THE ENTROPY FUNCTION OF AN INVARIANT MEASURE

NATHANAEL ACKERMAN, CAMERON FREER, AND REHANA PATEL

ABSTRACT. Given a countable relational language L, we consider probability measures on the space of L-structures with underlying set $\mathbb{N}$ that are invariant under the logic action. We study the growth rate of the entropy function of such a measure, defined to be the function sending $n \in \mathbb{N}$ to the entropy of the measure induced by restrictions to L-structures on $\{0, \ldots, n-1\}$. When L has finitely many relation symbols, all of arity $k \geq 1$, and the measure has a property called non-redundance, we show that the entropy function is of the form $Cn^k + o(n^k)$, generalizing a result of Aldous and Janson. When $k \geq 2$, we show that there are invariant measures whose entropy functions grow arbitrarily fast in $o(n^k)$, extending a result of Hatami–Norine. For possibly infinite languages L, we give an explicit upper bound on the entropy functions of non-redundant invariant measures in terms of the number of relation symbols in L of each arity; this implies that finite-valued entropy functions can grow arbitrarily fast.

1. INTRODUCTION

The information-theoretic notion of entropy captures the idea of the typical "uncertainty" in a sample of a given probability measure. An important class of probability measures in model theory, combinatorics, and probability theory are the $\mathrm{Sym}(\mathbb{N})$-invariant measures on the space of structures, in a countable language, with underlying set $\mathbb{N}$. We call such measures *invariant measures*, and consider the entropy of such measures.

When the support of a probability measure is uncountable, as is usually the case for invariant measures, the entropy of the measure is infinite. But any such measure μ can be approximated by its projections to the spaces of

structures on initial segments of $\mathbb{N}$. We define the *entropy function* of μ to be the function from $\mathbb{N}$ to $\mathbb{R} \cup \{\infty\}$ taking n to the entropy of the measure induced by μ on structures with underlying set $\{0, \ldots, n-1\}$. In this paper, we study the growth of entropy functions.

The growth of the entropy function of an invariant measure has been studied by Aldous [Ald85, Chapter 15] in the setting of exchangeable arrays and by Janson [Jan13, §10 and §D.2] and Hatami–Norine [HN13, Theorem 1.1] in the case where the measure is concentrated on the space of graphs. The leading coefficient of the entropy function has been used to study large deviations for Erdős–Rényi random graphs and exponential random graphs in [CV11] and [CD13], and phase transitions in exponential random graphs in [RS13]. For additional results on the entropy functions of invariant measures concentrated on the space of graphs, see [HJS18].

Further, ergodic invariant measures are the focus of a project that treats them model-theoretically as a notion of "symmetric probabilistic structures"; see [AFP16], [AFNP16], [AFKwP17], [AFP17], and [AFKrP17]. The entropy function of such a measure is one gauge of its complexity.

In this paper, we primarily consider invariant measures, for a relational language, that are *non-redundant*, namely those that concentrate on structures in which a relation can hold of a tuple only when the tuple consists of distinct elements. In §1.2 we develop machinery for quantifier-free interdefinitions that allows us to show, in Proposition 1.21, that in the setting of countable languages, every entropy function of an invariant measure is the entropy function of some non-redundant invariant measure for a relational language.

In Section 2, we study the case of invariant measures for finite relational languages L with all relations having the same arity $k \geq 1$. The Aldous–Hoover–Kallenberg theorem provides a representation, which we call an *extended L-hypergraphon*, of a non-redundant invariant measure for L. Our main technical result in this section, Theorem 2.7, establishes that the entropy function of the invariant measure corresponding to an extended L-hypergraphon W grows like $Cn^k + o(n^k)$, where the constant C can be calculated from W. When the invariant measure is concentrated on the space of graphs, Theorem 2.7 reduces to a result first observed by Aldous in [Ald85, Chapter 15], and also by Janson in [Jan13, Theorem D.5]. The proof of Theorem 2.7 follows closely that of Janson.

In Section 3 we consider invariant measures that arise via sampling from a Borel k-uniform hypergraph for $k \geq 2$. The entropy functions of such measures grow like $o(n^k)$, as we show in Lemma 3.2. We prove, in Theorem 3.5, that for every function $\gamma(n) \in o(n^k)$ there is an invariant measure sampled from a Borel k-uniform hypergraph whose entropy function grows faster than $\gamma(n)$. Our theorem generalizes a result of Hatami–Norine [HN13, Theorem 1.1], who

prove it in the case where $k = 2$ (hence where the Borel hypergraph is simply a Borel graph). Hatami–Norine's proof analyzes the random graph obtained by subsampling from an appropriate "blow up" of the Rado graph, i.e., the unique countable homogeneous-universal graph. Our proof takes an analogous path, using the unique countable homogeneous-universal k-uniform hypergraph.

Finally, in Section 4 we consider invariant measures for countable languages that may be infinite. We provide an upper bound on the entropy functions of non-redundant invariant measures for a relational language, in terms of the number of relations of each arity. We do so by calculating the entropy function of a particular non-redundant invariant measure whose entropy function is maximal among those in a given language. This calculation also demonstrates that whereas the growth of the entropy function of an invariant measure for a *finite* language is polynomial, there are $\mathbb{R}$-valued entropy functions that grow arbitrarily fast.

1.1. **Preliminaries.** *In this subsection, let L be a countable language and let $\mathfrak{n} \in \mathbb{N} \cup \{\mathbb{N}\}$.*

For $n \in \mathbb{N}$, write $[n] := \{0, \ldots, n-1\}$, and $[\mathbb{N}] := \mathbb{N}$. Let $\mathrm{Sym}(\mathfrak{n})$ denote the collection of permutations of $[\mathfrak{n}]$.

For $k \in \mathbb{N}$, define $\mathfrak{P}_{<k}(\mathfrak{n}) := \{Y \subseteq [\mathfrak{n}] : |Y| < k\}$ and $\mathfrak{P}_k(\mathfrak{n}) := \{Y \subseteq [\mathfrak{n}] : |Y| = k\}$, and let $\mathfrak{P}_{\leq k}(\mathfrak{n}) := \mathfrak{P}_{<k}(\mathfrak{n}) \cup \mathfrak{P}_k(\mathfrak{n})$. We order each of these sets using *shortlex* order, i.e., ordered by size, with sets of the same size ordered lexicographically.

We write $\mathcal{L}_{\omega,\omega}(L)$ to denote the collection of first-order L-formulas. An *L-theory* is a collection of first-order L-sentences. A theory is Π_1 when every sentence is quantifier-free or of the form $(\forall \overline{x})\varphi(\overline{x})$ where $\overline{x}$ is a tuple of variables and φ is quantifier-free. The *maximum arity* of L is the maximum arity, when it exists, of a relation symbol or function symbol in L, where we consider constant symbols to be function symbols of arity 0.

Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose $(D, \mathcal{D})$ is a measurable space. A *D-valued random variable Z*, also called a *random element in D*, is an $(\mathcal{F}, \mathcal{D})$-measurable function $Z \colon \Omega \to D$. The *distribution* of Z is the probability measure $\mathbb{P} \circ Z^{-1}$. Given an event $B \in \mathcal{D}$, we say that B holds *almost surely* when $\mathbb{P}(B) = 1$, and abbreviate this *a.s.* Typically B will be specified indirectly via some property of random variables; for example, we say that two random variables are equal a.s. when the subset of Ω on which they are equal has full measure. For a topological space S, let $\mathcal{P}(S)$ denote the space of Borel probability measures on S, with σ-algebra given by the weak topology. We use the symbol $\bigwedge$ for conjunctions of probabilistic events, as well as for conjunctions of logical formulas. See [Kal02] for further background and notation from probability theory.

We write λ to denote the uniform (Lebesgue) measure on $[0, 1]$ and on finite powers of $[0, 1]$. Suppose a variable x takes values in some finite power of $[0, 1]$. We say that an expression involving x holds *almost everywhere*, abbreviated *a.e.*, when it holds of x on all but a λ-null subset.

Let $\mathrm{Str}_L(\mathfrak{n})$ denote the collection of L-structures that have underlying set $[\mathfrak{n}]$. We will consider $\mathrm{Str}_L(\mathfrak{n})$ as a measure space with the σ-algebra generated by the topology given by basic clopen sets of the form

$$\{\mathcal{M} \in \mathrm{Str}_L(\mathfrak{n}) \,:\, \mathcal{M} \models \varphi(x_0, \ldots, x_{\ell-1})\}$$

when $\varphi \in \mathcal{L}_{\omega,\omega}(L)$ is a quantifier-free formula with ℓ free variables and $x_0, \ldots, x_{\ell-1} \in [\mathfrak{n}]$.

Definition 1.1. *Suppose $\varphi \in \mathcal{L}_{\omega,\omega}(L)$ has ℓ free variables and $r_0, \ldots, r_{\ell-1} \in [\mathfrak{n}]$. Define the **extent** (on $[\mathfrak{n}]$) of $\varphi(r_0, \ldots, r_{\ell-1})$ to be*

$$[\![\varphi(r_0, \ldots, r_{\ell-1})]\!]_{[\mathfrak{n}]} := \{\mathcal{M} \in \mathrm{Str}_L(\mathfrak{n}) \,:\, \mathcal{M} \models \varphi(r_0, \ldots, r_{\ell-1})\}.$$

When $\mathfrak{n} = \mathbb{N}$ we will sometimes omit the subscript $[\mathfrak{n}]$. For an L-theory T, define $[\![T]\!]_{[\mathfrak{n}]} := \bigcap_{\rho \in T} [\![\rho]\!]_{[\mathfrak{n}]}$.

There is a natural action of $\mathrm{Sym}(\mathfrak{n})$ on $\mathrm{Str}_L(\mathfrak{n})$ called **the logic action**, defined as follows: for $\sigma \in \mathrm{Sym}(\mathfrak{n})$ and $\mathcal{M} \in \mathrm{Str}_L(\mathfrak{n})$, let $\sigma \cdot \mathcal{M}$ be the structure $\mathcal{N} \in \mathrm{Str}_L(\mathfrak{n})$ for which

$$R^{\mathcal{N}}(r_0, \ldots, r_{k-1}) \quad \text{if and only if} \quad R^{\mathcal{M}}\big(\sigma^{-1}(r_0), \ldots, \sigma^{-1}(r_{k-1})\big)$$

for all relation symbols $R \in L$ and $r_0, \ldots, r_{k-1} \in \mathfrak{n}$, where k is the arity of R, and similarly with constant and function symbols. Note that the orbit under the logic action of any structure in $\mathrm{Str}_L(\mathfrak{n})$ is its isomorphism class. By Scott's isomorphism theorem, every such orbit is Borel. For more details on the logic action, see [Kec95, §16.C].

We say a probability measure μ on $\mathrm{Str}_L(\mathfrak{n})$ is **invariant** if it is invariant under the logic action of $\mathrm{Sym}(\mathfrak{n})$, i.e., if $\mu(B) = \mu(\sigma \cdot B)$ for every Borel $B \subseteq \mathrm{Str}_L(\mathfrak{n})$ and every $\sigma \in \mathrm{Sym}(\mathfrak{n})$. We call such a probability measure an *invariant measure for L*. An invariant measure μ is **ergodic** if $\mu(X) = 0$ or $\mu(X) = 1$ whenever $\mu(X \triangle \sigma(X)) = 0$ for all $\sigma \in \mathrm{Sym}(\mathbb{N})$. Every ergodic invariant measure on $\mathrm{Str}_L(\mathfrak{n})$ is an *extreme point* in the simplex of invariant measures on $\mathrm{Str}_L(\mathfrak{n})$, and any invariant measure on $\mathrm{Str}_L(\mathfrak{n})$ can be decomposed as a mixture of ergodic ones (see [Kal05, Lemma A1.2 and Theorem A1.3]).

For $n \in \mathbb{N}$, any probability measure μ on $\mathrm{Str}_L(\mathbb{N})$ induces a probability measure μ_n on $\mathrm{Str}_L(n)$ such that for any Borel set $B \subseteq \mathrm{Str}_L(n)$, we have

$$\mu_n(B) := \mu(\{\mathcal{M} \in \mathrm{Str}_L(\mathbb{N}) \,:\, \mathcal{M}|_{[n]} \in B\}),$$

where $\mathcal{M}|_{[n]} \in \mathrm{Str}_L(n)$ denotes the restriction of $\mathcal{M}$ to $[n]$. Further, by the Kolmogorov consistency theorem, μ is uniquely determined by the collection $\langle \mu_n \rangle_{n \in \mathbb{N}}$.

Towards defining the entropy function of an invariant measure, we give the standard definition of the entropy of a probability measure. For this definition, we use the convention that $-0 \log_2(0) = 0$.

