depends on $\mathcal{L}$, $c = c(\mathcal{H}, \mathcal{L}, \epsilon)$. Let n be sufficiently large. Then Theorem 6 applied to $W = [n]$ implies there is a collection $\mathcal{C}$ of $\mathcal{L}_{\mathcal{B}}$-templates with domain $[n]$ such that the following hold.

(i) For all $\mathcal{F}(K)$-free $\mathcal{L}$-structures M with domain $[n]$, there is $C \in \mathcal{C}$ such that $M \trianglelefteq_p C$.
(ii) For all $C \in \mathcal{C}$, $\mathrm{prob}(\widetilde{\mathcal{F}(K)}, C) \leq \epsilon$ and $\mathrm{prob}(\mathcal{E}, C) \leq \epsilon$.
(iii) $\log|\mathcal{C}| \leq cn^{r-\frac{1}{m}}\log n$.

We show this $\mathcal{C}$ satisfies the conclusions of Theorem 8 with c, m and δ. Note that because $K \geq r$, $S_r(\mathcal{H}) = S_r(\mathcal{B})$, so all $\mathcal{L}_{\mathcal{B}}$-templates are also $\mathcal{L}_{\mathcal{H}}$-templates. In particular the elements in $\mathcal{C}$ are all

$\mathcal{L}_{\mathcal{H}}$-templates. Clearly (iii) implies part (3) of Theorem 8 holds. For part (1), since any $H \in \mathcal{H}_n$ is $\mathcal{F}$-free, it is also $\mathcal{F}(K)$-free, so (i) implies there is $C \in \mathcal{C}$ such that $H \trianglelefteq_p C$. This shows part (1) of Theorem 8 holds. For part (2), fix $C \in \mathcal{C}$. Since C is an $\mathcal{L}_{\mathcal{H}}$-template, Lemma 4 implies $\text{prob}(G, C) = 0$ for all $G \in \text{FLAW}$. Then (ii) implies that for all $G \in \widetilde{\mathcal{F}(K)} \cup \mathcal{E}$, $\text{prob}(G, C) \leq \epsilon$. Since $\widetilde{\mathcal{F}(K)} = \tilde{\mathcal{F}}(K)$, these facts imply that for all $G \in \mathcal{A}(K)$, $\text{prob}(G, C) \leq \epsilon$. Thus Theorem 9 implies there is an $\mathcal{L}_{\mathcal{H}}$-structure C' with $dom(C) = dom(C') = [n]$ such that $\text{dist}(C, C') \leq \delta$ and $\text{prob}(\mathcal{A}, C') = 0$. Since $\text{prob}(\mathcal{A}, C') = 0$, C' is a FLAW-free, $\tilde{\mathcal{F}}$-free, and error-free $\mathcal{L}_{\mathcal{H}}$-structure with domain $[n]$, so by Corollary 1, $C' \in \mathcal{R}([n], \mathcal{H})$. This finishes the proof. □

8. A Reduction

We have now proved all the results in this paper except Theorem 6. In this section we prove Theorem 6 by reducing it to another result, Theorem 10 (which is proved in Section 9).

8.1. **Preliminaries.**

In this subsection we give preliminaries necessary for the statement of Theorem 10. Many of these notions are similar to definitions from Section 3. However, we will see that our proofs necessitate this more syntactic treatment.

Definition 22. Suppose C is a set of constants and $\sigma \subseteq S_r(C)$.

- $V(\sigma) = \{c \in C : c \text{ appears in some } p(\bar{c}) \in \sigma\}$.
- Given $A \in \binom{V(\sigma)}{r}$, let $Ch_\sigma(A) = \{p(\bar{c}) \in \sigma : \cup \bar{c} = A\}$. Elements of $Ch_\sigma(A)$ are *choices for A*.
- We say σ is *complete* if $Ch_\sigma(A) \neq \emptyset$, for all $A \in \binom{V(\sigma)}{r}$.

Example 8. Let $\mathcal{L}$ and $\mathcal{P}$ be as in Example 1 (i.e. metric spaces with distances in $\{1, 2, 3\}$). Let $W = \{u, v, w\}$ and $\sigma = \{p_1(c_u, c_v), p_2(c_u, c_v), p_2(c_u, c_w)\} \subseteq S_2(C_W, \mathcal{P})$. Then $V(\sigma) = \{c_u, c_v, c_w\}$ and it is easy to check that $Ch_\sigma(c_u c_v) = \{p_1(c_u, c_v), p_2(c_u, c_v)\}$, $Ch_\sigma(c_u c_w) = \{p_2(c_u, c_w)\}$, and $Ch_\sigma(c_v c_w) = \emptyset$. Observe, this σ is not complete.

Definition 23. Suppose C is a set of n constants and $\sigma \subseteq S_r(C)$. Given $m \leq n$, σ is a *syntactic m-diagram* if $|V(\sigma)| = m$ and for all $A \in \binom{V(\sigma)}{r}$, $|Ch_\sigma(A)| = 1$.

We will say $\sigma \subseteq S_r(C)$ is a *syntactic type diagram* if it is a syntactic $|V(\sigma)|$-diagram.

Example 9. Let $\mathcal{L}$ and $\mathcal{P}$ be as in Example 8, and let $W = \{t, u, v, w\}$ be a set of size 4. Set $\sigma' = \{p_1(c_u, c_v), p_2(c_u, c_w), p_3(c_v, c_w)\} \subseteq S_2(C_W, \mathcal{P})$. Then $V(\sigma') = \{c_u, c_v, c_w\}$ and we have that $Ch_\sigma(c_u c_v) = \{p_1(c_u, c_v)\}$, $Ch_\sigma(c_u c_w) = \{p_2(c_u, c_w)\}$, and $Ch_\sigma(c_v c_w) = \{p_3(c_v, c_w)\}$. This shows σ' is a syntactic 3-diagram.

Observe that if σ is a syntactic m-diagram, then by definition, $|V(\sigma)| = m$ and $|\sigma| = \binom{m}{r}$. Given a tuple of constants $\bar{c} = (c_1, \ldots, c_k)$, a first-order language $\mathcal{L}_0$ containing $\{c_1, \ldots, c_k\}$, and an $\mathcal{L}_0$-structure M, let $\bar{c}^M$ denote the tuple $(c_1^M, \ldots, c_k^M) \in dom(M)^k$.

Definition 24. Suppose C is a set of constants and $\sigma \subseteq S_r(C)$.

1. If M is an $\mathcal{L} \cup V(\sigma)$-structure, write $M \models \sigma^M$ if $M \models p(\bar{c}^M)$ for all $p(\bar{c}) \in \sigma$. Call σ *satisfiable* if there exists an $\mathcal{L} \cup V(\sigma)$-structure M such that $M \models \sigma^M$.
2. If M is an $\mathcal{L} \cup C$-structure, the *type-diagram* of M is the set

$$Diag^{tp}(M, C) = \{p(\bar{c}) \in S_r(C) : M \models p(\bar{c}^M)\}.$$

Suppose that M is an $\mathcal{L}$-structure with $dom(M) = W$. The *canonical type-diagram of M* is

$$Diag^{tp}(M) = \{p(c_{\bar{a}}) \in S_r(C_W) : M \models p(\bar{a})\}.$$

In other words, $Diag^{tp}(M) = Diag^{tp}(M, C_W)$ where M is considered with its natural $\mathcal{L} \cup C_W$-structure. Observe that $Diatg^{tp}(M)$ is always a syntactic $|dom(M)|$-diagram. The difference between $Diag^{tp}(M)$ and $Diag(M)$ is that elements of $Diag^{tp}(M)$ are types (with constants plugged in for the variables) while the elements of $Diag(M)$ are formulas (with constants plugged in for the variables). We now make a few observations which will be used in the remainder of the paper.

Ovservation 6. *Suppose M is an $\mathcal{L}$-structure with domain W of size n. Then the following hold.*

1. *Suppose $m \leq n$, $\sigma \subseteq S_r(C_W)$ is a syntactic m-diagram, and N is an $\mathcal{L} \cup V(\sigma)$-structure of size m. Then $N \models \sigma^N$ if and only if $\sigma = Diag^{tp}(N, V(\sigma))$.*
2. *Suppose N is an $\mathcal{L} \cup C_W$-structure of size n and $N \models Diag^{tp}(M)$. Then $M \cong_{\mathcal{L}} N$.*
3. *If $\sigma \subseteq S_r(C_W)$ and $Diag^{tp}(M) \subseteq \sigma$, then σ is complete.*

Proof. (1): Suppose first $\sigma = Diag^{tp}(N, V(\sigma))$. Then by Definition 24, $N \models \sigma^N$. Converesly, suppose $N \models \sigma^N$. By Definition 24, this implies $\sigma \subseteq Diag^{tp}(N, V(\sigma))$. To show the reverse inclusion, suppose $p(c_{\bar{a}}) \in Diag^{tp}(N, V(\sigma))$. By Definition 24, $N \models p(\bar{c}^N)$. Let $A = \cup \bar{c}^N \in \binom{dom(N)}{r}$ (since $p \in S_r(\mathcal{L})$ is proper and $N \models p(\bar{c}^N)$, the coordinates of $\bar{c}^N$ must all be distinct). Since σ is a syntactic m-diagram, $|Ch_\sigma(A)| = 1$, so there is $p'(\bar{x}) \in S_r(\mathcal{L})$ and $\mu \in Perm(r)$ such that $p'(\mu(\bar{c})) \in \sigma$. Since $N \models \sigma^N$, this implies $N \models p'(\mu(\bar{c}^N))$. Clearly $N \models p(\bar{c}^N)$ and $N \models p'(\mu(\bar{c}^N))$ implies $p(\bar{c}^N) = p'(\mu(\bar{c}^N))$. So we have $p(\bar{c}) = p'(\mu(\bar{c})) \in \sigma$ as desired.

(2): Clearly the map $f : W \rightarrow dom(N)$ sending $a \mapsto c_a^N$ is an $\mathcal{L}$-homomorphism of M into N. Since by assumption, M and N both have size n, it must be a bijection, and thus an $\mathcal{L}$-isomorphism.