Definition 1.2. *Let ν be a probability measure on a standard Borel space S, and let $A := \{s \in S : \nu(\{s\}) > 0\}$ be its (countable) set of atoms. If ν is purely atomic, i.e., $\nu(A) = 1$, then the* **entropy** *of ν is given by*

$$h(\nu) := -\sum_{x \in A} \nu(\{x\}) \log_2(\nu(\{x\})).$$

Otherwise, let

$$h(\nu) := \infty.$$

For any random variable X with distribution ν, define $h(X) := h(\nu)$.

Definition 1.3. *The* **joint entropy** *of a pair of random variables X and Y, written $h(X, Y)$, is defined to be the entropy of the joint distribution of (X, Y). Similarly, $h(\langle X_i \rangle_{i \in I})$ is defined to be the entropy of the joint distribution of the sequence $\langle X_i \rangle_{i \in I}$.*

Definition 1.4. *Given random variables X and Y, the* **conditional entropy** *of X given Y, written $h(X \in \cdot \,|\, Y)$, is defined to be the function*

$$y \mapsto h(\mathbb{P}(X \in \cdot \,|\, Y = y)).$$

We have defined the entropy of a random variable to be the entropy of its distribution. When the random variable takes values in a space of measures, instead of considering the entropy of the random variable directly, we sometimes need the random variable, defined below, that is obtained by taking the entropies of these measures themselves.

Definition 1.5. *Suppose χ is a measure-valued random variable. The* **random entropy** *of χ, written $H(\chi)$, is a random variable defined by*

$$H(\chi)(\varpi) = h(\chi(\varpi)).$$

for $\varpi \in \Omega$.

This notion allows us to define random conditional entropy.

Definition 1.6. *Given random variables X and Y, the* **random conditional entropy** *of X given Y, written $H(X \,|\, Y)$, is defined by*

$$H(X \,|\, Y) = H(\mathbb{E}(X \,|\, Y)),$$

i.e., the random entropy of the random measure $\mathbb{E}(X \,|\, Y)$, the conditional expectation of X given Y.

The following three lemmas are standard facts about entropy, which we will need later.

Lemma 1.7 (see [CT06, Theorem 2.2.1]). *For random variables X and Y, we have*

$$h(X, Y) = h(X) + \mathbb{E}(H(Y \mid X)).$$

Lemma 1.8 (see [CT06, Theorem 2.6.5]). *For random variables X and Y, we have*

$$h(X) \geq \mathbb{E}(H(X \mid Y)),$$

with equality if and only if X and Y are independent.

Lemma 1.9 (see [CT06, Theorem 2.6.6]). *Let $\langle X_i \rangle_{i \in I}$ be a sequence of random variables. Then*

$$h(\langle X_i \rangle_{i \in I}) \leq \sum_{i \in I} h(X_i),$$

with equality if and only if the X_i are independent.

We now define the entropy function of an invariant measure on $\mathrm{Str}_L(\mathbb{N})$.

Definition 1.10. *Let μ be an invariant measure on $\mathrm{Str}_L(\mathbb{N})$. The **entropy function** of μ is defined to be the function $\mathrm{Ent}(\mu)$ from $\mathbb{N}$ to $\mathbb{R} \cup \{\infty\}$ given by*

$$\mathrm{Ent}(\mu)(n) := h(\mu_n).$$

The following basic property of the entropy function is immediate.

Lemma 1.11. *Let μ be an invariant measure on $\mathrm{Str}_L(\mathbb{N})$. Then $\mathrm{Ent}(\mu)$ is a non-decreasing function.*

In this paper we will mainly be interested in invariant measures whose entropy functions take values in $\mathbb{R}$. The following lemma is immediate because in a finite language, there are only finitely many structures of each finite size.

Lemma 1.12. *Suppose L is finite, and let μ be an invariant measure on $\mathrm{Str}_L(\mathbb{N})$. Then $\mathrm{Ent}(\mu)$ is $\mathbb{R}$-valued.*

In Section 4 we will strengthen this lemma by showing that there is a polynomial upper bound of $O(n^k)$ on $\mathrm{Ent}(\mu)(n)$, where k is the maximum arity of L.

The notion of *non-redundance* for an invariant measure, which we introduce next, will be key throughout this paper.

Definition 1.13. *Suppose L is relational. For each $R \in L$ define ϑ_R to be the formula*

$$(\forall x_0, \ldots, x_{k-1})\Big(R(x_0, \ldots, x_{k-1}) \to \bigwedge_{i < j < k} (x_i \neq x_j)\Big),$$

where k is the arity of R. Define the Π_1 L-theory

$$\mathrm{Th}_{\mathrm{nr}}(L) := \{\vartheta_R \,:\, R \in L\}.$$

*An L-structure $\mathcal{M}$ is **non-redundant** when $\mathcal{M} \models \mathrm{Th}_{\mathrm{nr}}(L)$. An invariant measure μ on $\mathrm{Str}_L(\mathbb{N})$ is **non-redundant** when*

$$\mu(\llbracket\mathrm{Th}_{\mathrm{nr}}(L)\rrbracket) = 1.$$

For example, any k-uniform hypergraph is non-redundant, as is any directed graph without self-loops.

The following straightforward lemma provides conditions under which the entropy function takes values in $\mathbb{R}$.

Lemma 1.14. *Suppose L has finitely many relations of any given arity, and let μ be a non-redundant invariant measure on $\mathrm{Str}_L(\mathbb{N})$. Then for every $n \in \mathbb{N}$ there is a finite set $A_n \subseteq \mathrm{Str}_L(n)$ such that μ_n concentrates on A_n. In particular, $\mathrm{Ent}(\mu)(n) \in \mathbb{R}$, and so $\mathrm{Ent}(\mu)$ is $\mathbb{R}$-valued.*

In Section 4 we will consider the case of non-redundant invariant measures for a relational language, and will strengthen this lemma by providing an explicit upper bound on $\mathrm{Ent}(\mu)(n)$.

1.2. **Quantifier-free interdefinitions.** The notion of *quantifier-free inter-definability*, which we define below, is a variant of the standard notion of interdefinability from the setting of $\aleph_0$-categorical theories (see, e.g., [AZ86, §1]). It provides a method for translating invariant measures concentrated on the extent of a given Π_1 theory to invariant measures concentrated on the extent of a target Π_1 theory, in a way that preserves the entropy function. We use this machinery to show, in Proposition 1.21, that every entropy function of an invariant measure already occurs as the entropy function of a non-redundant invariant measure for a relational language.

Throughout this subsection, L_0 and L_1 will be countable languages, sometimes with further restrictions.

Definition 1.15. *Suppose T_0 is an L_0-theory and T_1 is an L_1-theory. A **quantifier-free interdefinition** between T_0 and T_1 is a pair $\Psi = (\Psi_0, \Psi_1)$ of maps*

$$\begin{aligned}\Psi_0 \;&:\; \mathcal{L}_{\omega,\omega}(L_0) \to \mathcal{L}_{\omega,\omega}(L_1) \qquad and\\ \Psi_1 \;&:\; \mathcal{L}_{\omega,\omega}(L_1) \to \mathcal{L}_{\omega,\omega}(L_0)\end{aligned}$$

such that for $j \in \{0,1\}$, the formula $\Psi_j(\eta)$ is quantifier-free whenever $\eta \in \mathcal{L}_{\omega,\omega}(L_{1-j})$ is quantifier-free, and further,

$$
\begin{aligned}
T_{1-j} &\vdash \quad \Psi_j \circ \Psi_{1-j}(\rho) \leftrightarrow \rho, \\
T_{1-j} &\vdash \quad \Psi_j(x = y) \leftrightarrow (x = y), \\
T_{1-j} &\vdash \quad \neg\Psi_j(\chi) \leftrightarrow \Psi_j(\neg\chi), \\
T_{1-j} &\vdash \quad \Psi_j(\chi \wedge \varphi) \leftrightarrow \big(\Psi_j(\chi) \wedge \Psi_j(\varphi)\big), \qquad \text{and} \\
T_{1-j} &\vdash \quad (\exists x)\Psi_j(\psi(x)) \leftrightarrow \Psi_j\big((\exists x)\psi(x)\big)
\end{aligned}
$$

for all $\rho \in \mathcal{L}_{\omega,\omega}(L_{1-j})$ and $\chi, \varphi, \psi(x) \in \mathcal{L}_{\omega,\omega}(L_j)$, and such that the free variables of $\Psi_j(\upsilon)$ are the same as those of υ for every $\upsilon \in \mathcal{L}_{\omega,\omega}(L_j)$.

For $\mathfrak{n} \in \mathbb{N} \cup \{\mathbb{N}\}$, the interdefinition Ψ induces maps $\Psi_{j,\mathfrak{n}}^* \colon [\![T_j]\!]_{[\mathfrak{n}]} \to [\![T_{1-j}]\!]_{[\mathfrak{n}]}$ for $j \in \{0,1\}$ satisfying, for any structure $\mathcal{N} \models T_j$, tuple $\boldsymbol{m} \in \mathfrak{n}$, and formula $\varphi \in \mathcal{L}_{\omega,\omega}(L_{1-j})$,

$$\Psi_{j,\mathfrak{n}}^*(\mathcal{N}) \models \varphi(\boldsymbol{m}) \qquad \text{if and only if} \qquad \mathcal{N} \models \Psi_{1-j}(\varphi)(\boldsymbol{m}).$$

It is immediate that each $\Psi_{j,\mathfrak{n}}^*$ is a bijection.

Quantifier-free interdefinitions between Π_1 theories preserve entropy functions, as we now show.

Lemma 1.16. *Let $\Psi = (\Psi_0, \Psi_1)$ be a quantifier-free interdefinition between a Π_1 L_0-theory T_0 and a Π_1 L_1-theory T_1. Suppose μ is an invariant measure on $\mathrm{Str}_{L_0}(\mathbb{N})$ concentrated on $[\![T_0]\!]$ and let ν be the pushforward of μ along $\Psi_{0,\mathbb{N}}^*$. Then ν is an invariant measure on $\mathrm{Str}_{L_1}(\mathbb{N})$ concentrated on $[\![T_1]\!]$, and $\mathrm{Ent}(\nu) = \mathrm{Ent}(\mu)$.*

Proof. It is immediate that ν is an invariant measure on $\mathrm{Str}_{L_1}(\mathbb{N})$ concentrated on $[\![T_1]\!]$, by the definition of pushforward and the fact that $\Psi_{0,\mathbb{N}}^*$ is a bijection between $[\![T_0]\!]$ and $[\![T_1]\!]$.

We will next show, for each $n \in \mathbb{N}$, that there is a measure-preserving bijection between the atoms of μ_n and the atoms of ν_n. By Definition 1.2, this will establish that $h(\mu_n) = h(\nu_n)$ for each $n \in \mathbb{N}$, and so $\mathrm{Ent}(\mu) = \mathrm{Ent}(\nu)$.

Let $n \in \mathbb{N}$. If $\mathcal{M} \in [\![T_0]\!]_{\mathbb{N}}$ then $\mathcal{M}|_{[n]} \in [\![T_0]\!]_{[n]}$, because μ concentrates on $[\![T_0]\!]_{\mathbb{N}}$ and T_0 is Π_1. Hence the measure μ_n concentrates on $[\![T_0]\!]_{[n]}$. Similarly, ν_n concentrates on $[\![T_1]\!]_{[n]}$. Therefore every atom of μ_n is in $[\![T_0]\!]_{[n]}$ and every atom of ν_n is in $[\![T_1]\!]_{[n]}$.

Write ν_n^* for the pushforward of μ_n along $\Psi_{0,[n]}^*$. It is clear that $\Psi_{0,[n]}^*$ is a measure-preserving bijection between the atoms of μ_n and the atoms of ν_n^*. Let $\mathcal{B}$ be an atom of ν_n; we have $\mathcal{B} \in [\![T_1]\!]_{[n]}$ because $\nu_n(\mathcal{B}) > 0$. We will show that $\nu_n(\mathcal{B}) = \nu_n^*(\mathcal{B})$.