(3): For each $A \in \binom{C_W}{r}$, $Diag^M(A) \in Ch_\sigma(A)$ implies $Ch_\sigma(A) \neq \emptyset$. □

Definition 25. Suppose $\mathcal{A}$ is a collection of finite $\mathcal{L}$-structures and C is a set of constant symbols.

1. We say $\sigma \subseteq S_r(C)$ is $\mathcal{A}$-satisfiable if $M \models \sigma^M$ and there is an $\mathcal{L} \cup V(\sigma)$-structure M such that $M \restriction_{\mathcal{L}} \in \mathcal{A}$.
2. Define $Diag^{tp}(\mathcal{A}, C) = \{\sigma \subseteq S_r(C) : \sigma$ is a syntactic type diagram which is $\mathcal{A}$-satisfiable$\}$.
3. Given $\sigma \subseteq S_r(C)$, set $Span(\sigma) = \{\sigma' \subseteq \sigma : \sigma'$ is a syntactic type diagram$\}$.

Example 10. Let $\mathcal{L}$ and $\mathcal{P}$ be as in Example 9. Let $C = \{c_1, c_2, c_3\}$ be a set of three constant symbols. Then $\sigma \subseteq S_2(C, \mathcal{P})$ is a syntactic 3-diagram if and only if $\sigma = \{p_i(c_1, c_2), p_j(c_1, c_3), p_k(c_2, c_3)\}$ for some $i, j, k \in [3]$. Clearly such a σ is $\mathcal{P}$-satisfiable if and only if $|i - j| \leq k \leq i + j$, that is, if and only if the numbers i, j, k do not violate the triangle inequality. Thus $Diag^{tp}(\mathcal{P}, C)$ consists of sets of the form $\sigma = \{p_i(c_1, c_2), p_j(c_1, c_3), p_k(c_2, c_3)\}$ where $i, j, k \in [3]$ satisfy $|i - j| \leq k \leq i + j$.

Fix $\sigma = \{p_1(c_1, c_2), p_2(c_1, c_2), p_3(c_1, c_2), p_1(c_2, c_3), p_1(c_1, c_3)\}$. Then $Span(\sigma)$ consists of the following three elements: $\sigma_1 = \{p_1(c_1, c_2), p_1(c_2, c_3), p_1(c_1, c_3)\}$, $\sigma_2 = \{p_2(c_1, c_2), p_1(c_2, c_3), p_1(c_1, c_3)\}$, and $\sigma_3 = \{p_3(c_1, c_2), p_1(c_2, c_3), p_1(c_1, c_3)\}$. Observe that σ_1 and σ_2 are $\mathcal{P}$-satisfiable, while σ_3 is not.

For the rest of this subsection, $\mathcal{H}$ is a fixed collection of finite $\mathcal{L}$-structures.

Lemma 5. *Suppose $X \subseteq W$ are finite sets, M is a complete $\mathcal{L}_{\mathcal{H}}$-structure with domain X, and $\chi \in Ch(M)$. Set $\sigma := \{\chi(A) : A \in \binom{X}{r}\} \subseteq S_r(C_W, \mathcal{H})$. Then*

1. *σ is a syntactic $|X|$-diagram.*
2. *If $F \trianglelefteq_\chi M$ then $\sigma = Diag^{tp}(F)$.*

Proof. Clearly $V(\sigma) = C_X$. Let $m = |C_X|$. Note $\binom{C_X}{r} = \{C_A : A \in \binom{X}{r}\}$. By definition of σ, for each $A \in \binom{X}{r}$, $\{\chi(A)\} = Ch_\sigma(C_A)$. Thus $|Ch_\sigma(C_A)| = 1$ for all $A \in \binom{X}{r}$ and σ is a syntactic m-diagram. This shows 1 holds. For 2, suppose $F \trianglelefteq_\chi M$. This means $dom(F) = X$ and for all $A \in \binom{X}{r}$, $Diag^F(A) = \chi(A)$. Clearly this implies $F \models \sigma^F$, where F is considered with its natural C_X-structure. Part 1 of Observation 6 then implies $\sigma = Diag^{tp}(F)$. □

Definition 26. Given an integer ℓ and a set of constants C, set

$$Err_\ell(C) = \{\sigma \subseteq S_r(C) : \sigma \text{ is an unsatisfiable syntactic } \ell\text{-diagram}\}.$$

Elements of $Err_\ell(C)$ are *syntactic C-errors of size ℓ*.

Example 11. Let $\mathcal{L} = \{E, d_1, d_2, d_3\}$ and $\mathcal{P}$ be as in Example 6. Let $C = \{c_1, c_2, c_3, c_4\}$ be a set of constants. Recall from Example 6, that $q_1(c_1, c_2, c_3) \cup q_2(c_1, c_2, c_4)$ is unsatisfiable. Therefore an example of a syntactic C-error of size 4 is the set $\{q_1(c_2, c_3, c_4), q_2(c_1, c_2, c_3), q_1(c_1, c_3, c_4), q_1(c_1, c_2, c_4)\}$.

Lemma 6. *Suppose W is finite a set. If $r + 1 \leq \ell < 2r$ and M is a complete $\mathcal{L}_{\mathcal{H}}$-structure which is an error of size ℓ and with domain $X \subseteq W$, then there is a choice function $\chi \in Ch(M)$ such that $\{\chi(A) : A \in \binom{X}{r}\}$ is a syntactic C_W-error of size ℓ.*

Proof. Since M is an error of size ℓ then there are $\bar{a}_1, \bar{a}_2 \in X^{\underline{r}}$ such that $\cup\bar{a}_1 \bigcup \cup\bar{a}_2 = X$ and $p_1(\bar{x}), p_2(\bar{x}) \in S_r(\mathcal{H})$ such that $M \models R_{p_1}(\bar{a}_1) \wedge R_{p_2}(\bar{a}_2)$ but $p_1(c_{\bar{a}_1}) \cup p_2(c_{\bar{a}_2})$ is unsatisfiable. Define a function $\chi : \binom{X}{r} \to S_r(C_W, \mathcal{H})$ as follows. Set $\chi(\cup\bar{a}_1) = p(c_{\bar{a}_1})$ and $\chi(\cup\bar{a}_2) = p(c_{\bar{a}_2})$. For all other $A \in \binom{X}{r}$ choose any $\chi(A) \in Ch_M(A)$ (this is possible because M is a complete). By construction, $\chi \in Ch(M)$. By part 1 of Lemma 5, $\sigma := \{\chi(A) : A \in \binom{X}{r}\}$ is a syntactic ℓ-diagram. Because σ contains $p_1(c_{\bar{a}_1}) \cup p_2(c_{\bar{a}_2})$, it is unsatisfiable. By definition, σ is a syntactic C_W-error of size ℓ. □

Note Lemma 6 (and any statement concerning an integer $r + 1 \leq \ell < 2r$) is vacuous when $r = 1$.

8.2. Proof of Theorem 6.

In this section we state Theorem 10 and use it to prove Theorem 6.

Theorem 10. *Let $0 < \epsilon < 1$. For all $k \geq r$, there is a positive constant $c = c(k, r, \mathcal{L}, \epsilon)$ and $m = m(k, r) \geq 1$ such that for all sufficiently large n the following holds. Suppose $\mathcal{F}$ is a collection of finite $\mathcal{L}$-structures, each of size at most k, and $\mathcal{H} := \mathrm{Forb}(\mathcal{F}) \neq \emptyset$. For any set W of size n, there is a set $\Sigma \subseteq \mathcal{P}(S_r(C_W, \mathcal{H}))$ such that the following hold.*

1. *For all $\mathcal{F}$-free $\mathcal{L}$-structures M with domain W, there is $\sigma \in \Sigma$ such that $Diag^{tp}(M) \subseteq \sigma$.*
2. *For all $\sigma \in \Sigma$ the following hold. For each $1 \leq \ell \leq k$, $|Diag^{tp}(\mathcal{F}(\ell), C_W) \cap Span(\sigma)| \leq \epsilon\binom{n}{\ell}$, and for each $r + 1 \leq \ell < 2r$, $|Err_\ell(C_W) \cap Span(\sigma)| \leq \epsilon\binom{n}{\ell}$.*
3. *$\log|\Sigma| \leq cn^{r-\frac{1}{m}} \log n$.*

Given a collection $\mathcal{H}$ of finite $\mathcal{L}$-structures, we now define a way of building an $\mathcal{L}_{\mathcal{H}}$-template from a complete subset of $S_r(C_W, \mathcal{H})$.

Definition 27. Suppose $\mathcal{H}$ is a nonempty collection of $\mathcal{L}$-structures, W is a set, and $\sigma \subseteq S_r(C_W, \mathcal{H})$ is such that $V(\sigma) = C_W$. Define an $\mathcal{L}_{\mathcal{H}}$-structure D_σ as follows. Set $dom(D_\sigma) = W$ and for each $\bar{a} \in W^r$, define $D_\sigma \models R_p(\bar{a})$ if and only if $p(c_{\bar{a}}) \in Ch_\sigma(\cup\bar{a})$.

In the notation of Definition 27, note that for all $A \in \binom{W}{r}$, $Ch_{D_\sigma}(A) = Ch_\sigma(A)$ (here $Ch_{D_\sigma}(A)$ is in the sense of Definition 10 and $Ch_\sigma(A)$ is in the sense of Definition 22). We now prove two lemmas.

Lemma 7. *Suppose $\mathcal{F}$ is collection of finite $\mathcal{L}$-structures and $\mathcal{H} = \mathrm{Forb}(\mathcal{F}) \neq \emptyset$. For any set W and complete $\sigma \subseteq S_r(C_W, \mathcal{H})$, D_σ is an $\mathcal{L}_{\mathcal{H}}$-template.*

Proof. First, observe that D_σ is a complete $\mathcal{L}_{\mathcal{H}}$-structure because for every $A \in \binom{W}{r}$, we have $Ch_{D_\sigma}(A) = Ch_\sigma(C_A)$, and $Ch_\sigma(C_A) \neq \emptyset$ because σ is complete by assumption (in the sense of Definition 22). Suppose now $\bar{a} \in W^r \setminus W^{\underline{r}}$. Then because $S_r(\mathcal{H})$ contains only proper types, there is no $p(\bar{x}) \in S_r(\mathcal{H})$ such that $p(c_{\bar{a}}) \in S_r(C_W, \mathcal{H})$. Thus $D_\sigma \models \neg R_p(\bar{a})$ for all $p(\bar{x}) \in S_r(\mathcal{H})$, so D_σ satisfies part (1) of Definition 9. Suppose $p(\bar{x}), p'(\bar{x}) \in S_r(\mathcal{H})$ and $\mu \in Perm(r)$ are such that $p(\bar{x}) = p'(\mu(\bar{x}))$. Suppose $a \in W^{\underline{r}}$. Then by definition of D_σ, $D_\sigma \models R_p(\bar{a})$ if and only if $p(c_{\bar{a}}) \in \sigma$. Since $p(c_{\bar{a}}) = p'(c_{\mu(\bar{a})})$, $p(c_{\bar{a}}) \in \sigma$ if and only if $p'(c_{\mu(\bar{a})}) \in \sigma$. By definition of D_σ, $p'(c_{\mu(\bar{a})}) \in \sigma$ if and only if $D_\sigma \models R_{p'}(\mu(\bar{a}))$. Thus we've shown $D_\sigma \models R_p(\bar{a})$ if and only if $D_\sigma \models R_{p'}(\mu(\bar{a}))$, so D_σ satisfies part (2) of Definition 9. This finishes the verification that D_σ is an $\mathcal{L}_{\mathcal{H}}$-template. □

Lemma 8. *Suppose $k \geq r$, W is a finite set, $\mathcal{H}$ is a nonempty collection of finite $\mathcal{L}$-structures, and $\sigma \subseteq S_r(C_W, \mathcal{H})$ is complete. Suppose $\mathcal{F}$ is a collection of finite $\mathcal{L}$-structures, each of size at most k. Then for each $1 \leq \ell \leq k$, there is an injection*

$$\Phi : \text{cop}(\tilde{\mathcal{F}}(\ell), D_\sigma) \rightarrow Diag^{tp}(\mathcal{F}(\ell), C_W) \cap Span(\sigma).$$

and for each $r+1 \leq \ell < 2r$, there is an injection

$$\Theta : \text{cop}(\mathcal{E}(\ell), D_\sigma) \rightarrow Err_\ell(C_W) \cap Span(\sigma).$$

Proof. Without loss of generality, assume $\mathcal{F}$ is closed under isomorphism (we can do this because it does not change either of the sets $\text{cop}(\tilde{\mathcal{F}}(\ell), D_\sigma)$ or $Diag^{tp}(\mathcal{F}(\ell), C_W) \cap Span(\sigma)$). Suppose $1 \leq \ell \leq k$ and $G \in \text{cop}(\tilde{\mathcal{F}}(\ell), D_\sigma)$. Then $G \subseteq_{\mathcal{L}_\mathcal{H}} D_\sigma$ and $G \cong_{\mathcal{L}_\mathcal{H}} B$, for some $B \in \tilde{\mathcal{F}}(\ell)$. It is straightforward to check that since $\mathcal{F}$ is closed under isomorphism, this implies $G \in \tilde{\mathcal{F}}(\ell)$. So without loss of generality we may assume that $B = G$. Then there is some $F \in \mathcal{F}(\ell)$ such that $F \trianglelefteq_p G$. Choose any such F and let $\chi \in Ch(G)$ be such that $F \trianglelefteq_\chi G$. Define $\Phi(G) = \{\chi(A) : A \in \binom{dom(G)}{r}\}$. By part 2 of Lemma 5, $\Phi(G) = Diag^{tp}(F)$. Thus by definition, $\Phi(G) \in Diag^{tp}(\mathcal{F}(\ell), C_W)$. By definition of D_σ and because $\chi \in Ch(G)$, $G \subseteq_{\mathcal{L}_\mathcal{H}} D_\sigma$ implies $\Phi(G) \subseteq \sigma$, so $\Phi(G) \in Diag^{tp}(\mathcal{F}(\ell), C_W) \cap Span(\sigma)$, as desired. To see that Φ is injective, note that for all $G \in \text{cop}(\tilde{\mathcal{F}}(\ell), D_\sigma)$, $V(\Phi(G)) = dom(G)$. Therefore if $G_1 \neq G_2 \in \text{cop}(\tilde{\mathcal{F}}(\ell), D_\sigma)$, $dom(G_1) \neq dom(G_2)$ implies $V(\Phi(G_1)) \neq V(\Phi(G_2))$, so $\Phi(G_1) \neq \Phi(G_2)$.

Suppose now $r+1 \leq \ell < 2r$ and $G \in \text{cop}(\mathcal{E}(\ell), D_\sigma)$. Then G is a complete $\mathcal{L}_\mathcal{H}$-structure which is an error of size ℓ. Lemma 6 implies there is $\chi \in Ch(G)$ such that $\{\chi(A) : A \in \binom{dom(G)}{r}\}$ is a syntactic C_W-error of size ℓ. Set $\Theta(G) = \{\chi(A) : A \in \binom{dom(G)}{r}\}$. Then this shows $\Theta(G) \in Err_\ell(C_W)$. By definition of D_σ and because $\chi \in Ch(G)$, $G \subseteq_{\mathcal{L}_\mathcal{H}} D_\sigma$ implies $\Theta(G) \subseteq \sigma$, so $\Theta(G) \in Err_\ell(C_W) \cap Span(\sigma)$, as desired. To see that Θ is injective, note that for all $G \in \text{cop}(\mathcal{E}(\ell), D_\sigma)$, $V(\Theta(G)) = dom(G)$. Therefore if $G_1 \neq G_2 \in \text{cop}(\mathcal{E}(\ell), D_\sigma)$, $dom(G_1) \neq dom(G_2)$ implies $V(\Theta(G_1)) \neq V(\Theta(G_2))$, so $\Theta(G_1) \neq \Theta(G_2)$. ☐

Proof of Theorem 6 from Theorem 10. Let $0 < \epsilon < 1$ and let $k \geq r$ be an integer. Choose the constants $c = c(k, r, \mathcal{L}, \epsilon)$ and $m = m(k, r)$ to be the ones given by Theorem 10. Suppose $\mathcal{F}$ is a collection of finite $\mathcal{L}$-structures, each of size at most k, and $\mathcal{B} := \text{Forb}(\mathcal{F}) \neq \emptyset$. Suppose n is sufficiently large and W is a set of size n. Theorem 10 applied to $\mathcal{B}$ implies there exists a set $\Sigma \subseteq \mathcal{P}(S_r(C_W, \mathcal{B}))$ such that the following hold.

- (i) For all $\mathcal{F}$-free $\mathcal{L}$-structures M with domain W, there is $\sigma \in \Sigma$ such that $Diag^{tp}(M) \subseteq \sigma$.
- (ii) For all $\sigma \in \Sigma$ the following hold. For any $1 \leq \ell \leq k$, $|Diag^{tp}(\mathcal{F}(\ell), C_W) \cap Span(\sigma)| \leq \epsilon\binom{n}{\ell}$, and for any $r+1 \leq \ell < 2r$, $|Err_\ell(C_W) \cap Span(\sigma)| \leq \epsilon\binom{n}{\ell}$.
- (iii) $\log|\Sigma| \leq cn^{r-\frac{1}{m}} \log n$.

Set $\mathcal{D} = \{D_\sigma : \sigma \in \Sigma\}$, where for each $\sigma \in \Sigma$, D_σ is the $\mathcal{L}_\mathcal{B}$-structure from Definition 27. We claim this $\mathcal{D}$ satisfies conclusions of Theorem 6. First note (i) and part 3 of Observation 6 imply that every $\sigma \in \Sigma$ is complete in the sense of Definition 22. Therefore Lemma 7 implies each $D_\sigma \in \mathcal{D}$ is an $\mathcal{L}_\mathcal{B}$-template. We now verify parts (1)-(3) of Theorem 6 hold for this $\mathcal{D}$.

Clearly $|\mathcal{D}| \leq |\Sigma|$, so (iii) implies part (3) of Theorem 6 is satisfied. Suppose now M is an $\mathcal{F}$-free $\mathcal{L}$-structure with $dom(M) = W$. By (i), there is $\sigma \in \Sigma$ such that $Diag^{tp}(M) \subseteq \sigma$. We claim that $M \trianglelefteq_p D_\sigma$. Let $A \in \binom{W}{r}$ and suppose $p(\bar{x}) \in S_r(\mathcal{H})$ is such that $M \models p(\bar{a})$ for some enumeration $\bar{a}$ of A. Then $Diag^M(A) = p(c_{\bar{a}}) \in Diag^{tp}(M) \subseteq \sigma$ implies by definition of D_σ, $D_\sigma \models R_p(\bar{a})$, so $p(c_{\bar{a}}) \in Ch_{D_\sigma}(A)$. This shows $M \leq_p D_\sigma$. Then $M \trianglelefteq_p D_\sigma$ holds because by assumption $dom(M) = dom(D_\sigma) = W$. Thus part (1) of Theorem 6 is satisfied.

We now verify part (2) of Theorem 6. Let $D_\sigma \in \mathcal{D}$. We need to show $\text{prob}(\tilde{\mathcal{F}}, D_\sigma) \leq \epsilon$ and $\text{prob}(\mathcal{E}, D_\sigma) \leq \epsilon$. For each $1 \leq \ell \leq k$, we have

$$|\text{cop}(\tilde{\mathcal{F}}(\ell), D_\sigma)| \leq |Diag^{tp}(\mathcal{F}(\ell), C_W) \cap Span(\sigma)| \leq \epsilon \binom{n}{\ell},$$

where the first inequality is because of Lemma 8 and the second inequality is by (ii). This implies that for all $1 \leq \ell \leq k$, $|\text{cop}(\tilde{\mathcal{F}}(\ell), D_\sigma)| \leq \epsilon\binom{n}{\ell}$, so $\text{prob}(\tilde{\mathcal{F}}(\ell), D_\sigma) \leq \epsilon$. Since every element in $\tilde{\mathcal{F}}$ has size at most k, we have $\text{prob}(\tilde{\mathcal{F}}, D_\sigma) \leq \epsilon$. Similarly, for any $r+1 \leq \ell < 2r$,

$$|\text{cop}(\mathcal{E}(\ell), D_\sigma)| \leq |Err_\ell(C_W) \cap Span(\sigma)| \leq \epsilon \binom{n}{\ell},$$

where the first inequality is by Lemma 8 and the second inequality is by (ii). This implies for all $r+1 \leq \ell < 2r$, $|\text{cop}(\mathcal{E}(\ell), D_\sigma)| \leq \epsilon\binom{n}{\ell}$, so $\text{prob}(\mathcal{E}(\ell), D_\sigma) \leq \epsilon$. Since every element in $\mathcal{E}$ has size at least $r+1$ and at most $2r$, we have $\text{prob}(\mathcal{E}, D_\sigma) \leq \epsilon$. This finishes the proof. $\square$

9. Applying Hypergraph Containers to Prove Theorem 10

In this section we prove Theorem 10. We will use the hypergraph containers theorem. We begin with a definition.