Let $\mathcal{A} = \Psi^*_{1,[n]}(\mathcal{B})$. For $\mathcal{N} \in [\![T_1]\!]_{\mathbb{N}}$, by Definition 1.15 we have $\mathcal{N}|_{[n]} = \mathcal{B}$ if and only if $\Psi^*_{1,\mathbb{N}}(\mathcal{N})|_{[n]} = \mathcal{A}$. By this fact, the definitions of μ_n, ν_n, and pushforward, the surjectivity of $\Psi^*_{1,\mathbb{N}}$, and the fact that μ and ν are concentrated on $[\![T_0]\!]_{\mathbb{N}}$ and $[\![T_1]\!]_{\mathbb{N}}$ respectively, we have

$$
\begin{aligned}
\nu_n(\mathcal{B}) &= \nu(\{\mathcal{N} \in [\![T_1]\!]_{\mathbb{N}} : \mathcal{N}|_{[n]} = \mathcal{B}\}) \\
&= \nu(\{\mathcal{N} \in [\![T_1]\!]_{\mathbb{N}} : \Psi^*_{1,\mathbb{N}}(\mathcal{N})|_{[n]} = \mathcal{A}\}) \\
&= \mu(\{\mathcal{M} \in [\![T_0]\!]_{\mathbb{N}} : \mathcal{M}|_{[n]} = \mathcal{A}\}) \\
&= \mu_n(\mathcal{A}).
\end{aligned}
$$

But $\mu_n(\mathcal{A}) = \nu^*_n(\mathcal{B})$ by the definition of pushforward. Hence $\nu_n(\mathcal{B}) = \nu^*_n(\mathcal{B})$, as desired. $\qquad\square$

Non-redundant invariant measures are defined only for relational languages. Hence towards showing, in Proposition 1.21, that every entropy function is the entropy function of a non-redundant invariant measure, we need a quantifier-free interdefinition between the empty theory in an arbitrary language and a particular Π_1 theory in a relational language. There is a standard such interdefinition that maps every function symbol to a relation symbol representing the graph of the function, but it entails an increase in arity, as each function symbol of arity ℓ is mapped to a relation symbol of arity $\ell + 1$.

While this standard interdefinition would suffice for our purposes in Proposition 1.21, here we provide a more parsimonious interdefinition, to highlight a connection with what is known about invariant measures for languages containing function symbols. Namely, using results from [AFP17], in Lemmas 1.18 and 1.19 we show how to avoid increasing the arity of symbols, by providing a quantifier-free interdefinition that replaces each function symbol of arity ℓ with finitely many relation symbols, each of arity ℓ.

As we will see in Lemma 1.18, every invariant measure is concentrated on structures in which every function is a "selector", sometimes called a "choice function", i.e., a function for which the output is always one of the inputs. For example, the only unary selector is the identity function. We consider constant symbols to be 0-ary function symbols; observe that no constant is a selector.

Definition 1.17. *Let $f \in L_0$ be a function symbol, and let ℓ be the arity of f. Define θ_f to be the sentence*

$$
(\forall x_0, \ldots, x_{\ell-1}) \left(\bigvee_{i \in [\ell]} f(x_0, \ldots, x_{\ell-1}) = x_i \right),
$$

asserting that f is a selector, *and define the Π_1 L_0-theory*

$$
\mathrm{Th}_{\mathrm{sel}}(L_0) := \{\theta_f : f \in L_0 \text{ is a function symbol}\}.
$$

Lemma 1.18. *If μ is an invariant measure on $\mathrm{Str}_{L_0}(\mathbb{N})$ then $\mu(\llbracket \mathrm{Th}_{\mathrm{sel}}(L_0) \rrbracket) = 1$.*

Proof. Let ν be an ergodic invariant measure on $\mathrm{Str}_{L_0}(\mathbb{N})$. By [AFP17, Lemma 2.4],

$$\mathrm{Th}(\nu) := \{\varphi \in \mathcal{L}_{\omega,\omega}(L) \,:\, \nu(\llbracket \varphi \rrbracket) = 1\},$$

is a complete deductively-closed first-order theory. For a function symbol $f \in L_0$, if $\neg\theta_f \in \mathrm{Th}(\nu)$ then $\mathrm{Th}(\nu)$ has non-trivial definable closure, contradicting [AFP17, Proposition 6.1], and so we must have $\theta_f \in \mathrm{Th}(\nu)$.

Therefore $\nu(\llbracket \mathrm{Th}_{\mathrm{sel}}(L_0) \rrbracket) = 1$. Because μ is a mixture of ergodic invariant measures ν, we also have $\mu(\llbracket \mathrm{Th}_{\mathrm{sel}}(L_0) \rrbracket) = 1$. $\qquad\square$

We now provide the desired quantifier-free interdefinition.

Lemma 1.19. *Suppose L_1 is the relational language consisting of*

- *each relation symbol in L_0, along with*
- *a relation symbol $E_{f,i}$ for each function symbol $f \in L_0$ and $i \in [\ell]$, where ℓ is the arity of f.*

Let T be the Π_1 L_1-theory consisting of the sentences

$$(\forall x_0, \ldots, x_{\ell-1}) \bigvee_{i \in [\ell]} E_{f,i}(x_0, \ldots, x_{\ell-1}),$$

and

$$(\forall x_0, \ldots, x_{\ell-1}) \bigwedge_{i < j \in [\ell]} \neg\big(E_{f,i}(x_0, \ldots, x_{\ell-1}) \wedge E_{f,j}(x_0, \ldots, x_{\ell-1})\big),$$

for each function symbol $f \in L_0$, where ℓ is the arity of f.

Then there is a quantifier-free interdefinition (Ψ_0, Ψ_1) between $\mathrm{Th}_{\mathrm{sel}}(L_0)$ and T.

Proof. For each relation symbol $R \in L_0$, let $\Psi_1(R(x_0, \ldots, x_{k-1}))$ be the formula $R(x_0, \ldots, x_{k-1})$, where k is the arity of R. For each function symbol $f \in L_0$ and $i \in [\ell]$, let $\Psi_1(E_{f,i}(x_0, \ldots, x_{\ell-1}))$ be the formula $f(x_0, \ldots, x_{\ell-1}) = x_i$, where ℓ is the arity of f. It is easy to define the analogous map Ψ_0 and check that (Ψ_0, Ψ_1) is a quantifier-free interdefinition between $\mathrm{Th}_{\mathrm{sel}}(L_0)$ and T. $\quad\square$

Next we provide a quantifier-free interdefinition which, combined with the previous results, will allow us to prove Proposition 1.21.

Lemma 1.20. *Suppose L_0 is relational. Then there is a countable relational language L_1 and a quantifier-free interdefinition (Ψ_0, Ψ_1) between the empty L_0-theory and the Π_1 L_1-theory $\mathrm{Th}_{\mathrm{nr}}(L_1)$.*

Proof. For each relation symbol $R \in L_0$ and equivalence relation E on $[k]$, where $k \in \mathbb{N}$ is the arity of R, we define the following. Let ℓ be the number of E-equivalence classes. Let $f_E \colon [k] \to [k]$ send each $i \in [k]$ to the least element of its E-equivalence class, and let $\langle y_0^E, \ldots, y_{\ell-1}^E \rangle$ be the increasing enumeration of the image of f_E. Let R_E be an ℓ-ary relation symbol and let

$$\phi_E(x_0, \ldots, x_{k-1}) := R_E(x_{y_0^E}, \ldots, x_{y_{\ell-1}^E}) \wedge \bigwedge_{\substack{i,j \in [k] \\ iEj}} x_i = x_j \wedge \bigwedge_{\substack{i,j \in [k] \\ \neg(iEj)}} x_i \neq x_j.$$

Define L_1 to be the collection of all such symbols R_E. Recall the Π_1 theory $\mathrm{Th}_{\mathrm{nr}}(L_1)$, as defined in Definition 1.13.

Define the map Ψ_0 on atomic L_0-formulas by

$$\Psi_0\big(R(x_0, \ldots, x_{k-1})\big) := \bigvee_{\substack{E \text{ is an equivalence} \\ \text{relation on } [k]}} \phi_E(x_0, \ldots, x_{k-1}),$$

for each $R \in L_0$, where k is the arity of R. Define the map Ψ_1 on atomic L_1-formulas by

$$\Psi_1\big(R_E(x_0, \ldots, x_{\ell-1})\big) := R(x_{f_E(0)}, \ldots, x_{f_E(k-1)})$$

for each $R_E \in L_1$, where ℓ is the arity of R_E. Extend Ψ_0 and Ψ_1 to all formulas in $\mathcal{L}_{\omega,\omega}(L_0)$ and $\mathcal{L}_{\omega,\omega}(L_1)$, respectively, in the natural way.

Observe that for any $\mathcal{M} \in \mathrm{Str}_{L_0}(\mathbb{N})$, the L_1-structure $\Psi_{0,\mathbb{N}}^*(\mathcal{M})$ is non-redundant, and conversely, every non-redundant structure in $\mathrm{Str}_{L_1}(\mathbb{N})$ is in the image of $\Psi_{0,\mathbb{N}}^*$. One can check that (Ψ_0, Ψ_1) is a quantifier-free interdefinition between the empty L_0-theory and $\mathrm{Th}_{\mathrm{nr}}(L_1)$. $\qquad\square$

We can now show that every entropy function is the entropy function of some non-redundant invariant measure.

Proposition 1.21. *Let μ be an invariant measure on $\mathrm{Str}_{L_0}(\mathbb{N})$. Then there is a countable relational language L_1 and a non-redundant invariant measure ν on $\mathrm{Str}_{L_1}(\mathbb{N})$ such that $\mathrm{Ent}(\mu) = \mathrm{Ent}(\nu)$. Further, if L_0 is finite and of maximum arity k, then so is L_1.*

Proof. Recall the Π_1 L_0-theory $\mathrm{Th}_{\mathrm{sel}}(L_0)$ from Definition 1.17. First observe by Lemma 1.18 that $\mu([\![\mathrm{Th}_{\mathrm{sel}}(L_0)]\!]) = 1$. By Lemma 1.19 there is a quantifier-free interdefinition (Ψ_0, Ψ_1) between $\mathrm{Th}_{\mathrm{sel}}(L_0)$ and some Π_1 theory T in a countable relational language L'. When L_0 is finite and of maximum arity k, so is L'.

By Lemma 1.20 there is a countable relational language L_1, and a quantifier-free interdefinition $\Theta = (\Theta_0, \Theta_1)$ between the empty L'-theory and the Π_1 L_1-theory $\mathrm{Th}_{\mathrm{nr}}$. It is easy to see from Definition 1.15 that Θ is also a quantifier-free interdefinition between T and its image $\Theta_0(T)$, which is a Π_1 L_1-theory. Again, when L' is finite and of maximum arity k, so is L_1.

By two applications of Lemma 1.16, there is an invariant measure ν on $\mathrm{Str}_{L_1}(\mathbb{N})$ that is concentrated on $[\![\Theta_0(T)]\!]$ and that satisfies $\mathrm{Ent}(\mu) = \mathrm{Ent}(\nu)$. Finally, ν is non-redundant because $\Theta_0(T)$ contains $\mathrm{Th}_{\mathrm{nr}}$. $\qquad\square$

2. Invariant measures arising from extended L-hypergraphons

In this section we study the growth of entropy functions of non-redundant invariant measures for a finite relational language L whose relation symbols all have the same arity $k \geq 1$. By a variant of the Aldous–Hoover–Kallenberg theorem, these invariant measures are precisely the ones that arise as the distribution of a certain random L-structure $G(\mathbb{N}, W)$, where W is a type of measurable function called an *extended L-hypergraphon*. In Theorem 2.7 we express the growth of the entropy function of $G(\mathbb{N}, W)$ in terms of W. This result generalizes a theorem of Aldous [Ald85, Chapter 15] and Janson [Jan13, Theorem D.5], and our argument mirrors that of Janson. As an immediate consequence of Theorem 2.7 and the Aldous–Hoover–Kallenberg theorem, we see in Corollary 2.8 that the entropy function of any non-redundant invariant measure for L is $O(n^k)$.

For the rest of this section, fix $k \geq 1$ and let L be a finite relational language all of whose relations have arity k.

We will define an extended L-hypergraphon to be a probability kernel from $[0,1]^{\mathfrak{P}_{<k}(k)}$ to a space of quantifier-free k-types, satisfying a specific coherence condition.

Definition 2.1. *A **complete non-redundant quantifier-free L-type** with free variables $x_0, \ldots, x_{k-1}$ is a maximal consistent collection of atomic formulas or negations of atomic formulas containing $\{x_i \neq x_j : i < j < k\}$, and whose free variables are all contained in $\{x_0, \ldots, x_{k-1}\}$.*

Let $\mathfrak{qf}_L^{\mathrm{nr}}$ be the space of complete non-redundant quantifier-free L-types with free variables $x_0, \ldots, x_{k-1}$, where a subbasic clopen set consists of complete non-redundant quantifier-free L-types containing a given atomic formula or negated atomic formula all of whose free variables are contained in $\{x_0, \ldots, x_{k-1}\}$.

Note that every complete non-redundant quantifier-free L-type in the above sense implies a complete quantifier-free $\mathcal{L}_{\omega_1, \omega}(L)$-type.

In the next definition, we introduce the notion of an extended L-hypergraphon. As we describe below, this generalizes the standard notion of a hypergraphon [Lov12, §23.3], which is a higher-arity version of a graphon.