Definition 28. Suppose K is a positive integer and $\mathcal{A}$ is a collection of finite $\mathcal{L}$-structures each of size at most K. Set $cl_K(\mathcal{A}) = \{M : M \text{ is an } \mathcal{L}\text{-structure of size } K \text{ such that } prob(\mathcal{A}, M) > 0\}$.

Observe that in the above notation, an $\mathcal{L}$-structure of size at least K is $\mathcal{A}$-free if and only if it is $cl_K(\mathcal{A})$-free. We now state a key lemma.

Lemma 9. *Assume $n \geq k \geq r$ and $\mathcal{F}$ is a nonempty collection of $\mathcal{L}$-structures, each of size at most k. Suppose $\mathcal{H} := \text{Forb}(\mathcal{F}) \neq \emptyset$ and W is a set of size n. Fix $0 < \epsilon < 1/2$. Suppose $\sigma \subseteq S_r(C_W, \mathcal{H})$ is complete and satisfies $V(\sigma) = C_W$. If*

$$|(Diag^{tp}(cl_k(\mathcal{F}), C_W) \cup Err_k(C_W)) \cap Span(\sigma)| \leq \epsilon \binom{n}{k}$$

holds, then for all $1 \leq \ell \leq k$, $|(Diag^{tp}(\mathcal{F}(\ell), C_W) \cup Err_\ell(C_W)) \cap Span(\sigma)| \leq \epsilon\binom{n}{\ell}$.

Proof. For $1 \leq \ell < k$, set $\Gamma(\ell) = (Diag^{tp}(\mathcal{F}(\ell), C_W) \cup Err_\ell(C_W)) \cap Span(\sigma)$ and let

$$\Gamma(k) = (Diag^{tp}(cl_k(\mathcal{F}), C_W) \cup Err_k(C_W)) \cap Span(\sigma).$$

We want to show that $|\Gamma(k)| \leq \epsilon\binom{n}{k}$ implies that for all $\ell \in [k]$, $|\Gamma(\ell)| \leq \epsilon\binom{n}{\ell}$. If $\ell = k$, this is immediate. Fix $1 \leq \ell < k$. We claim the following holds.

$$\text{For all } S_0 \in \Gamma(\ell),\ |\{S_1 \in \Gamma(k) : S_0 \subseteq S_1\}| \geq \binom{n-\ell}{r-\ell}. \tag{9}$$

Suppose $S_0 \in \Gamma(\ell)$. Consider the following procedure for constructing a set $S_1 \subseteq S_r(C_W, \mathcal{H})$.

- Choose $X \in \binom{C_W}{k}$ such that $V(S_0) \subseteq X$. There are $\binom{n-\ell}{k-\ell}$ choices.
- For each $A \in \binom{X}{r} \setminus \binom{V(S_0)}{r}$, choose some $p_A \in Ch_\sigma(A)$ (this is possible since σ is complete).
- Set $S_1 = S_0 \cup \{p_A : A \in \binom{X}{r} \setminus \binom{V(S_0)}{r}\}$.

Suppose S_1 is constructed from S_0 in this way. We claim $S_1 \in \Gamma(k)$. By construction and because S_0 is a syntactic ℓ-diagram, S_1 is a syntactic k-diagram. Also by construction, $S_1 \subseteq \sigma$, so by definition, $S_1 \in Span(\sigma)$. If S_1 is unsatisfiable, then by definition $S_1 \in Err_k(C_W) \cap Span(\sigma) \subseteq \Gamma(k)$, so we are done. Suppose now S_1 is satisfiable and M is an $\mathcal{L} \cup V(S_1)$-structure such that $M \models S_1^M$. Let $N = M[V(S_0)^M]$. Then considered as an $\mathcal{L} \cup V(S_0)$-structure, $N \models S_0^N$, so part 1 of Observation 6 implies $Diag^{tp}(N) = S_0$. On the other hand, $S_0 \in Diag^{tp}(\mathcal{F}(\ell), C_W)$ implies there is $F \in F(\ell)$ which can be made into an $\mathcal{L} \cup V(S_0)$-structure such that $F \models S_0^F$. Part 2 of Observation 6 then implies $N \cong_{\mathcal{L}} F$. Since $\mathcal{F}$ is closed under isomorphism, $N \in \mathcal{F}$. Since $N \subseteq_{\mathcal{L}} M$ and M has size k, this implies by definition that $M \in cl_k(\mathcal{F})$. Since $S_1 = Diag^{tp}(M, V(S_1))$, we have $S_1 \in Diag^{tp}(cl_k(\mathcal{F}), C_W)$ by definition. Thus we have shown that S_1 is in $\Gamma(k)$. Observe that every distinct choice of X produces a distinct S_1, so this construction produces at least $\binom{n-\ell}{k-\ell}$ distinct $S_1 \in \Gamma(k)$ such that $S_0 \subseteq S_1$. We we have proved (9) holds for all $1 \le \ell < k$. Consider the following procedure for constructing element in $S_0 \in \Gamma(\ell)$:

- Choose $S_1 \in \Gamma(k)$. There are $|\Gamma(k)|$ choices.
- Choose $S_0 \subseteq S_1$ such that $S_0 \in \Gamma(\ell)$ (if one exists). There are at most $\binom{V(S_1)}{\ell} = \binom{k}{\ell}$ choices.

By (9), this construction produces every element $S_0 \in \Gamma(\ell)$ at least $\binom{n-\ell}{k-\ell}$ times. This shows

$$|\Gamma(\ell)| \le |\Gamma(k)|\binom{k}{\ell}\Big/\binom{n-\ell}{k-\ell} \le \epsilon\binom{n}{k}\binom{k}{\ell}\Big/\binom{n-\ell}{k-\ell} = \epsilon\binom{n}{\ell},$$

where the second inequality is because $|\Gamma(k)| \le \epsilon\binom{n}{k}$ by assumption. □

We now present a computational lemma which will be used in the proof of Theorem 10.

Lemma 10. *For all integers* $1 \le x < y$, $m(y,x) := \max\left\{\frac{\binom{\ell}{x}-1}{\ell-x} : x < \ell \le y\right\} \ge 1$.

Proof. We show that for all $1 \le x < y$, $\binom{y}{x} \ge y - x + 1$. Since by definition, $m(y,x) \ge \frac{\binom{y}{x}-1}{y-x}$, this will imply $m(y,x) \ge 1$, as desired. Fix $x \ge 1$. We induct on t where $y = x+t$. Suppose first $y = x+1$. Then $\binom{y}{x} = \frac{(x+1)!}{x!} = x+1 \ge 2 = y-x+1$, where the inequality is because $x \ge 1$. Assume now that $y = x+t$ where $t > 1$, and suppose by induction $\binom{y-1}{x} \ge (y-1)-x+1$. Then $\binom{y}{x} = \frac{(y-1)!y}{x!(y-x-1)!(y-x)} = \binom{y-1}{x}\frac{y}{y-x}$. By our induction hypothesis,

$$\binom{y-1}{x}\frac{y}{y-x} \ge ((y-1)-x+1)\left(\frac{y}{y-x}\right) = (y-x)\frac{y}{y-x} = y \ge y-x+1,$$

where the last inequality is because $x \ge 1$. Thus $\binom{y}{x} \ge y-x+1$, as desired. □

Defining the Hypergraph.

We now give a procedure for defining a special hypergraph given a finite set and a collection of $\mathcal{L}$-structures satisfying certain properties. Assume we are given the following.

- An integer $k \ge 2r$ and a nonempty collection, $\mathcal{F}$, of finite $\mathcal{L}$-structures, each of size at most k.
- A set W of size n, where $n \ge k$ is an integer.

Let $\mathcal{H}$ be the class of all finite $cl_k(\mathcal{F})$-free $\mathcal{L}$-structures. Define the hypergraph $H = H(\mathcal{F}, W)$ as follows.

$$V(H) = S_r(C_W, \mathcal{H}) \qquad \text{and} \qquad E(H) = Diag^{tp}(cl_k(\mathcal{F}), C_W) \cup Err_k(C_W).$$

We now make a few observations about H. First, note that the edges of H are syntactic k-diagrams, so H is a $\binom{k}{r}$-uniform hypergraph. By definition $|V(H)| = \binom{n}{r}|S_r(\mathcal{H})|$. If X and X' are both in $\binom{C_W}{k}$, then since relabeling constants does not change the satisfiability properties of a collection of $\mathcal{L}\cup C_W$-sentences, we must have $|Diag^{tp}(Cl_k(\mathcal{F}), X) \cup Err_k(X)| = |Diag^{tp}(Cl_k(\mathcal{F}), X') \cup Err_k(X')|$. Therefore, the following is well defined.

Definition 29. Let $\alpha = \alpha(\mathcal{F})$ be such that for all $X \in \binom{C_W}{k}$, $|Diag^{tp}(Cl_k(\mathcal{F}), X) \cup Err_k(X)| = \alpha$.

We claim that $|E(H)| = \alpha\binom{n}{k}$. Indeed, any $\sigma \in E(H)$ can be constructed as follows.

- Choose $X \in \binom{C_W}{k}$. There are $\binom{n}{k}$ choices.
- Choose an element $\sigma \in Diag^{tp}(Cl_k(\mathcal{F}), C_W) \cup Err_k(C_W)$ such that $V(\sigma) = X$, i.e. choose an element $\sigma \in Diag^{tp}(Cl_k(\mathcal{F}), X) \cup Err_k(X)$. There are α choices.

Each of these choices lead to distinct subsets $\sigma \in E(H)$, so this shows $|E(H)| = \alpha\binom{n}{k}$. Note that because $\mathcal{F} \neq \emptyset$, $\alpha \geq 1$. On the other hand, there are at most $|S_r(\mathcal{H})|^{\binom{k}{r}}$ syntactic k-diagrams σ with $V(\sigma) = X$, so $\alpha \leq |S_r(\mathcal{H})|^{\binom{k}{r}} \leq |S_r(\mathcal{L})|^{\binom{k}{r}}$. We now make a key observation about this hypergraph.

Proposition 4. *For any $\mathcal{F}$-free $\mathcal{L}$-structure M with domain W, $Diag^{tp}(M)$ is an independent subset of $V(H)$.*

Proof. Suppose towards a contradiction that $Diag^{tp}(M)$ contains an edge $\sigma \in E(H)$. Then σ is a syntactic k-diagram which is either in $Err_k(C_W)$ or $Diag^{tp}(cl_k(\mathcal{F}), C_W)$. Clearly $\sigma \notin Err_k(C_W)$, since $M \models \sigma^M$ implies σ is satisfiable. Thus $\sigma \in Diag^{tp}(cl_k(\mathcal{F}), C_W)$. So there is an $\mathcal{L} \cup V(\sigma)$-structure B such that $B \restriction_{\mathcal{L}} \in cl_k(\mathcal{F})$ and $Diag^{tp}(B, V(\sigma)) = \sigma$. Let $A = \{a : c_a \in V(\sigma)\} \subseteq W$ and let $N = M[A]$. Suppose $p(c_{\bar{a}}) \in \sigma$. Since $\sigma \subseteq Diag^{tp}(M)$, $M \models p(\bar{a})$. Since $N \subseteq_{\mathcal{L}} M$ and $\cup\bar{a} \subseteq A = dom(N)$, we have $N \models p(\bar{a})$. This shows that with its canonical $\mathcal{L} \cup V(\sigma)$-structure, $N \models \sigma^N$. Since σ is a syntactic k-diagram and N has size k, part 1 of Observation 6 implies $\sigma = Diag^{tp}(N)$. Now $Diag^{tp}(N) = \sigma = Diag^{tp}(B, V(\sigma))$ implies by part 2 of Observation 6, that $N \cong_{\mathcal{L}} B$. But now N is an $\mathcal{L}$- substructure of M isomorphic to an element of $cl_k(\mathcal{F})$, contradicting our assumption that M is $\mathcal{F}$-free (recall $|dom(M)| = n \geq k$ implies M is $cl_k(\mathcal{F})$-free if and only if M if $\mathcal{F}$-free). □

Observe that by definition of H, if $S \subseteq V(H)$, then

$$E(H[S]) = \Big(Diag^{tp}(cl_k(\mathcal{F}), C_W) \cup Err_k(C_W)\Big) \cap Span(S). \tag{10}$$

We are now ready to prove Theorem 10. At this point the reader may wish to review Theorem 1 and its corresponding notation in Subsection 2.2, as Theorem 10 is the key tool used in this proof.

Proof of Theorem 10. Clearly it suffices to show the theorem holds for all $0 < \epsilon < 1/2$. We claim that further, it suffices to show the theorem holds for any $k \geq 2r$. Indeed, suppose $k < 2r$ and Theorem 10 holds for all $k' \geq 2r$. Suppose $\mathcal{F}$ is a nonempty collection of $\mathcal{L}$-structures, each of size at most k and $\mathcal{H} := \text{Forb}(\mathcal{F}) \neq \emptyset$. Then $\mathcal{F}$ is also a collection of $\mathcal{L}$-structures, each of size at most $k' = 2r$. Apply Theorem 10 to $k' = 2r$ to obtain constants $c = c(2r, r, \mathcal{L}, \epsilon)$ and $m = m(2r, r)$. Since $k < 2r$, it is clear the conclusions of Theorem 10 for $\mathcal{H}$ and $2r$ imply the conclusions of Theorem 10 for $\mathcal{H}$ and k. Thus we may take $c(k, r, \epsilon) = c(2r, r, \epsilon)$ and $m(k, r) = m(2r, r)$. We now prove the theorem holds for all $0 < \epsilon < 1/2$ and $k \geq 2r$.

Fix $0 < \epsilon < 1/2$ and $k \geq 2r$. Apply Theorem 1 to $s := \binom{k}{r}$ to obtain the constant $c_0 = c_0(s)$ (note $k \geq 2r$ implies $s \geq 2$). Set

$$m = m(k, r) = \max\Big\{\frac{\binom{\ell}{r} - 1}{\ell - r} : r < \ell \leq k\Big\}.$$

By Lemma 10, since $1 \leq r < k$, $m \geq 1$. Set $\epsilon' = \epsilon/|S_r(\mathcal{L})|^{\binom{k}{r}}$ and choose $0 < \gamma < 1$ sufficiently small so that

$$2^{\binom{\binom{k}{r}}{2}+1}|S_r(\mathcal{L})|r!(k-r)^{k-r}\gamma \leq \frac{\epsilon'}{12\binom{k}{2}!}. \tag{11}$$

Now set $c = c(k, r, \mathcal{L}, \epsilon) = (c_0 \log(\frac{1}{\epsilon'})|S_r(\mathcal{L})|)/\gamma m$. Observe that c actually depends on $\mathcal{L}$, not just $r_{\mathcal{L}}$. Let $M \geq k$ be such that $n \geq M$ implies $(n-r)^{k-r} \geq n^{k-r}/2$, and $n^{-\frac{1}{m}}\gamma^{-1} < 1/2$. We show Theorem 10 holds for this c and m for all $n \geq M$.

Fix $n \geq M$. Suppose $\mathcal{F}$ a nonempty collection of finite $\mathcal{L}$-structures, each of size at most k, such that $\mathcal{H} := \text{Forb}(\mathcal{F}) \neq \emptyset$. Let W be a set of size n and let $H = H(\mathcal{F}, W)$ be the $\binom{k}{r}$-uniform hypergraph described above. Set $\tau = n^{\frac{-1}{m}}\gamma^{-1}$. By our assumptions we have that $0 < \epsilon', \tau < \frac{1}{2}$.

We show that $\delta(H,\tau) \leq \frac{\epsilon'}{12\binom{k}{r}!}$ so that we may apply Theorem 1 to H. Let $\alpha = \alpha(\mathcal{F})$ be as in Definition 29 so that $E(H) = \alpha\binom{n}{k}$ and let $N = |V(H)|$. Given $2 \leq j \leq \binom{k}{r}$, set

$$f(j) = \min\{\ell : \binom{\ell}{r} \geq j\}. \tag{12}$$

Observe that for each $2 \leq j \leq \binom{k}{r}$, $r < f(j) \leq k$. Indeed, $r < f(j)$ holds since $\binom{f(j)}{r} \geq j \geq 2$, and $f(j) \leq k$ holds since $k \in \{\ell : \binom{\ell}{r} \geq j\}$. Thus by definition of m, for each $2 \leq j \leq \binom{k}{r}$,

$$m \geq \frac{\binom{f(j)}{r} - 1}{f(j) - r} \geq \frac{j-1}{f(j)-r}, \tag{13}$$

where the inequality is because by (12), $\binom{f(j)}{r} \geq j$.

Claim 1. *For all $\sigma \subseteq V(H)$ with $2 \leq |\sigma| \leq \binom{k}{r}$, $d(\sigma) \leq \alpha n^{k-f(|\sigma|)}$.*

Proof. Fix $\sigma \subseteq V(H)$ so that $2 \leq |\sigma| \leq \binom{k}{r}$. If $|V(\sigma)| > k$, then $\{e \in E(H) : \sigma \subseteq e\} = \emptyset$ since every $e \in E(H)$ is a syntactic k-diagram, so must satisfy $|V(e)| = k$. Therefore $d(\sigma) = 0 \leq \alpha n^{k-f(|\sigma|)}$. Similarly, if there is $A \in \binom{V(\sigma)}{r}$ such that $|Ch_\sigma(A)| \geq 2$, then $\{e \in E(H) : \sigma \subseteq e\} = \emptyset$, since every $e \in E(H)$ is a syntactic k-diagram, so must satisfy $|Ch_e(A)| = 1$. Therefore, $d(\sigma) = 0 \leq \alpha n^{k-f(|\sigma|)}$. Suppose now $|V(\sigma)| \leq k$ and $|Ch_\sigma(A)| \leq 1$ for all $A \in \binom{V(\sigma)}{r}$. This implies $|\sigma| \leq \binom{|V(\sigma)|}{r}$, so by (12), $f(|\sigma|) \leq |V(\sigma)|$. Every edge in H containing σ can be constructed as follows.

- Choose a set $X \in \binom{C_W}{k}$ such that $V(\sigma) \subseteq X$ (this makes sense since $|V(\sigma)| \leq k$). There are $\binom{n-|V(\sigma)|}{k-|V(\sigma)|}$ ways to do this.
- Choose an element of $Diag^{tp}(cl_k(\mathcal{F}), X) \cup Err_k(X)$ containing σ. By definition of α, there are at most α choices for this.