The definition of extended L-hypergraphon involves a coherence condition, specified in terms of an action of $\mathrm{Sym}(k)$. In Definition 2.5, we describe how an extended L-hypergraphon gives rise to a distribution on L-structures by determining, for each k-tuple of elements, the distribution on its quantifier-free

type. The coherence condition ensures that the order in which the k-tuple is specified does not affect the resulting distribution on its quantifier-free type.

Consider the action of $\mathrm{Sym}(k)$ on $\mathfrak{qf}_L^{\mathrm{nr}}$ given by

$$\sigma \cdot u = \{\varphi(x_{\sigma(i_0)}, \ldots, x_{\sigma(i_{\ell-1})}) \ : \ \varphi(x_{i_0}, \ldots, x_{i_{\ell-1}}) \in u \text{ and } \ell \leq k\}$$

for $\sigma \in \mathrm{Sym}(k)$ and $u \in \mathfrak{qf}_L^{\mathrm{nr}}$. Note that this action of $\mathrm{Sym}(k)$ extends to $\mathcal{P}(\mathfrak{qf}_L^{\mathrm{nr}})$ in the natural way, namely,

$$(\sigma \cdot \nu)(B) = \nu(\{\sigma \cdot u \ : \ u \in B\})$$

for $\nu \in \mathcal{P}(\mathfrak{qf}_L^{\mathrm{nr}})$ and Borel $B \subseteq \mathfrak{qf}_L^{\mathrm{nr}}$.

Definition 2.2. *An* **extended L-hypergraphon** *is a measurable map*

$$W \colon [0,1]^{\mathfrak{P}_{<k}(k)} \to \mathcal{P}(\mathfrak{qf}_L^{\mathrm{nr}})$$

such that for any $\langle x_F \rangle_{F \in \mathfrak{P}_{<k}(k)} \in [0,1]^{\mathfrak{P}_{<k}(k)}$ *and* $\sigma \in \mathrm{Sym}(k)$,

$$W(\langle x_{\sigma(F)} \rangle_{F \in \mathfrak{P}_{<k}(k)}) = \sigma \cdot W(\langle x_F \rangle_{F \in \mathfrak{P}_{<k}(k)}).$$

The next technical lemma and definitions show how an extended L-hypergraphon W gives rise to a random non-redundant L-structure $G(\mathbb{N}, W)$ whose distribution is an invariant measure on $\mathrm{Str}_L(\mathbb{N})$.

The following is a special case of a standard result from probability theory about the randomization of a kernel.

Lemma 2.3 (see [Kal02, Lemma 3.22]). *Let* $W \colon [0,1]^{\mathfrak{P}_{<k}(k)} \to \mathcal{P}(\mathfrak{qf}_L^{\mathrm{nr}})$ *be an extended L-hypergraphon. There is a measurable function* $W^* \colon [0,1]^{\mathfrak{P}_{<k}(k)} \times [0,1] \to \mathfrak{qf}_L^{\mathrm{nr}}$ *such that whenever* ζ *is a uniform random variable in* $[0,1]$, *then for all* $t \in [0,1]^{\mathfrak{P}_{<k}(k)}$, *the random variable* $W^*(t, \zeta)$ *has distribution* $W(t)$.

Definition 2.4 introduces notation that we will use throughout this section and the next.

Definition 2.4. *Let* $\mathfrak{n} \in \mathbb{N} \cup \{\mathbb{N}\}$, *and suppose* $\mathfrak{n} \geq k$ *(when* $\mathfrak{n} \in \mathbb{N}$). *Let* $J \in \mathfrak{P}_k(\mathfrak{n})$, *and let* $\tau_J \colon [k] \to J$ *be the unique increasing bijection from* $[k]$ *to* J. *Define* $\widehat{X}_J$ *to be the sequence* $\langle X_{\tau_J(F)} \rangle_{F \in \mathfrak{P}_{<k}(k)}$ *consisting of terms of the form* X_I, *where* I *is from* $\mathfrak{P}_{<k}(\mathfrak{n})$.

Recall that L consists of relation symbols all of arity k, and so in order to define a non-redundant L-structure, it suffices to describe the complete quantifier-free type of every strictly increasing k-tuple in the structure.

Definition 2.5. *Let* W *be an extended L-hypergraphon and* $\mathfrak{n} \in \mathbb{N} \cup \{\mathbb{N}\}$, *and suppose* $\mathfrak{n} \geq k$ *(when* $\mathfrak{n} \in \mathbb{N}$).

- *Define* $M(\mathfrak{n}, W) \colon [0,1]^{\mathfrak{P}_{\leq k}(\mathfrak{n})} \to \mathrm{Str}_L(\mathfrak{n})$ *to be the map such that for all* $\langle x_D \rangle_{D \in \mathfrak{P}_{\leq k}(\mathfrak{n})} \in [0,1]^{\mathfrak{P}_{\leq k}(\mathfrak{n})}$ *and every* $J \in \mathfrak{P}_k(\mathfrak{n})$, *the quantifier-free type of the tuple* $\langle \tau_J(0), \ldots, \tau_J(k-1) \rangle$ *in the L-structure* $M(\mathfrak{n}, W)(\langle x_D \rangle_{D \in \mathfrak{P}_{\leq k}(\mathfrak{n})})$ *is* $W^*(\widehat{x}_J, x_J)$.

- *Let $\langle \zeta_D \rangle_{D \in \mathfrak{P}_{\leq k}(\mathfrak{n})}$ be an i.i.d. sequence of uniform random variables in $[0,1]$. Define $G(\mathfrak{n}, W)$ to be the random L-structure $M(\mathfrak{n}, W)(\langle \zeta_D \rangle_{D \in \mathfrak{P}_{\leq k}(\mathfrak{n})})$.*
- *For $J \in \mathfrak{P}_k(\mathfrak{n})$, define the random variable $E_J^W := W^*(\widehat{\zeta}_J, \zeta_J)$.*

In summary, $G(\mathfrak{n}, W)$ is the random L-structure with underlying set $[\mathfrak{n}]$ whose quantifier-free k-types are given by the random variables E_J^W for $J \in \mathfrak{P}_k(\mathfrak{n})$. It is easy to check that the distribution of $G(\mathfrak{n}, W)$ does not depend on the specific choice of i.i.d. uniform $\langle \zeta_D \rangle_{D \in \mathfrak{P}_{\leq k}(\mathfrak{n})}$ or on the specific function W^* satisfying Lemma 2.3.

Observe that for an extended L-hypergraphon W, the distribution of $G(\mathbb{N}, W)$ is a probability measure on $\mathrm{Str}_L(\mathbb{N})$ that is invariant because $\langle \zeta_D \rangle_{D \in \mathfrak{P}_{\leq k}(\mathbb{N})}$ is i.i.d., and is non-redundant because W^* takes values in $\mathfrak{qf}_L^{\mathrm{nr}}$. In fact, as Theorem 2.6 asserts, every non-redundant invariant measure on $\mathrm{Str}_L(\mathbb{N})$ arises from some W in this way; this result is a variant of the usual Aldous–Hoover–Kallenberg theorem, and is a specialization of [Ack15, Theorem 2.37].

Theorem 2.6 (Aldous–Hoover–Kallenberg theorem). *A non-redundant probability measure μ on $\mathrm{Str}_L(\mathbb{N})$ is invariant under the action of $\mathrm{Sym}(\mathbb{N})$ if and only if μ is the distribution of $G(\mathbb{N}, W)$ for some extended L-hypergraphon W.*

For more details on the usual Aldous–Hoover–Kallenberg theorem, which is stated in terms of exchangeable arrays, see [Kal05, Chapter 7] and the historical notes to that chapter. There is an essentially equivalent statement, in terms of hypergraphons, that provides a correspondence between hypergraphons (of arity k) and ergodic invariant measures on the space of k-uniform graphs (see [DJ08] and [Aus08]). The version that we have stated above in Theorem 2.6 provides an analogous correspondence between extended L-hypergraphons and non-redundant invariant measures on $\mathrm{Str}_L(\mathbb{N})$. Note that, unlike those invariant measures that arise from hypergraphons, the ones arising in Theorem 2.6 can in general be mixtures of ergodic invariant measures, and can be concentrated on structures that include multiple relations (though finitely many, of the same arity) which need not be symmetric.

In fact, we could have defined extended L-hypergraphons, and then described the associated sampling procedure $G(\mathbb{N}, W)$ and stated the Aldous–Hoover–Kallenberg theorem, for the even more general case that does not stipulate non-redundance. However, the invariant measure resulting from the sampling procedure would then be sensitive to measure 0 changes in the extended L-hypergraphon, unlike the situation for standard hypergraphons or the usual Aldous–Hoover–Kallenberg representation, and so we have restricted our definitions and results to the non-redundant case.

Using the machinery we have developed above, we can now prove the main theorems of this section. By Theorem 2.6, in order to study the entropy function of a non-redundant invariant measure on $\mathrm{Str}_L(\mathbb{N})$, we may ask for a suitable extended L-hypergraphon W and then analyze the entropy function of $G(\mathbb{N}, W)$. In Theorem 2.7, we consider such a $G(\mathbb{N}, W)$, and express the leading term of its entropy function as a function of W. As a consequence, in Corollary 2.8, we obtain bounds on the entropy function of a non-redundant invariant measure for L.

The proof of Theorem 2.7 closely follows that of [Jan13, Theorem D.5].

Theorem 2.7. *Fix an extended L-hypergraphon W. Then*

$$\lim_{n \to \infty} \frac{\mathrm{Ent}(G(\mathbb{N}, W))(n)}{|\mathfrak{P}_k(n)|} = \int h(W(\widehat{z}_{[k]})) \; d\lambda(\widehat{z}_{[k]}).$$

In particular, $\mathrm{Ent}(G(\mathbb{N}, W))(n) = Cn^k + o(n^k)$ *for some constant C.*

Proof. Observe that for all $n \in \mathbb{N}$, we have $\mathrm{Ent}(G(\mathbb{N}, W))(n) = h(G(n, W))$. We first show a lower bound on $h(G(n, W))$ for all $n \geq k$. For $J \in \mathfrak{P}_k(n)$, the random variables E_J^W are independent conditioned on $\langle \zeta_I \rangle_{I \in \mathfrak{P}_{<k}(n)}$ and so, as the conditional entropy of conditionally independent random variables is additive, we have

$$\begin{aligned}
H\big(G(n, W) \mid \langle \zeta_I \rangle_{I \in \mathfrak{P}_{<k}(n)}\big) &= \sum_{J \in \mathfrak{P}_k(n)} H\big(E_J^W \mid \langle \zeta_I \rangle_{I \in \mathfrak{P}_{<k}(n)}\big) \\
&= \sum_{J \in \mathfrak{P}_k(n)} H\big(\mathbb{E}\big(W(\widehat{\zeta}_J) \mid \langle \zeta_I \rangle_{I \in \mathfrak{P}_{<k}(n)}\big)\big) \\
&= \sum_{J \in \mathfrak{P}_k(n)} H(W(\widehat{\zeta}_J))
\end{aligned}$$

a.s., where the last equality follows from the fact that every random variable in the sequence $\widehat{\zeta}_J$ occurs within $\langle \zeta_I \rangle_{I \in \mathfrak{P}_{<k}(n)}$.

By Lemma 1.8, and then taking expectations of the first and last terms in the previous chain of equalities, we have

$$\begin{aligned}
h\big(G(n, W)\big) &\geq \mathbb{E}\big(H\big(G(n, W) \mid \langle \zeta_I \rangle_{I \in \mathfrak{P}_{<k}(n)}\big)\big) \\
&= \sum_{J \in \mathfrak{P}_k(n)} \mathbb{E}\big(H(W(\widehat{\zeta}_J))\big).
\end{aligned}$$

Because the distribution of the random variable $\widehat{\zeta}_J$ is the same for all $J \in \mathfrak{P}_k(n)$, we have

$$\sum_{J \in \mathfrak{P}_k(n)} \mathbb{E}\big(H(W(\widehat{\zeta}_J))\big) \;=\; |\mathfrak{P}_k(n)| \cdot \mathbb{E}\big(H(W(\widehat{\zeta}_{[k]}))\big)$$

$$=\; |\mathfrak{P}_k(n)| \cdot \int h(W(\widehat{z}_{[k]})) \; d\lambda(\widehat{z}_{[k]}).$$

This equation, together with the previous inequality, yields

$$\frac{h(G(n,W))}{|\mathfrak{P}_k(n)|} \;\geq\; \int h(W(\widehat{z}_{[k]})) \; d\lambda(\widehat{z}_{[k]}).$$

Now we show an upper bound on $h(G(n,W))$ for all $n \geq k$. Let $r \in \mathbb{N}$ be positive.