Therefore, $d(\sigma) = |\{e \in E(H) : \sigma \subseteq e\}| \leq \alpha\binom{n-|V(\sigma)|}{k-|V(\sigma)|} \leq \alpha n^{k-|V(\sigma)|} \leq \alpha n^{k-f(|\sigma|)}$, where the last inequality is because $f(|\sigma|) \leq |V(\sigma)|$. Thus we have shown that for any $2 \leq j \leq \binom{k}{r}$ and $\sigma \subseteq V(H)$ with $|\sigma| = j$, $d(\sigma) \leq \alpha n^{k-f(j)}$. □

By Claim 1, for any $2 \leq j \leq \binom{k}{r}$ and $v \in V(H)$, $d^{(j)}(v) = \max\{d(\sigma) : v \in \sigma, |\sigma| = j\} \leq \alpha n^{k-f(j)}$. Combining this with the definition of τ yields that for all $2 \leq j \leq \binom{k}{r}$,

$$\delta_j = \frac{\sum_{v \in V(H)} d^{(j)}(v)}{N d \tau^{j-1}} \leq \frac{N n^{k-f(j)}}{N d \tau^{j-1}} = \frac{n^{k-f(j)+(j-1)\frac{1}{m}}\gamma^{j-1}}{d}. \tag{14}$$

Note the average degree of H is

$$d = \binom{k}{r}|E(H)|/|V(H)| = \frac{\binom{k}{r}\alpha\binom{n}{k}}{|S_r(\mathcal{H})|\binom{n}{r}} = \frac{\alpha}{|S_r(\mathcal{H})|}\binom{n-r}{k-r} \geq \frac{\alpha}{|S_r(\mathcal{H})|}\Big(\frac{n-r}{k-r}\Big)^{k-r}.$$

Combining this with our assumption n, we obtain the following inequality.

$$d \geq \frac{\alpha}{|S_r(\mathcal{H})|}\Big(\frac{n-r}{k-r}\Big)^{k-r} = \frac{\alpha}{|S_r(\mathcal{H})|(k-r)^{k-r}}(n-r)^{k-r} \geq \frac{\alpha}{2|S_r(\mathcal{H})|(k-r)^{k-r}} n^{k-r}. \tag{15}$$

Combining (15) and (14) yields that for all $2 \leq j \leq \binom{k}{r}$,

$$\delta_j \leq 2|S_r(\mathcal{H})|(k-r)^{k-r}\gamma^{j-1}\alpha^{-1}n^{k-f(j)+\frac{j-1}{m}-k+r} = 2|S_r(\mathcal{H})|(k-r)^{k-r}\gamma^{j-1}\alpha^{-1}n^{r-f(j)+\frac{j-1}{m}}.$$

By (13), $\frac{j-1}{m} \leq f(j) - r$, so this implies δ_j is at most

$$2|S_r(\mathcal{H})|(k-r)^{k-r}\gamma^{j-1}\alpha^{-1}n^{r-f(j)+f(j)-r} = 2|S_r(\mathcal{H})|(k-r)^{k-r}\gamma^{j-1}\alpha^{-1} \leq 2|S_r(\mathcal{H})|(k-r)^{k-r}\gamma,$$

where the last inequality is because $j \geq 2$ and $\gamma < 1$ implies $\gamma^{j-1} < \gamma$ and $\mathcal{F} \neq \emptyset$ implies $\alpha^{-1} \leq 1$. Therefore

$$\delta(H,\tau) = 2^{\binom{\binom{k}{r}}{2}-1}\sum_{j=2}^{\binom{k}{r}} 2^{-\binom{j-1}{2}}\delta_j \leq 2^{\binom{\binom{k}{r}}{2}-1}2|S_r(\mathcal{H})|(k-r)^{k-r}\gamma\sum_{j=2}^{\binom{k}{r}} 2^{-\binom{j-1}{2}}. \tag{16}$$

If $t = \binom{\binom{k}{r}}{2}$, then $\sum_{j=2}^{\binom{k}{r}} 2^{-\binom{j-1}{2}} \leq \sum_{j=0}^{t} 2^{-t}$. Using the formula for summing finite geometric series, $\sum_{j=0}^{t} 2^{-t} = \frac{1-2^{-t-1}}{1-2^{-1}} = 2(1-2^{-t-1}) < 2$. Plugging this into (16) yields

$$\delta(H,\tau) \leq 2^{\binom{\binom{k}{r}}{2}-1}2|S_r(\mathcal{H})|(k-r)^{k-r}\gamma 2 = 2^{\binom{\binom{k}{r}}{2}+1}|S_r(\mathcal{H})|(k-r)^{k-r}\gamma \leq 2^{\binom{\binom{k}{r}}{2}+1}|S_r(\mathcal{L})|(k-r)^{k-r}\gamma.$$

By (11), the right hand side above is at most $\frac{\epsilon'}{12\binom{k}{r}!}$, so we have shown that $\delta(H,\tau) \leq \frac{\epsilon'}{12\binom{k}{r}!}$. Thus Theorem 1 implies there is $\Sigma_0 \subseteq \mathcal{P}(V(H))$ with the following properties.

(i) For every independent set $I \subseteq V(H)$, there is $\sigma \in \Sigma_0$ such that $I \subseteq \sigma$.
(ii) For every $\sigma \in \Sigma_0$, $e(H[\sigma]) \leq \epsilon' e(H)$.
(iii) $\log|\Sigma_0| \leq c_0 \log(1/\epsilon')N\tau\log(1/\tau)$.

Define $\Sigma = \{\sigma \in \Sigma_0 : \exists$ an $\mathcal{F}$-free $\mathcal{L}$-structure M with domain W such that $Diag^{tp}(M) \subseteq \sigma\}$. Observe that every $\sigma \in \Sigma$ is complete by part 3 of Observation 6. We show Σ satisfies the conclusions (1)-(3) of Theorem 10. Suppose M is an $\mathcal{F}$-free $\mathcal{L}$-structure with domain W. Proposition 4 implies $Diag^{tp}(M)$ is an independent subset of $V(H)$, so by (i) and definition of Σ, there is $\sigma \in \Sigma$ such that $Diag^{tp}(M) \subseteq \sigma$. Thus part (1) of Theorem 10 holds. We now show part (2) holds. Fix $\sigma \in \Sigma$. By (10), $(Diag^{tp}(cl_k(\mathcal{F}), C_W) \cup Err_k(C_W)) \cap Span(\sigma) = E(H[\sigma])$. So (ii) implies

$$e(H[\sigma]) = |(Diag^{tp}(cl_k(\mathcal{F}), C_W) \cup Err_k(C_W)) \cap Span(\sigma)| \leq \epsilon' e(H).$$

By definition of ϵ' and because $\alpha \leq |S_r(\mathcal{L})|^{\binom{k}{r}}$,

$$\epsilon' e(H) = \epsilon'\alpha\binom{n}{k} = \frac{\epsilon}{|S_r(\mathcal{L})|^{\binom{k}{r}}}\alpha\binom{n}{k} \leq \epsilon\binom{n}{k}.$$

Thus $|(Diag^{tp}(cl_k(\mathcal{F}), C_W) \cup Err_k(C_W)) \cap Span(\sigma)| \leq \epsilon\binom{n}{k}$. By Lemma 9, for all $1 \leq \ell \leq k$,

$$|(Diag^{tp}(\mathcal{F}(\ell), C_W) \cup Err_\ell(C_W)) \cap Span(\sigma)| \leq \epsilon\binom{n}{\ell}.$$

Since $k \geq 2r$, this immediately implies part (2) of Theorem 10 holds. By (iii), definition of c, and because $\Sigma \subseteq \Sigma_0$ we have that

$$\begin{aligned}|\Sigma| \leq |\Sigma_0| \leq c_0\log(1/\epsilon')N\tau\log(1/\tau) &= c_0\log(1/\epsilon')|S_r(\mathcal{H})|\binom{n}{r}\tau\log(1/\tau)\\ &\leq c_0\log(1/\epsilon')|S_r(\mathcal{L})|\binom{n}{r}\tau\log(1/\tau) = c\gamma m\binom{n}{r}\tau\log\left(\frac{1}{\tau}\right).\end{aligned}$$

This shows $|\Sigma| \leq c\gamma m n^r\tau\log(\frac{1}{\tau})$. By definition of τ,

$$c\gamma m n^r\tau\log\left(\frac{1}{\tau}\right) = cmn^{r-\frac{1}{m}}\left(\frac{\log n}{m} + \log\gamma\right) = cn^{r-\frac{1}{m}}\Big(\log n + m\log\gamma\Big) \leq cn^{r-\frac{1}{m}}\log n,$$

where the last inequality is because $\gamma \leq 1 \leq m$ implies $m\log\gamma \leq 0$. This shows part (3) of Theorem 10 holds, so we are done. ☐

10. Concluding Remarks

We end with some questions and conjectures. In an extended version of this paper appearing on the arxiv [41], we consider in detail how our results relate to the example of discrete metric spaces. We give a brief outline of some interesting behavior arising from this anaylysis as motivation for Conjecture 1 below. Given integers $r \geq 3$ and $n \geq 1$, let $M_r(n)$ be the set of metric spaces with distances all in $[r]$ and underlying set $[n]$. In [31] the author and Mubayi proved the following theorem, where $\mathcal{L}_r = \{d_i : i \in [r]\}$ is a relational language with one relation symbol for each distance in $[r]$.

Theorem 11 (**Mubayi-Terry [31]**). *If $r \geq 2$ is even, then $M_r = \bigcup_{n\in\mathbb{N}} M_r(n)$ has a 0-1 law in the language $\mathcal{L}_r$.*

It was then conjectured in [31] that this theorem is false in the case when r is odd (to our knowledge, this is still open). In Section 10 of [41], we prove the hereditary property associated to $\bigcup_{n\in\mathbb{N}} M_r(n)$ has a stability theorem in the sense of Definition 19, but when r is odd, we prove this is not the case. These facts lead us to make the following conjecture.

Conjecture 1. *Suppose $\mathcal{L}$ is a finite relational language, and $\mathcal{H}$ is a fast-growing hereditary $\mathcal{L}$-property such that $\bigcup_{n\in\mathbb{N}} \mathcal{H}_n$ has a 0-1 law. Then $\mathcal{H}$ has a stability theorem.*

The idea behind this conjecture is that if $\mathcal{H}$ has nice asymptotic structure, it should reflect the structure of $\mathcal{R}_{ex}([n], \mathcal{H})$. Another phenomenon which can be observed from known examples is that the structures in $\mathcal{R}_{ex}(n, \mathcal{H})$ are not very complicated. The following questions are various ways of asking if this is always the case.

Question 1. *Suppose $\mathcal{L}$ is a finite relational language, and $\mathcal{H}$ is a fast-growing hereditary $\mathcal{L}$-property. For each n, let $\mathcal{P}_n = \mathcal{R}_{ex}([n], \mathcal{H})$. Does $\bigcup_{n\in\mathbb{N}} \mathcal{P}_n$ always have a 0-1 law?*

We direct the reader to [39] for the definition of a formula having the k-order property.