For $I \in \mathfrak{P}_{<k}(n)$, define $Y_I := \lfloor r \cdot \zeta_I \rfloor$, so that $Y_I = \ell$ precisely when $\frac{\ell}{r} \leq \zeta_I < \frac{\ell+1}{r}$. Then by Lemma 1.7, we have

$$h\big(G(n,W)\big) \;\leq\; h\big(G(n,W), \langle Y_I\rangle_{I \in \mathfrak{P}_{<k}(n)}\big)$$

$$=\; h\big(\langle Y_I\rangle_{I \in \mathfrak{P}_{<k}(n)}\big) + \mathbb{E}\Big(H\big(G(n,W) \mid \langle Y_I\rangle_{I \in \mathfrak{P}_{<k}(n)}\big)\Big).$$

The first term in this last expression is straightforward to calculate; we have

$$h\big(\langle Y_I\rangle_{I \in \mathfrak{P}_{<k}(n)}\big) = \big|\mathfrak{P}_{<k}(n)\big| \cdot \log_2(r),$$

as the Y_I are uniformly distributed on $[r]$ and independent (because the ζ_I are i.i.d. uniform on $[0,1]$).

We next calculate the second term; we will show that

$$\mathbb{E}\Big(H\big(G(n,W) \mid \langle Y_I\rangle_{I \in \mathfrak{P}_{<k}(n)}\big)\Big) \;=\; |\mathfrak{P}_k(n)| \cdot \int h(W_r(\widehat{z}_{[k]})) \; d\lambda(\widehat{z}_{[k]}).$$

By (the standard conditional extension of) Lemma 1.9, we have

$$H\big(G(n,W) \mid \langle Y_I\rangle_{I \in \mathfrak{P}_{<k}(n)}\big) \;\leq\; \sum_{J \in \mathfrak{P}_k(n)} H\big(E_J^W \mid \langle Y_I\rangle_{I \in \mathfrak{P}_{<k}(n)}\big)$$

$$=\; \sum_{J \in \mathfrak{P}_k(n)} H\big(E_J^W \mid \widehat{Y}_J\big)$$

a.s., where the last equality follows from the fact that the only random variables in $\langle \zeta_I\rangle_{I \in \mathfrak{P}_{<k}(n)}$ on which a given E_J^W depends are those among $\widehat{\zeta}_J$.

Given a function $\alpha \colon \mathfrak{P}_{<k}(k) \to [r]$, define

$$w_r(\alpha) \;:=\; r^{|\mathfrak{P}_{<k}(k)|} \cdot \int W(\widehat{z}_{[k]}) \cdot \prod_{F \in \mathfrak{P}_{<k}(k)} \mathbb{1}_{\left[\frac{\alpha(F)}{r}, \frac{\alpha(F)+1}{r}\right)}(z_F) \; d\lambda(\widehat{z}_{[k]}),$$

where $\mathbb{1}_S$ denotes the characteristic function of a set S (in this case, a half-open interval). Observe that

$$w_r(\alpha) \;=\; \mathbb{E}\Big(W(\widehat{\zeta}_{[k]}) \,\Big|\, \bigwedge_{F \in \mathfrak{P}_{<k}(k)} Y_F = \alpha(F)\Big)$$

a.s. (We may form this conditional expectation because the set of signed measures on $\mathfrak{qf}_L^{\mathrm{nr}}$ is a Banach space, and hence integration on this space is well-defined.)

For $\widehat{z}_{[k]} \in [0,1]^{\mathfrak{P}_{<k}(k)}$, define $\beta_{\widehat{z}_{[k]}} \colon \mathfrak{P}_{<k}(k) \to [r]$ by $\beta_{\widehat{z}_{[k]}}(F) = \lfloor r \cdot z_F \rfloor$ for $F \in \mathfrak{P}_{<k}(k)$, and let $W_r(\widehat{z}_{[k]}) := w_r(\beta_{\widehat{z}_{[k]}})$. Observe that while we have defined continuum-many instances of $\beta_{\widehat{z}_{[k]}}$, they range over the merely finitely many functions from $\mathfrak{P}_{<k}(k)$ to $[r]$. Further note that W_r is a step function, and that as $r \to \infty$, the function W_r converges to W pointwise a.e. We can think of W_r as the result of discretizing W along blocks of width $1/r$.

Recall the function τ_J defined in Definition 2.4. For any $J \in \mathfrak{P}_k(n)$, we have

$$\mathbb{P}\Big(E_J^W \in \cdot \,\Big|\, \bigwedge_{F \in \mathfrak{P}_{<k}(k)} Y_{\tau_J(F)} = \beta_{\widehat{z}_{[k]}}(F)\Big) \;=\; \mathbb{E}\Big(W(\widehat{\zeta}_J) \,\Big|\, \bigwedge_{F \in \mathfrak{P}_{<k}(k)} Y_{\tau_J(F)} = \beta_{\widehat{z}_{[k]}}(F)\Big)$$

$$=\; \mathbb{E}\Big(W(\widehat{\zeta}_{[k]}) \,\Big|\, \bigwedge_{F \in \mathfrak{P}_{<k}(k)} Y_F = \beta_{\widehat{z}_{[k]}}(F)\Big)$$

$$=\; w_r(\beta_{\widehat{z}_{[k]}})$$

a.s., because τ_J is a bijection from $[k]$ to J and the distribution of the random variable $\widehat{\zeta}_J$ is the same as that of $\widehat{\zeta}_{[k]}$. Therefore

$$h\Big(E_J^W \in \cdot \,\Big|\, \bigwedge_{F \in \mathfrak{P}_{<k}(k)} Y_{\tau_J(F)} = \beta_{\widehat{z}_{[k]}}(F)\Big) = h(w_r(\beta_{\widehat{z}_{[k]}})).$$

Summing both sides over all possible choices of $\beta_{\widehat{z}_{[k]}}$, we have

$$\sum_{\alpha \colon \mathfrak{P}_{<k}(k) \to [r]} h\Big(E_J^W \in \cdot \,\Big|\, \bigwedge_{F \in \mathfrak{P}_{<k}(k)} Y_{\tau_J(F)} = \alpha(F)\Big) = \sum_{\alpha \colon \mathfrak{P}_{<k}(k) \to [r]} h(w_r(\alpha))).$$

As before,

$$H\big(E_J^W \,\big|\, \langle Y_I \rangle_{I \in \mathfrak{P}_{<k}(n)}\big) = H\big(E_J^W \,\big|\, \widehat{Y}_J\big).$$

One can directly show that

$$\mathbb{E}\Big(H\big(E_J^W \,\big|\, \widehat{Y}_J\big)\Big) = r^{-|\mathfrak{P}_{<k}(k)|} \sum_{\alpha \colon \mathfrak{P}_{<k}(k) \to [r]} h\Big(E_J^W \in \cdot \,\Big|\, \bigwedge_{F \in \mathfrak{P}_{<k}(k)} Y_{\tau_J(F)} = \alpha(F)\Big).$$

Hence

$$\mathbb{E}\Big(H\big(E_J^W \mid \langle Y_I \rangle_{I \in \mathfrak{P}_{<k}(n)}\big)\Big) = r^{-|\mathfrak{P}_{<k}(k)|} \cdot \sum_{\alpha: \, \mathfrak{P}_{<k}(k) \to [r]} h(w_r(\alpha))$$

$$= \int h\big(W_r(\widehat{z}_{[k]})\big) \ d\lambda(\widehat{z}_{[k]}).$$

But once again by (an extension of) Lemma 1.9, we have

$$\mathbb{E}\Big(H\big(G(n,W) \mid \langle Y_I \rangle_{I \in \mathfrak{P}_{<k}(n)}\big)\Big) \leq \sum_{J \in \mathfrak{P}_k(n)} \mathbb{E}\Big(H\big(E_J^W \mid \langle Y_I \rangle_{I \in \mathfrak{P}_{<k}(n)}\big)\Big)$$

$$= |\mathfrak{P}_k(n)| \cdot \int h(W_r(\widehat{z}_{[k]})) \ d\lambda(\widehat{z}_{[k]}).$$

Putting together our two calculations, we obtain

$$h(G(n,W)) \leq |\mathfrak{P}_{<k}(n)| \cdot \log_2(r) + |\mathfrak{P}_k(n)| \cdot \int h(W_r(\widehat{z}_{[k]})) \ d\lambda(\widehat{z}_{[k]}).$$

Hence for each positive r, we have

$$\frac{h(G(n,W))}{|\mathfrak{P}_k(n)|} \leq \frac{|\mathfrak{P}_{<k}(n)|}{|\mathfrak{P}_k(n)|} \cdot \log_2(r) + \int h(W_r(\widehat{z}_{[k]})) \ d\lambda(\widehat{z}_{[k]}),$$

and so

$$\limsup_{n \to \infty} \frac{h(G(n,W))}{|\mathfrak{P}_k(n)|} \leq \int h(W_r(\widehat{z}_{[k]})) \ d\lambda(\widehat{z}_{[k]}).$$

Letting $r \to \infty$, recall that $W_r \to W$ a.e., and so by the dominated convergence theorem, we have

$$\limsup_{n \to \infty} \frac{h(G(n,W))}{|\mathfrak{P}_k(n)|} \leq \int h(W(\widehat{z}_{[k]})) \ d\lambda(\widehat{z}_{[k]}).$$

Combining our lower and upper bounds, we have

$$\lim_{n \to \infty} \frac{h(G(\mathbb{N},W))}{|\mathfrak{P}_k(n)|} = \int h(W(\widehat{z}_{[k]})) \ d\lambda(\widehat{z}_{[k]}),$$

as desired.

Finally, because $\lim_{n \to \infty} \frac{|\mathfrak{P}_k(n)|}{n^k} = 1$, we have

$$\text{Ent}(G(\mathbb{N},W))(n) = Cn^k + o(n^k),$$

where $C = \int h(W(\widehat{z}_{[k]})) \ d\lambda(\widehat{z}_{[k]})$. $\qquad \square$

As a corollary, we obtain a bound on the growth of entropy functions of non-redundant invariant measures for L.

Corollary 2.8. *Let μ be a non-redundant invariant measure on $\text{Str}_L(\mathbb{N})$. Then $\text{Ent}(\mu)(n) = Cn^k + o(n^k)$ for some constant C. In particular, $\text{Ent}(\mu)(n) = O(n^k)$.*

Proof. By Theorem 2.6 there is some extended L-hypergraphon W such that μ is the distribution of the random L-structure $G(\mathbb{N}, W)$. Its entropy function $\text{Ent}(\mu)$ is of the desired form by Theorem 2.7. □

3. Invariant measures sampled from a Borel hypergraph

We have seen in Corollary 2.8 that for a finite relational language L all of whose relation symbols have the same arity $k \geq 1$, the entropy function of an invariant measure on $\text{Str}_L(\mathbb{N})$ is of the form $Cn^k + o(n^k)$, where C is a constant depending on the invariant measure. In this section we consider the situation where $C = 0$.

For $k = 1$, consider an invariant measure μ on Str_L, and let W be an extended L-hypergraphon such that μ is the distribution of $G(\mathbb{N}, W)$. By Theorem 2.7, we have $\text{Ent}(\mu)(n) = Cn + o(n)$, where $C = \int h(W(z_\emptyset)) \, d\lambda(z_\emptyset)$. Suppose $C = 0$. Then $h(W(z_\emptyset)) = 0$ for a.e. $z_\emptyset$, and so $W(z_\emptyset)$ is a point mass a.e. Hence $G(\mathbb{N}, W)$ is a random L-structure where every element of $\mathbb{N}$ has the same quantifier-free 1-type, a.s. Therefore μ is a mixture of finitely many point masses, and $\text{Ent}(\mu)(n)$ is a constant that does not depend on n. In summary, for $k = 1$, the only possible entropy functions of sublinear growth are the constant functions.

Theorem 3.5, the main result of this section, states that in contrast, for $k > 1$, Corollary 2.8 is tight in the sense that for any given function γ that is $o(n^k)$, there is some non-redundant invariant measure whose entropy function is $o(n^k)$ but grows faster than γ. As discussed in the introduction, this result is a generalization of [HN13, Theorem 1.1], and the arguments in this section closely follow their proof.

For the rest of this section, fix $k \geq 2$ and let L be the language of k-uniform hypergraphs, i.e., $L = \{E\}$ where E is a k-ary relation symbol.

A *k-uniform hypergraph* is a non-redundant L-structure satisfying

$$\bigwedge_{\sigma \in \text{Sym}(k)} (\forall x_0, \ldots, x_{k-1}) \left(E(x_0, \ldots, x_{k-1}) \leftrightarrow E(x_{\sigma(0)}, \ldots, x_{\sigma(k-1)}) \right),$$

and we call the instantiation of E its *edge set*.

By a *Borel hypergraph*, we mean a k-uniform hypergraph $\mathcal{M}$ whose underlying set is $[0, 1]$ and such that for any atomic formula φ in the language of hypergraphs, the set $\{\bar{a} \in \mathcal{M} : \mathcal{M} \models \varphi(\bar{a})\}$ of realizations of φ in $\mathcal{M}$ is Borel.

Any extended L-hypergraphon W *yields* a non-redundant invariant measure, namely the distribution of the random L-structure $G(\mathbb{N}, W)$. We say that W *induces a Borel hypergraph* when this invariant measure can be obtained by sampling a random subhypergraph of some Borel hypergraph, as we make precise in Definition 3.1. We will see, in Lemma 3.2, that when W induces a Borel hypergraph, it yields an invariant measure whose entropy function is $o(n^k)$.

The main construction of this section, in Theorem 3.5, builds non-redundant invariant measures of arbitrarily high growth within $o(n^k)$ by sampling from certain extended L-hypergraphons that induce Borel hypergraphs.

Observe that there are only two non-redundant quantifier-free k-types in L that are consistent with the theory of k-uniform hypergraphs. Let $u_\top, u_\perp \in \mathfrak{qf}_L^{nr}$ be the unique non-redundant quantifier-free types containing

$$\bigwedge_{\sigma \in \mathrm{Sym}(k)} E(x_{\sigma(0)}, \dots, x_{\sigma(k-1)})$$

and

$$\bigwedge_{\sigma \in \mathrm{Sym}(k)} \neg E(x_{\sigma(0)}, \dots, x_{\sigma(k-1)}),$$

respectively, and let $\delta_\top, \delta_\perp \in \mathcal{P}(\mathfrak{qf}_L^{nr})$ be the respective point masses concentrated on them.

Recall from Definition 2.4 that $\widehat{x}_{[k]}$ denotes the tuple of variables $\langle x_I \rangle_{I \in \mathfrak{P}_{<k}(k)}$.

Definition 3.1. *We say that an extended L-hypergraphon $W \colon [0,1]^{\mathfrak{P}_{<k}(k)} \to \mathcal{P}(\mathfrak{qf}_L^{nr})$* **induces a Borel hypergraph** *if*

- *for a.e. pair of sequences $\widehat{x}_{[k]}$, $\widehat{y}_{[k]}$ of elements of $[0,1]$ with $x_{\{i\}} = y_{\{i\}}$ for all $i \in [k]$, we have $W(\widehat{x}_{[k]}) = W(\widehat{y}_{[k]})$, and*
- *for a.e. sequence $\widehat{x}_{[k]}$ of elements of $[0,1]$, the distribution $W(\widehat{x}_{[k]})$ is either $\delta_\top$ or $\delta_\perp$.*

It follows that an extended L-hypergraphon W induces a Borel hypergraph precisely when there is a Borel hypergraph $\mathcal{B}$ such that $W(\widehat{x}_{[k]})$ is a point mass concentrated on the quantifier-free type of $\langle x_{\{0\}}, \dots, x_{\{k-1\}} \rangle$ in $\mathcal{B}$. In this case we say that W *induces the Borel hypergraph* $\mathcal{B}$.

The notion of an extended L-hypergraphon inducing a Borel hypergraph is closely related, in the case $k = 2$, to that of a graphon being random-free. A *graphon* is a symmetric Borel function from $[0,1]^2$ to $[0,1]$; as described in [LS10] and [Jan13], it is called *random-free* when it is $\{0,1\}$-valued a.e. Every graphon gives rise to a random undirected graph on $\mathbb{N}$ whose distribution is an invariant measure. For $k = 2$, an extended L-hypergraphon W that yields an invariant measure concentrated on undirected graphs can be expressed as a mixture of invariant measures, each obtained via a graphon. In the case where such a W corresponds to a single graphon, W induces a Borel hypergraph precisely when the corresponding graphon is random-free.

The notion of a random-free graphon also essentially appeared, in the context of separate exchangeability, in work of Aldous in [Ald81, Proposition 3.6] and [Ald85, (14.15) and p. 133], and Diaconis–Freedman [DF81, (4.10)]. Further, Kallenberg [Kal99] describes, for all $k \geq 2$, the similar notion of a *simple array*; this corresponds to our notion of inducing a Borel hypergraph, in the case

where the distribution of the simple array is ergodic. Random-free graphons arise as well in [PV10] and [AFP16, §6.1], which consider invariant measures that are concentrated on a given orbit of the logic action.

Theorem 2.7 implies that any extended L-hypergraphon W yields an invariant measure whose entropy function is $O(n^k)$. Observe that there are extended L-hypergraphons achieving this upper bound, i.e., that yield an invariant measure whose entropy function is $\Omega(n^k)$. For example, one can directly calculate that the Erdős–Rényi extended L-hypergraphon given by the constant function

$$W_{\mathrm{ER}}(\widehat{x}_{[k]}) = \mathrm{Uniform}(\{u_\top, u_\bot\})$$

satisfies $\mathrm{Ent}(G(\mathbb{N}, W_{\mathrm{ER}}))(n) = \binom{n}{k}$.

However, we now show that this growth rate cannot be achieved for an extended L-hypergraphon that induces a Borel hypergraph.

Lemma 3.2. *Let W be an extended L-hypergraphon, and suppose W induces a Borel hypergraph. Then $\mathrm{Ent}(G(\mathbb{N}, W))(n) = o(n^k)$.*

Proof. Because W induces a Borel hypergraph, it takes the value $\delta_\top$ or $\delta_\bot$ a.e. But $h(\delta_\top) = h(\delta_\bot) = 0$, and so $\int h(W(\widehat{z}_{[k]}))\ d\lambda(\widehat{z}_{[k]}) = 0$. Hence by Theorem 2.7, we have $\mathrm{Ent}(G(\mathbb{N}, W))(n) = o(n^k)$. $\square$

As noted previously, Aldous [Ald85] and Janson [Jan13] have versions of Theorem 2.7 for $k = 2$. They also observe that their respective results immediately imply that a random-free graphon yields an invariant measure with entropy function that is $o(n^2)$; their proofs are similar to that of Lemma 3.2. Their setting involves working with graphons, which yield ergodic invariant measures. Under the restriction of ergodicity, it is easily seen that the converse of Lemma 3.2 holds for $k = 2$, as noted by Janson [Jan13, Theorem 10.16].

In contrast, the converse of Lemma 3.2 itself does not hold, as our notion of extended L-hypergraphon allows for ones that yield non-ergodic invariant measures. For example, consider the extended L-hypergraphon

$$W_{\blacksquare\square}(\widehat{x}_{[k]}) = \begin{cases} \delta_\top & \text{if } x_\emptyset < \frac{1}{2}, \\ \delta_\bot & \text{otherwise.} \end{cases}$$

Define the extended L-hypergraphons $W_\blacksquare(\widehat{x}_{[k]}) = \delta_\top$ and $W_\square(\widehat{x}_{[k]}) = \delta_\bot$, which each induce a Borel hypergraph. The random hypergraph $G(\mathbb{N}, W_{\blacksquare\square})$ is the complete hypergraph or the empty hypergraph, each with probability $\frac{1}{2}$, and so its distribution is a non-trivial mixture of the distributions of $G(\mathbb{N}, W_\blacksquare)$ and $G(\mathbb{N}, W_\square)$, hence a non-ergodic invariant measure. The extended L-hypergraphon $W_{\blacksquare\square}$ does not induce a Borel hypergraph as it depends on the variable $x_\emptyset$, yet the entropy function of $G(\mathbb{N}, W_{\blacksquare\square})$ is $o(n^k)$.

But in fact, for $k \geq 3$, there is a more interesting obstruction to a converse of Lemma 3.2, even among extended L-hypergraphons that yield ergodic invariant measures. The following example for $k = 3$ (which is easily generalized to larger values of k) makes fundamental use of the variables indexed by pairs from [3], and yet also yields an invariant measure whose entropy function is $o(n^3)$:

$$W_\triangle(\widehat{x}_{[3]}) = \begin{cases} \delta_\top & \text{if } x_{\{0,1\}} < \frac{1}{2} \text{ and } x_{\{0,2\}} < \frac{1}{2} \text{ and } x_{\{1,2\}} < \frac{1}{2}, \\ \delta_\perp & \text{otherwise.} \end{cases}$$

The random hypergraph $G(\mathbb{N}, W_\triangle)$ can be thought of as first building a "virtual" Erdős–Rényi graph with independent 2-edge probabilities $\frac{1}{2}$, then adding a 3-edge for each triangle existing in the graph, and then throwing away the virtual 2-edges. (For more about this example, see [Aus08, p. 92] and [Lov12, Example 23.11].)

In Lemma 3.2, we established that any extended L-hypergraphon W that induces a Borel hypergraph is such that the entropy function of $G(\mathbb{N}, W)$ is $o(n^k)$. We now proceed to show that there are such W for which the growth of $\mathrm{Ent}(G(\mathbb{N}, W))(n)$ is arbitrarily close to n^k.

We first define a kind of "blow up" that creates an extended L-hypergraphon from a countably infinite L-structure. We will use this notion in Lemma 3.4. Blow ups are a standard technique for expanding a countable structure into a continuum-sized structure that has a positive-measure worth of "copies" of each point from the original. (See, e.g., the use of step function graphons in [Lov12].)

Definition 3.3. *Let $\mathcal{M} \in \mathrm{Str}_L(\mathbb{N})$ and let $\pi \colon [0,1] \to \mathbb{N}$ be a Borel map such that $\lambda(\pi^{-1}(i)) > 0$ for all $i \in \mathbb{N}$. The π-**blow up of** $\mathcal{M}$ is defined to be the extended L-hypergraphon $W \colon [0,1]^{\mathfrak{P}_{<k}(k)} \to \{\delta_\top, \delta_\perp\}$ given by*

$$W(\widehat{x}_{[k]}) = \begin{cases} \delta_\top & \text{if } \mathcal{M} \models E\big(\pi(x_{\{0\}}), \pi(x_{\{1\}}), \ldots, \pi(x_{\{k-1\}})\big), \\ \delta_\perp & \text{otherwise.} \end{cases}$$

Observe that such a W induces a Borel hypergraph, as on every input it outputs either the value $\delta_\top$ or $\delta_\perp$, and it depends only on the variables $x_{\{0\}}, x_{\{1\}}, \ldots, x_{\{k-1\}}$. In particular, W induces the Borel hypergraph $\mathcal{B}$ with underlying set $[0,1]$ given by:

$$\mathcal{B} \models E(x_0, x_1, \ldots, x_{k-1})$$

if and only if

$$\mathcal{M} \models E\big(\pi(x_0), \pi(x_1), \ldots, \pi(x_{k-1})\big).$$

Hence we can think of π^{-1} as a Borel partition of the unit interval into positive measure pieces, such that each element of $\mathcal{M}$ is "blown up" into a piece of the partition in $\mathcal{B}$, and each piece of the partition arises in this way.

Next, we proceed to construct an extended L-hypergraphon W that yields an invariant measure whose entropy function has the desired growth. Analogously to Hatami–Norine [HN13], we let W be the blow up of a particular countably infinite structure, the *Rado k-hypergraph*, a well-known generalization of the *transversal-uniform graph* used in [HN13]. The Rado k-hypergraph is the countable homogeneous-universal k-uniform hypergraph, namely, the unique (up to isomorphism) countable k-uniform hypergraph satisfying the so-called "Alice's restaurant" axioms. These axioms state that for any possible way of extending a finite induced subhypergraph of the Rado k-hypergraph by one vertex to obtain a k-uniform hypergraph, there is some element of the Rado k-hypergraph that realizes this extension.

We now describe an inductive construction of an instantiation $\mathcal{R}_k$ of the Rado k-hypergraph, with underlying set $\mathbb{N}$. At each stage $\ell \in \mathbb{N}$ we define a finite set A_ℓ of new vertices, which we call *generation ℓ*, and build a k-uniform hypergraph G_ℓ with underlying set $V_\ell := \bigcup_{j \le \ell} A_j$.

Stage 0: Let $A_0 := \{0\}$ consist of a single vertex, and let G_0 be the empty hypergraph with vertex set $V_0 = A_0$.

Stage $\ell > 0$: Let A_ℓ consist of one new vertex a_X for each subset X of unordered $(k-1)$-tuples from $V_{\ell-1}$, with the elements of A_ℓ chosen to be the consecutive least elements of $\mathbb{N}$ not yet used. Let G_ℓ be the k-uniform hypergraph on vertex set $V_\ell = V_{\ell-1} \cup A_\ell$ whose edges are those of $G_{\ell-1}$ along with, for each such X and every unordered tuple $d \in X$, an edge consisting of a_X and the $k-1$ vertices in d.