Question 2. *Suppose $\mathcal{L}$ is a finite relational language, and $\mathcal{H}$ is a fast-growing hereditary $\mathcal{L}$-property. Is there a finite $k = k(\mathcal{H})$ such that for all n and $M \in \mathcal{R}_{ex}(n, \mathcal{H})$, every atomic formula $\phi(x; y) \in \mathcal{L}_{\mathcal{H}}$ does not have the k-order property in M?*

A weaker version of this question is the following. We direct the reader to [40] for the definition of the VC-dimension of a formula.

Question 3. *Suppose $\mathcal{L}$ is a finite relational language, and $\mathcal{H}$ is a fast-growing hereditary $\mathcal{L}$-property. Is there a finite $k = k(\mathcal{H})$ such that for all n and $M \in \mathcal{R}_{ex}(n, \mathcal{H})$, every atomic formula $\phi(x; y) \in \mathcal{L}_{\mathcal{H}}$ has VC-dimension bounded by k in M?*

Appendix A. Proof of Theorem 9 from [5]

In this section we will prove Theorem 9 from [5]. Throughout this section, $\mathcal{L}_0$ is a fixed finite relational language with $r_{\mathcal{L}_0} = r$. Note $\mathcal{L}_0$ is not necessarily the same as $\mathcal{L}$, although we are assuming $r_{\mathcal{L}_0} = r_{\mathcal{L}} = r$. Given a partition p of a finite set X, $||p||$ denotes the number of parts in p.

Definition 30. Let $Index = \{(R, p) : R \in \mathcal{L}_0 \text{ and } p \text{ is a partition of } [\ell] \text{ where } \ell \text{ is the arity of } R\}$. Suppose $(R, p) \in Index$ and R has arity ℓ.

1. $C_p(x_1, \ldots, x_\ell)$ is the subtuple of $(x_1, \ldots, x_\ell)$ obtained by replacing each x_i with $x_{p(i)}$ where $p(i) = \min\{j : x_j \text{ is in the same part of } p \text{ as } i\}$, then deleting all but the first occurance of each variable in the tuple $(x_{p(1)}, \ldots, x_{p(\ell)})$.
2. $R_p(C(\bar{x}))$ is the $||p||$-ary relation obtained from $R(x_1, \ldots, x_\ell)$ by replacing each x_i with $x_{p(i)}$ where $p(i) = \min\{j : x_j \text{ is in the same part of } p \text{ as } i\}$.
3. If N is an $\mathcal{L}_0$-structure, define $DH^R_p(N) = \{\bar{a} \in dom(N)^{\underline{||p||}} : N \models R_p(\bar{a})\}$.

Now we can define the notion of distance between two $\mathcal{L}_0$-structures from [5].

Definition 31. Given $(R,p) \in Index$ and M, N two finite $\mathcal{L}_0$-structures with the same universe W, set

$$d_p^R(M,N) = \frac{|DH_p^R(M)\Delta DH_p^R(N)|}{|W|^{||p||}} \qquad \text{and set} \qquad d(M,N) = \sum_{(R,p)\in Index} d_p^R(M,N).$$

We now state the graph removal lemma of Aroskar and Cummings (Theorem 2 from [5]).

Theorem 12 (**Aroskar-Cummings [5]**). *Suppose $\mathcal{A}$ is a collection of finite $\mathcal{L}_0$-structures. For every $\delta > 0$ there exists $\epsilon > 0$ and K such that the following holds. For all sufficiently large finite $\mathcal{L}_0$-structures M, if* $\text{prob}(\mathcal{A}(K), M) < \epsilon$*, then there is an $\mathcal{L}_0$-structure M' with $dom(M') = dom(M)$ such that $d(M', M) < \delta$ and* $\text{prob}(\mathcal{A}, M') = 0$.

We now prove a key fact about relationship between $d(M,N)$ and $\text{dist}(M,N)$. Given a tuple $\bar{x} = (x_1, \dots, x_\ell)$, a *subtuple* of $\bar{x}$ is any tuple $\bar{x}' = (x_{i_1}, \dots, x_{i_{\ell'}})$ where $1 \leq i_1 < \ldots < i_{\ell'} \leq \ell$. If $\ell' < \ell$, we say $\bar{x}'$ is a *proper subtuple* of $\bar{x}$, denoted $\bar{x}' \subsetneq \bar{x}$.

Lemma 11. *If M and N are $\mathcal{L}_0$-structures with the same finite domain W of size at least $2r$, then*

$$\text{dist}(M,N) \leq (r!)^2 2^r d(M,N).$$

Proof. Let $n = |W|$. Note that $n \geq 2r$ implies for all $1 \leq \ell \leq r$,

$$\frac{n!}{(n-\ell)!} = n \cdot (n-1) \cdots (n-\ell+1) \geq (n-\ell+1)^\ell \geq (n/2)^\ell = n^\ell/2^\ell. \tag{17}$$

Given $1 \leq \ell \leq r$, define

$$\text{diff}^\ell(M,N) = \{\bar{a} \in W^{\underline{\ell}} : qftp^M(\bar{a}) \neq qftp^N(\bar{a}) \text{ and for all } \bar{a}' \subsetneq \bar{a},\ qftp^M(\bar{a}') = qftp^N(\bar{a}')\}.$$

Observe that elements in $\text{diff}(M,N)$ are *sets* of elements from W, while elements in $\text{diff}^\ell(M,N)$ are *tuples* of elements of W. Clearly if $A \in \text{diff}(M,N)$, there is some $\ell \in [r]$ and a tuple $\bar{a} \in A^{\underline{\ell}}$ such that $\bar{a} \in \text{diff}^\ell(M,N)$. Define $\Psi : \text{diff}(M,N) \to \bigcup_{\ell\in[r]} \text{diff}^\ell(M,N)$ to be any map which sends each $A \in \text{diff}(M,N)$ to some such tuple. Given $\ell \in [r]$ and $\bar{a} = (a_1, \dots, a_\ell) \in \text{diff}^\ell(M,N)$, note that

$$\Psi^{-1}(\bar{a}) \subseteq \{A \in \binom{W}{r} : \cup\bar{a} \subseteq A\}.$$

Since the right hand side has size $\binom{n-\ell}{r-\ell}$, we have that for all $\bar{a} \in \text{diff}^\ell(M,N)$, $|\Psi^{-1}(\bar{a})| \leq \binom{n-\ell}{r-\ell}$.

For each $\ell \in [r]$, we now define a map $f_\ell : \text{diff}^\ell(M,N) \to \bigcup_{(R,p)\in Index, ||p||=\ell} DH_p^R(M)\Delta DH_p^R(N)$. Let $\bar{a} \in \text{diff}^\ell(M,N)$. Since $\bar{a} \in \text{diff}^\ell(M,N)$, there is a relation $R(x_1, \dots, x_t) \in \mathcal{L}_0$ and a map $h : [\ell] \to [t]$ such that $M \models R(a_{h(1)}, \dots, a_{h(t)})$ and $N \models \neg R(a_{h(1)}, \dots, a_{h(t)})$ or vice versa. If h is not surjective, then some permutation of $C_p(a_{h(1)}, \dots, a_{h(t)})$ is a proper subtuple $\bar{a}'$ of $\bar{a}$ such that $qftp^M(\bar{a}') \neq qftp^N(\bar{a}')$. But this contradicts that $\bar{a} \in \text{diff}^\ell(M,N)$. Thus h is surjective. Let p be the partition of $[t]$ with parts $h^{-1}(\{1\}), \dots, h^{-1}(\{\ell\})$. Since h is surjective, the parts are all nonempty, so $||p|| = \ell$. Then by definition, $C_p(a_{h(1)}, \dots, a_{h(t)}) \in DH_p^R(M)\Delta DH_p^R(N)$. Define $f_\ell(\bar{a}) = C_p(a_{h(1)}, \dots, a_{h(t)})$. Observe that $\cup C_p(a_{h(1)}, \dots, a_{h(t)}) = \cup\bar{a}$ implies

$$f_\ell^{-1}(f_\ell(\bar{a})) \subseteq \{\bar{b} \in W^\ell : \cup\bar{b} = \cup\bar{a}\},$$

so $|f_\ell^{-1}(f_\ell(\bar{a}))| \leq \ell!$. Thus $f_\ell : \text{diff}^\ell(M,N) \to \bigcup_{(R,p)\in Index, ||p||=\ell} DH_p^R(M)\Delta DH_p^R(N)$ and

$$\text{for all } \bar{c} \in \bigcup_{(R,p)\in Index, ||p||=\ell} DH_p^R(M)\Delta DH_p^R(N), \quad |f_\ell^{-1}(\bar{c})| \leq \ell!. \tag{18}$$

Define a map $\beta : \text{diff}(M,N) \to \bigcup_{(R,p)\in Index} DH^R_p(M)\Delta DH^R_p(N)$ as follows. Given $A \in \text{diff}(M,N)$, apply Ψ to obtain $\Psi(A) \in \text{diff}^\ell(M,N)$ for some $\ell \in [r]$. Then define

$$\beta(\bar{a}) := f_\ell(\Psi(\bar{a})) \in \bigcup_{(R,p)\in Index, ||p||=\ell} DH^R_p(M)\Delta DH^R_p(N).$$

Suppose $\bar{c} \in \bigcup_{(R,p)\in Index} DH^R_p(M)\Delta DH^R_p(N)$ and $\ell := |\bar{c}|$. Then $\bar{c} \in DH^R_p(M)\Delta DH^R_p(N)$ for some $(R,p) \in Index$ with $||p|| = \ell$. By definition of β, $\beta^{-1}(\bar{c}) = \Psi^{-1}(f_\ell^{-1}(\bar{c}))$. Combining (18) and the fact that $|\Psi^{-1}(\bar{a})| \leq \binom{n-\ell}{r-\ell}$ for all $\bar{a} \in \text{diff}^\ell(M,N)$, we have that

$$|\beta^{-1}(\bar{c})| = |\Psi^{-1}(f_\ell^{-1}(\bar{c}))| \leq \binom{n-\ell}{r-\ell}\ell!.$$

This shows that $|\text{diff}(M,N)| \leq \sum_{\ell\in[r]}\sum_{(R,p)\in Index,||p||=\ell}\binom{n-\ell}{r-\ell}\ell!|DH^R_p(M)\Delta DH^R_p(N)|$. Dividing both sides of this by $\binom{n}{r}$, we obtain the following.

$$\text{dist}(M,N) \leq \sum_{\ell\in[r]}\sum_{(R,p)\in Index,||p||=\ell}\frac{\binom{n-\ell}{r-\ell}\ell!}{\binom{n}{r}}|DH^R_p(M)\Delta DH^R_p(N)|. \tag{19}$$

Note that for all $1 \leq \ell < r$,

$$\frac{\binom{n-\ell}{r-\ell}\ell!}{\binom{n}{r}} = \frac{(n-\ell)!}{n!}\frac{\ell!r!}{(r-\ell)!} \leq \frac{2^\ell}{n^\ell}\frac{\ell!r!}{(r-\ell)!} < \frac{(r!)^2 2^r}{n^\ell},$$

where the first inequality is by (17) and the last is because $\ell < r$. If $\ell = r$, then

$$\frac{\binom{n-\ell}{r-\ell}\ell!}{\binom{n}{r}} = \frac{r!}{\binom{n}{r}} = \frac{(r!)^2(n-r)!}{n!} \leq \frac{(r!)^2 2^r}{n^r},$$

where the inequality is by (17). Thus for all $\ell \in [r]$, $\frac{\binom{n-\ell}{r-\ell}\ell!}{\binom{n}{r}} \leq \frac{(r!)^2 2^r}{n^\ell}$. Combining this with (19) yields

$$\text{dist}(M,N) \leq (r!)^2 2^r \sum_{\ell\in[r]}\sum_{(R,p)\in Index,||p||=\ell}\frac{|DH^R_p(M)\Delta DH^R_p(N)|}{n^\ell} = (r!)^2 2^r d(M,N).$$

□

Proof of Theorem 9. Fix $\delta > 0$. Let $\delta' = \frac{\delta}{(r!)^2 2^r}$ and choose $K = K(\delta')$ and $\epsilon = \epsilon(\delta')$ by applying Theorem 12 to δ' and $\mathcal{A}$. Suppose n is sufficiently large so that Theorem 12 applies to structures of size n. Suppose M is an $\mathcal{L}_0$-structure of size n such that $\text{prob}(\mathcal{A}(K), M) < \epsilon$. Then Theorem 12 implies there is an $\mathcal{L}_0$-structure M' with $dom(M') = dom(M)$ such that $d(M', M) < \delta'$ and $\text{prob}(\mathcal{A}, M') = 0$. Combining this with Lemma 11 yields

$$\text{dist}(M', M) \leq (r!)^2 2^r d(M', M) < (r!)^2 2^r \delta' = \delta.$$

□

References

[1] V. E. Alekseev, *Hereditary classes and coding of graphs*, Problemy Kibernet. (1982), no. 39, 151–164.
[2] ———, *Range of values of entropy of hereditary classes of graphs*, Diskret. Mat. **4** (1992), no. 2, 148–157.
[3] Noga Alon, József Balogh, Béla Bollobás, and Robert Morris, *The structure of almost all graphs in a hereditary property*, J. Combin. Theory Ser. B **101** (2011), no. 2, 85–110.
[4] Noga Alon and Raphael Yuster, *The number of orientations having no fixed tournament*, Combinatorica **26** (2006), no. 1, 1–16.
[5] Ashwini Aroskar and James Cummings, *Limits, regularity and removal for finite structures*, arXiv:1412.808v1 [math.LO], 2014.

[6] József Balogh, Béla Bollobás, and Robert Morris, *Hereditary properties of combinatorial structures: posets and oriented graphs*, J. Graph Theory **56** (2007), no. 4, 311–332.

[7] ———, *Hereditary properties of tournaments*, Electron. J. Combin. **14** (2007), no. 1, Research Paper 60, 25 pp. (electronic).

[8] József Balogh, Béla Bollobás, and Miklós Simonovits, *The number of graphs without forbidden subgraphs*, J. Combin. Theory Ser. B **91** (2004), no. 1, 1–24.

[9] József Balogh, Béla Bollobás, and David Weinreich, *The speed of hereditary properties of graphs*, J. Combin. Theory Ser. B **79** (2000), no. 2, 131–156.

[10] ———, *The penultimate rate of growth for graph properties*, European J. Combin. **22** (2001), no. 3, 277–289.

[11] Jòzsef Balogh, Béla Bollobás, and David Weinreich, *Measures on monotone properties of graphs*, Discrete Appl. Math. **116** (2002), no. 1-2, 17–36.

[12] József Balogh, Neal Bushaw, Maurício Collares Neto, Hong Liu, Robert Morris, and Maryam Sharifzadeh, *The typical structure of graphs with no large cliques*, to appear, Combinatorica, arXiv: 1406.6961v2 [math.CO], 2015.

[13] József Balogh, Hong Liu, Šárka Petříčková, and Maryam Sharifzadeh, *The typical structure of maximal triangle-free graphs*, to appear, Forum of Mathematics, Sigma, arXiv: 1501.02849 [math.CO], 2015.

[14] József Balogh, Robert Morris, and Wojciech Samotij, *Independent sets in hypergraphs*, J. Amer. Math. Soc. **28** (2015), no. 3, 669–709.

[15] József Balogh and Dhruv Mubayi, *Almost all triple systems with independent neighborhoods are semi-bipartite*, J. Combin. Theory Ser. A **118** (2011), no. 4, 1494–1518.

[16] ———, *Almost all triangle-free triple systems are tripartite*, Combinatorica **32** (2012), no. 2, 143–169.

[17] József Balogh and Adam Zsolt Wagner, *Further applications of the container method*, pp. 191–213, Springer International Publishing, Cham, 2016.

[18] Béla Bollobás and Andrew Thomason, *Projections of bodies and hereditary properties of hypergraphs*, Bull. London Math. Soc. **27** (1995), no. 5, 417–424.

[19] ———, *Hereditary and monotone properties of graphs*, The mathematics of Paul Erdős, II, Algorithms Combin., vol. 14, Springer, Berlin, 1997, pp. 70–78.

[20] Ryan Dotson and Brendan Nagle, *Hereditary properties of hypergraphs*, J. Combin. Theory Ser. B **99** (2009), no. 2, 460–473.

[21] V. Falgas-Ravry, O'Connell K., Strömberg J., and Uzzell A., *Multicolour containers and the entropy of decorated graph limits*, arXiv:1607.08152 [math.CO], 2016.

[22] Christoph Hundack, Hans Jürgen Prömel, and Angelika Steger, *Extremal graph problems for graphs with a color-critical vertex*, Combin. Probab. Comput. **2** (1993), no. 4, 465–477.

[23] Yoshiyasu Ishigami, *The number of hypergraphs and colored hypergraphs with hereditary properties*, arXiv:0712.0425 [math.CO], 2007.

[24] Jaehoon Kim, Daniela Kühn, Deryk Osthus, and Timothy Townsend, *Forbidding induced even cycles in a graph: typical structure and counting*, arXiv:1507.04944 [math.CO], 2015.

[25] Yoshiharu Kohayakawa, Brendan Nagle, and Vojtěch Rödl, *Hereditary properties of triple systems*, Combin. Probab. Comput. **12** (2003), no. 2, 155–189.

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

[27] Vera Koponen, *Asymptotic probabilities of extension properties and random l-colourable structures*, Ann. Pure Appl. Logic **163** (2012), no. 4, 391–438.

[28] Daniela Kühn, Deryk Osthus, Timothy Townsend, and Yi Zhao, *On the structure of oriented graphs and digraphs with forbidden tournaments or cycles*, arxiv:1404.6178 [math.CO], 2014.

[29] David Marker, *Model theory*, Graduate Texts in Mathematics, vol. 217, Springer-Verlag, New York, 2002, An introduction.

[30] Frank Mousset, Rajko Nenadov, and Angelika Steger, *On the number of graphs without large cliques*, SIAM J. Discrete Math. **28** (2014), no. 4, 1980–1986.

[31] D. Mubayi and C. Terry, *Discrete metric spaces: structure, enumeration, and 0-1 laws*, under review, J. Symbolic Logic, arXiv:1502.01212 [math.CO], 2015.

[32] Brendan Nagle and Vojtěch Rödl, *The asymptotic number of triple systems not containing a fixed one*, Discrete Math. **235** (2001), no. 1-3, 271–290, Combinatorics (Prague, 1998).

[33] Yury Person and Mathias Schacht, *Almost all hypergraphs without Fano planes are bipartite*, Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, PA, 2009, pp. 217–226.

[34] H. J. Prömel and A. Steger, *Excluding induced subgraphs. III. A general asymptotic*, Random Structures Algorithms **3** (1992), no. 1, 19–31.

[35] ———, *Excluding induced subgraphs. II. Extremal graphs*, Discrete Appl. Math. **44** (1993), no. 1-3, 283–294.

[36] Hans Jürgen Prömel and Angelika Steger, *Excluding induced subgraphs: quadrilaterals*, Random Structures Algorithms **2** (1991), no. 1, 55–71.

[37] David Saxton and Andrew Thomason, *Hypergraph containers*, Inventiones mathematicae **201** (2015), no. 3, 925–992 (English).

[38] Edward R. Scheinerman and Jennifer Zito, *On the size of hereditary classes of graphs*, J. Combin. Theory Ser. B **61** (1994), no. 1, 16–39.

[39] S. Shelah, *Classification theory: and the number of non-isomorphic models*, Studies in Logic and the Foundations of Mathematics, Elsevier Science, 1990.

[40] Pierre Simon, *A guide to nip theories*, Cambridge University Press, 2015.

[41] Caroline Terry, *Structure and enumeration theorems for hereditary properties in finite relational languages*, arXiv: 1607.04902 [math.LO], 2016.

[42] Peter Winkler, *Random structures and zero-one laws*, Finite and infinite combinatorics in sets and logic (Banff, AB, 1991), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., vol. 411, Kluwer Acad. Publ., Dordrecht, 1993, pp. 399–420.