Define $\mathcal{R}_k$ to be the union of the hypergraphs G_ℓ, i.e., the hypergraph with vertex set $\bigcup_{\ell \in \mathbb{N}} V_\ell$ and edge set $\bigcup_{\ell \in \mathbb{N}} E^{G_\ell}$. Observe that $\mathcal{R}_k$ is a k-uniform hypergraph with underlying set $\mathbb{N}$ that satisfies the Alice's restaurant axioms: Given a finite induced subhypergraph D of $\mathcal{R}_k$, all one-vertex extensions of D are realized in stage $\ell + 1$, for any ℓ such that V_ℓ contains the vertices of D.

The following lemma, which we will use in the proof of Theorem 3.5, provides a lower bound on the entropy function of a random hypergraph sampled from a blow up of $\mathcal{R}_k$ in the case where elements of $\mathcal{R}_k$ belonging to the same generation get blown up to sets of equal measure. Both the statement and proof of the lemma are directly analogous to those of [HN13, Lemma 2.1].

Recall from Definition 2.5 that $\langle \zeta_D \rangle_{D \in \mathfrak{P}_{\le k}(\mathbb{N})}$ is the collection of i.i.d. uniform random variables in $[0, 1]$ in terms of which the random L-structure $G(\mathbb{N}, W)$ is defined.

Lemma 3.4. *Let $\pi\colon [0,1] \to \mathbb{N}$ be a Borel map such that $\lambda(\pi^{-1}(i)) > 0$ for all $i \in \mathbb{N}$, and let W be a π-blow up of $\mathcal{R}_k$. Suppose that $\lambda(\pi^{-1}(a)) = \lambda(\pi^{-1}(b))$ for all $\ell \in \mathbb{N}$ and $a, b \in A_\ell$. Then for all $n \in \mathbb{N}$ and $\rho\colon [n] \to \mathbb{N}$, we have*

$$H\Big(G(n, W) \,\Big|\, \bigwedge_{j \in [n]} \pi(\zeta_{\{j\}}) \in A_{\rho(j)}\Big) \geq \binom{|\rho([n])|}{k},$$

a.s.

Proof. Let $S \subseteq [n]$ be maximal such that ρ is injective on S, and write $G(S, W)$ to denote the random induced substructure of $G(n, W)$ with underlying set S. Then

$$H\Big(G(n, W) \,\Big|\, \bigwedge_{j \in [n]} \pi(\zeta_{\{j\}}) \in A_{\rho(j)}\Big) \geq H\Big(G(S, W) \,\Big|\, \bigwedge_{j \in S} \pi(\zeta_{\{j\}}) \in A_{\rho(j)}\Big).$$

Consider the random measure

$$\mathbb{E}\Big(G(S, W) \,\Big|\, \bigwedge_{j \in S} \pi(\zeta_{\{j\}}) \in A_{\rho(j)}\Big). \tag{$\dagger$}$$

Since W is a blow up of $\mathcal{R}_k$, the random hypergraph $G(S, W)$ is a $|S|$-element sample with replacement from $\mathcal{R}_k$, with the vertices relabeled by S. By the injectivity of ρ, the condition in ($\dagger$) constrains the elements of $G(S, W)$ to be obtained from distinct generations of $\mathcal{R}_k$, which implies that the random distribution on hypergraphs given by ($\dagger$) is actually sampled from $\mathcal{R}_k$ without replacement a.s.

Further, each $\zeta_{\{j\}}$ is uniform, and π^{-1} assigns sets of equal measure to vertices in $\mathcal{R}_k$ of the same generation. So ($\dagger$) is a.s. the distribution of the random hypergraph Q with underlying set S obtained by, for each $j \in S$, uniformly selecting a vertex of $\mathcal{R}_k$ from among those in generation $\rho(j)$, and taking the edges induced from $\mathcal{R}_k$.

Let $\ell \in \mathbb{N}$, and consider a subset $U \subseteq V_\ell$. Write T for the set of unordered $(k-1)$-tuples from U. Suppose $\ell' > \ell$. For every subset $X \subseteq T$, exactly a $2^{-|T|}$-fraction of the vertices in $A_{\ell'}$ form an edge with every $(k-1)$-tuple in X and with no $(k-1)$-tuple in $T \setminus X$. For a vertex v selected uniformly at random from $A_{\ell'}$, let $G_{U,v}$ be the (not necessarily induced) random subhypergraph of $G_{\ell'}$ that has vertex set $U \cup \{v\}$ and whose edges are those in $G_{\ell'}$ that consist of v along with $k-1$ vertices from U. Then $G_{U,v}$ is equally likely to be any of the hypergraphs with vertex set $U \cup \{v\}$ whose edges all include v.

For each $j \in S$, define $S_j := \{i \in S : \rho(i) < \rho(j)\}$, and let Q_j be the (not necessarily induced) random subhypergraph of Q that has vertex set $S_j \cup \{j\}$ and whose edges are those in Q that consist of j along with $k-1$ vertices from S_j. Then for each $j \in S$, the random hypergraph Q_j is equally likely to be any given hypergraph F_j with vertex set $S_j \cup \{j\}$ all of whose edges

include j. Further, for any choice of hypergraphs $\{F_j : j \in S\}$ as above, the events $Q_j = F_j$ for $j \in S$ are independent because the random variables $\zeta_{\{j\}}$ are independent.

Now, for any hypergraph F with underlying set S, we can write F as the union, over $j \in S$, of the subhypergraph of F that has vertex set $S_j \cup \{j\}$ and whose edges are those in F consisting of j along with $k-1$ vertices from S_j. We therefore see that Q is equally likely to be any hypergraph on vertex set S.

In summary, the random distribution ($\dagger$) is a.s. the uniform measure on k-uniform hypergraphs with underlying set S. Hence

$$H\Big(G(S,W) \ \Big| \ \bigwedge_{j \in S} \pi(\zeta_{\{j\}}) \in A_{\rho(j)}\Big) \ = \ \binom{|S|}{k}$$

a.s., establishing the lemma. $\qquad\square$

We now prove the main result of this section, which asserts that the entropy function of an invariant measure can have arbitrarily large growth rate within $o(n^k)$. This result is a higher-arity version of [HN13, Theorem 1.1], and its proof proceeds via the same steps. We include the proof here (with appropriately modified parameters and notation) for completeness.

Theorem 3.5. *Suppose $\gamma\colon \mathbb{N} \to [0,1]$ is a function such that $\lim_{n\to\infty} \gamma(n) = 0$. Then there is an extended L-hypergraphon W that induces a Borel hypergraph and is such that $\mathrm{Ent}(G(\mathbb{N},W))(n) = o(n^k)$ and $\mathrm{Ent}(G(\mathbb{N},W))(n) = \Omega(\gamma(n) \cdot n^k)$.*

Proof. We will define a Borel map $\pi\colon [0,1] \to \mathbb{N}$ satisfying $\lambda(\pi^{-1}(i)) > 0$ for all $i \in \mathbb{N}$ in such a way that the π-blow up of $\mathcal{R}_k$, which we denote by W, has the desired properties.

By the observation that follows Definition 3.3, if $\lambda(\pi^{-1}(i)) > 0$ for all $i \in \mathbb{N}$, then the π-blow up W induces a Borel hypergraph, and so by Lemma 3.2, we have $\mathrm{Ent}(G(\mathbb{N},W))(n) = o(n^k)$. Hence it suffices to construct π satisfying $\lambda(\pi^{-1}(i)) > 0$ for all $i \in \mathbb{N}$ in such a way that $\mathrm{Ent}(G(\mathbb{N},W))(n) = \Omega(\gamma(n) \cdot n^k)$.

For positive $r \in \mathbb{N}$, define

$$g_r := \max\big\{\{2^{r+3}k\} \cup \{n \in \mathbb{N} : \gamma(n) > 2^{-(r+1)k-3k-1}k^{-k}\}\big\}.$$

Note that $\lim_{n\to\infty} \gamma(n) = 0$ and so for each r there are only finitely many n such that $\gamma(n) > 2^{-(r+1)k-3k-1}k^{-k}$; hence g_r is well-defined.

Observe that for all $n \geq g_1 + 1$, the inequalities

$$n > 2^{r+2}k \qquad \text{and} \qquad \gamma(n) \leq 2^{-rk-3k-1}k^{-k} \qquad\qquad (\star)$$

hold when $r = 1$. The remainder of the proof establishes that for all such n, we have $\mathrm{Ent}(G(\mathbb{N},W))(n) \ \geq \ n^k \cdot \gamma(n)$.

Fix $n \geq g_1 + 1$. We have seen that there is at least one r satisfying $(\star)$; on the other hand, there are only finitely many choices of r for which $(\star)$ holds. Let q be the largest such r. Then either $n < 2^{q+3}k$ or $\gamma(n) > 2^{-(q+1)k-3k-1}k^{-k}$, and so $n \leq g_q$ by the definition of g_q.

For each $r \geq 1$, define

$$\Gamma_r := \big\{\ell \in \mathbb{N} \,:\, \textstyle\sum_{i=1}^{r-1} g_i \leq \ell < \sum_{i=1}^{r} g_i\big\},$$

so that $\{\Gamma_r\}_{r \geq 1}$ is a partition of $\mathbb{N}$. For every $\ell \in \Gamma_r$, let $\alpha_\ell := \frac{1}{g_r 2^r}$. Observe that for each $r \geq 1$, we have $|\Gamma_r| = g_r$, and so $\sum_{\ell \in \mathbb{N}} \alpha_\ell = 1$.

As a consequence, there is a Borel map $\pi \colon [0,1] \to \mathbb{N}$ such that for all $a \in \mathcal{R}_k$, we have

$$\lambda(\pi^{-1}(a)) = \frac{\alpha_\ell}{|A_\ell|}$$

where ℓ is such that $a \in A_\ell$. In other words, vertices in $\mathcal{R}_k$ of the same generation A_ℓ are blown up to sets of the same positive measure, and the entire generation A_ℓ is blown up to a set of measure α_ℓ. We may therefore apply Lemma 3.4. Hence for any $\rho \colon [n] \to \mathbb{N}$, we have

$$H\Big(G(n,W) \,\Big|\, \bigwedge_{j \in [n]} \pi(\zeta_{\{j\}}) \in A_{\rho(j)}\Big) \geq \binom{|\rho([n])|}{k}$$

a.s.

By Lemma 1.8, we have

$$\mathrm{Ent}(G(\mathbb{N},W))(n) \;\geq\; \mathbb{E}\Big(H\Big(G(n,W) \,\Big|\, \bigvee_{\rho \colon [n] \to \mathbb{N}} \big(\bigwedge_{j \in [n]} \pi(\zeta_{\{j\}}) \in A_{\rho(j)}\big)\Big)\Big)$$

$$\geq \sum_{\rho \colon [n] \to \mathbb{N}} \mathbb{P}\Big(\bigwedge_{j \in [n]} \pi(\zeta_{\{j\}}) \in A_{\rho(j)}\Big) \cdot \binom{|\rho([n])|}{k}$$

$$\geq \mathbb{P}\Big(|Z| \geq n \cdot 2^{-q-2}\Big) \cdot \binom{\lceil n \cdot 2^{-q-2} \rceil}{k}, \tag{$\ddagger$}$$

where we define the random set $Z := \bigcup_{j \in [n]} \{\ell \in \mathbb{N} \,:\, \pi(\zeta_{\{j\}}) \in A_\ell\}$.

Now define the random quantity $X := |\{Z \cap \Gamma_q\}|$, and note that we always have $X \leq |Z|$. Because $\langle \zeta_{\{j\}} \rangle_{j \in [n]}$ is an i.i.d. uniform sequence, we have

$$\mathbb{E}(X) \;=\; \sum_{\ell \in \Gamma_q} \mathbb{P}\Big(\bigvee_{j \in [n]} \pi(\zeta_{\{j\}}) \in A_\ell\Big)$$

$$= \sum_{\ell \in \Gamma_q} (1 - (1 - \alpha_\ell)^n)$$

$$= g_q \cdot \Big(1 - \Big(1 - \frac{1}{g_q 2^q}\Big)^n\Big).$$

Observe that $(1-x)^n \leq 1 - nx + n^2x^2 \leq 1 - \frac{nx}{2}$ holds for all $x \in [0, \frac{1}{2n}]$. Since $n \leq g_q$ and $q \geq 1$, we have $\frac{1}{g_q 2^q} \in [0, \frac{1}{2n}]$, and so

$$1 - \left(1 - \frac{1}{g_q 2^q}\right)^n \geq \frac{n}{g_q 2^{q+1}}.$$

Putting these together, we get

$$\mathbb{E}(X) \geq n \cdot 2^{-q-1}.$$

By Chebyshev's inequality, we have

$$\mathbb{P}\left[|X - \mathbb{E}(X)| \geq \frac{\mathbb{E}(X)}{2}\right] \leq \frac{4 \operatorname{Var}(X)}{(\mathbb{E}(X))^2}.$$

Hence,

$$\begin{aligned}
1 - \frac{4 \operatorname{Var}(X)}{(\mathbb{E}(X))^2} &\leq \mathbb{P}\left[|X - \mathbb{E}(X)| < \frac{\mathbb{E}(X)}{2}\right] \\
&\leq \mathbb{P}\left[X > \frac{\mathbb{E}(X)}{2}\right] \\
&\leq \mathbb{P}(X > n \cdot 2^{-q-2}) \\
&\leq \mathbb{P}(|Z| > n \cdot 2^{-q-2}).
\end{aligned}$$

For distinct $\ell, \ell' \in \Gamma_q$, the events $\ell \in Z$ and $\ell' \in Z$ have negative correlation, which implies that $\operatorname{Var}(X) \leq \mathbb{E}(X)$. Hence

$$\mathbb{P}(|Z| > n \cdot 2^{-q-2}) \geq 1 - \frac{4}{\mathbb{E}(X)} \geq 1 - \frac{4}{n \cdot 2^{-q-1}} \geq \frac{1}{2}.$$

Finally, substituting in (‡) and recalling that, by (⋆) for $r = q$, we have $n \cdot 2^{-q-2} > k$ and $\gamma(n) \leq 2^{-qk-3k-1}k^{-k}$, we obtain

$$\begin{aligned}
\operatorname{Ent}(G(\mathbb{N}, W))(n) &\geq \frac{1}{2} \cdot \binom{\lceil n \cdot 2^{-q-2}\rceil}{k} \\
&\geq \frac{1}{2} \cdot \frac{(n \cdot 2^{-q-2} - k)^k}{k^k} \\
&\geq \frac{1}{2} \cdot (n \cdot 2^{-q-3})^k k^{-k} \\
&= n^k \cdot 2^{-qk-3k-1}k^{-k} \\
&\geq n^k \cdot \gamma(n),
\end{aligned}$$

as desired. $\qquad\square$

4. Non-redundant invariant measures

In this section we consider entropy functions of invariant measures for countable languages that may be infinite, possibly of unbounded arity. If an invariant measure fails to be non-redundant, then its entropy function may take the value ∞ even when there are only finitely many relation symbols of each arity in the language. Hence we restrict to the case of non-redundant invariant measures for relational languages, and provide an upper bound on the entropy function in terms of the number of relation symbols of each arity.

For the rest of this section, let L be a countable relational language, possibly infinite.

We show, in Proposition 4.3, that the entropy function of a non-redundant invariant measure for L is dominated by that of a particular random L-structure that generalizes the Erdős–Rényi random graph having edge probability $\frac{1}{2}$. We also calculate, in Lemma 4.2, the entropy function of such a maximal entropy structure explicitly in terms of the number of relation symbols of each arity in L. In the case where L has finitely many relation symbols of each arity, this provides a more precise version of Lemma 1.14, which states that such an entropy function takes values in $\mathbb{R}$. Moreover, this calculation shows that there are $\mathbb{R}$-valued entropy functions that grow arbitrarily fast, in contrast to the situation for finite languages, where the growth is at most polynomial.

We now define the *uniform non-redundant measure* for L, which is the distribution of a random structure obtained by independently flipping a fair coin to decide every relation on a tuple of distinct elements, and setting all relations on tuples with repeated elements to false.

For a relation symbol $R \in L$, write $\mathrm{arity}(R)$ to denote its arity.

Definition 4.1. *Given a set X, define*

$$\mathrm{NR}_{L,X} := \{\langle R, \overline{x} \rangle \, : \, R \in L, \ \overline{x} \in X \text{ has distinct entries, and } |\overline{x}| = \mathrm{arity}(R)\}.$$

Let $\{\xi_{R(\overline{x})} \, : \, \langle R, \overline{x} \rangle \in \mathrm{NR}_{L,\mathbb{N}}\}$ be a collection of i.i.d. uniform $\{\top, \bot\}$-valued random variables and let Ξ be the $\mathrm{Str}_L(\mathbb{N})$-valued random variable given by

$$\Xi \models R(\overline{x}) \qquad \textit{if and only if} \qquad \langle R, \overline{x} \rangle \in \mathrm{NR}_{L,\mathbb{N}} \quad \textit{and} \quad \xi_{R(\overline{x})} = \top$$

*for every $R \in L$ and $\mathrm{arity}(R)$-tuple $\overline{x}$ of elements from $\mathbb{N}$. Define $\mu_{\mathrm{un},L}$ to be the distribution of Ξ, and call it the **uniform non-redundant measure** for L.*

It is easy to see that $\mu_{\mathrm{un},L}$ is both invariant and non-redundant. Note that for $n \in \mathbb{N}$, the measure $(\mu_{\mathrm{un},L})_n$ is the uniform distribution on non-redundant structures in $\mathrm{Str}_L(n)$.

Let $\mathfrak{a}_L \colon \mathbb{N} \to \mathbb{N} \cup \{\infty\}$ be the function sending each $n \in \mathbb{N}$ to the number of relation symbols in L having arity n. In the following lemma we calculate the entropy function of $\mu_{\mathrm{un},L}$ in terms of the function $\mathfrak{a}_L$.

Lemma 4.2. *For any $n \in \mathbb{N}$, we have*

$$\mathrm{Ent}(\mu_{\mathrm{un},L})(n) \;=\; \sum_{r \le n} \binom{n}{r} \cdot r! \cdot \mathfrak{a}_L(r).$$

Proof. The invariant measure $\mu_{\mathrm{un},L}$ is non-redundant, and so for each $n \in \mathbb{N}$, the invariant measure $(\mu_{\mathrm{un},L})_n$ is concentrated on elements of $\mathrm{Str}_L(n)$ in which no relations of arity greater than n hold. Hence $(\mu_{\mathrm{un},L})_n$ is determined by the set $\{\xi_{R(\overline{x})} : \langle R, \overline{x}\rangle \in \mathrm{NR}_{L,[n]}\}$ of random variables. Because these random variables are independent, we have

$$
\begin{aligned}
\mathrm{Ent}(\mu_{\mathrm{un},L})(n) \;&=\; \sum_{(R,\overline{x}) \in \mathrm{NR}_{L,[n]}} h(\xi_{R(\overline{x})}) \\
&=\; \big|\mathrm{NR}_{L,[n]}\big| \\
&=\; \sum_{r \le n} \binom{n}{r} \cdot r! \cdot \mathfrak{a}_L(r),
\end{aligned}
$$

as the number of r-tuples from $[n]$ consisting of distinct elements is $\binom{n}{r} \cdot r!$, and there are $\mathfrak{a}_L(r)$-many relation symbols of arity r in L. $\qquad\square$

Observe that when L has only finitely many relation symbols of each arity, Lemma 4.2 shows that $\mathrm{Ent}(\mu_{\mathrm{un},L})$ is $\mathbb{R}$-valued. Hence, by varying the choice of such an L (and hence the function $\mathfrak{a}_L$), we can obtain $\mathbb{R}$-valued entropy functions that grow arbitrarily fast.

We now show that $\mu_{\mathrm{un},L}$ has the fastest growing entropy function among non-redundant invariant measures on $\mathrm{Str}_L(\mathbb{N})$.

Proposition 4.3. *Let ν be a non-redundant invariant measure on $\mathrm{Str}_L(\mathbb{N})$. Then*

$$\mathrm{Ent}(\nu)(n) \;\le\; \mathrm{Ent}(\mu_{\mathrm{un},L})(n)$$

for all $n \in \mathbb{N}$.

Proof. Suppose $\mathcal{N}$ is a random $\mathrm{Str}_L(\mathbb{N})$-structure with distribution ν. For $R \in L$ and $\overline{x} \in \mathbb{N}$ with $|\overline{x}| = \mathrm{arity}(R)$, recall that the random instantiation $R^{\mathcal{N}}$ satisfies

$$R^{\mathcal{N}}(\overline{x}) = \top \qquad \text{if and only if} \qquad \mathcal{N} \models R(\overline{x}).$$

By Lemma 1.9 we then have, for $n \in \mathbb{N}$,

$$\mathrm{Ent}(\nu)(n) \;=\; \mathrm{Ent}(\mathcal{N})(n) \;\leq\; \sum_{\substack{R \in L,\; \overline{x} \in [n],\; \text{and} \\ |\overline{x}| = \mathrm{arity}(R)}} h\big(R^{\mathcal{N}}(\overline{x})\big).$$

If $\overline{x}$ has duplicate entries then we know that $h\big(R^{\mathcal{N}}(\overline{x})\big) = 0$ as $\mathcal{N}$ is non-redundant a.s. Further, if $\overline{x}$ has no duplicate entries then $h\big(R^{\mathcal{N}}(\overline{x})\big) \leq h(\xi_{R(\overline{x})})$ as the distribution of $\xi_{R(\overline{x})}$ is uniform on $\{\top, \bot\}$, and this is the distribution with maximal entropy on $\{\top, \bot\}$.

Therefore, we have

$$\sum_{\substack{R \in L,\; \overline{x} \in [n],\; \text{and} \\ |\overline{x}| = \mathrm{arity}(R)}} h\big(R^{\mathcal{N}}(\overline{x})\big) \;\leq\; \sum_{(R,\overline{x}) \in \mathrm{NR}_{L,[n]}} h(\xi_{R(\overline{x})})$$

$$= \;\mathrm{Ent}(\Xi)(n)$$

$$= \;\mathrm{Ent}(\mu_{\mathrm{un},L})(n),$$

and so $\mathrm{Ent}(\nu)(n) \;\leq\; \mathrm{Ent}(\mu_{\mathrm{un},L})(n)$, as desired. $\qquad\square$

Putting together the previous lemma and proposition we immediately obtain the following bound.

Corollary 4.4. *Let ν be a non-redundant invariant measure on $\mathrm{Str}_L(\mathbb{N})$. Then*

$$\mathrm{Ent}(\nu)(n) \leq \sum_{r \leq n} \binom{n}{r} \cdot r! \cdot \mathfrak{a}_L(r)$$

for all $n \in \mathbb{N}$. In particular, when L is finite with maximum arity k, we have $\mathrm{Ent}(\nu)(n) = O(n^k)$.

Recall from Lemma 1.14 that when L has only finitely many relation symbols of each arity, any non-redundant invariant measure on $\mathrm{Str}_L(\mathbb{N})$ has an $\mathbb{R}$-valued entropy function. Corollary 4.4 improves this by providing an explicit upper bound.

In fact, we do not need the invariant measure to be non-redundant, nor the language to be relational, to obtain the final line of Corollary 4.4; by Proposition 1.21 we see that the polynomial bound $O(n^k)$ holds for the entropy function of an arbitrary invariant measure for a finite language (possibly with constant or function symbols) of maximum arity k, thereby strengthening Lemma 1.12.

An interesting question for future work is to characterize precisely those functions from $\mathbb{N}$ to $\mathbb{R} \cup \{\infty\}$ that can be the entropy function of an invariant measure. By Proposition 1.21, it suffices to consider entropy functions of non-redundant invariant measures for relational languages.

Acknowledgements

References

[Ack15]   N. Ackerman, *Representations of Aut(M)-invariant measures: Part 1*, arXiv e-print 1509.06170v1 (2015).

[AFKrP17]   N. Ackerman, C. Freer, A. Kruckman, and R. Patel, *Properly ergodic structures*, ArXiv e-print 1710.09336 (2017).

[AFKwP17]   N. Ackerman, C. Freer, A. Kwiatkowska, and R. Patel, *A classification of orbits admitting a unique invariant measure*, Ann. Pure Appl. Logic **168** (2017), no. 1, 19–36.

[AFNP16]   N. Ackerman, C. Freer, J. Nešetřil, and R. Patel, *Invariant measures via inverse limits of finite structures*, European J. Combin. **52** (2016), 248–289.

[AFP16]   N. Ackerman, C. Freer, and R. Patel, *Invariant measures concentrated on countable structures*, Forum Math. Sigma **4** (2016), e17, 59pp.

[AFP17]   _____, *Countable infinitary theories admitting an invariant measure*, arXiv e-print 1710.06128v1 (2017).

[Ald81]   D. J. Aldous, *Representations for partially exchangeable arrays of random variables*, J. Multivariate Anal. **11** (1981), no. 4, 581–598.

[Ald85]   _____, *Exchangeability and related topics*, École d'été de probabilités de Saint-Flour, XIII (1983), Lecture Notes in Math., vol. 1117, Springer, Berlin, 1985, pp. 1–198.

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

[AZ86]   G. Ahlbrandt and M. Ziegler, *Quasi-finitely axiomatizable totally categorical theories*, Ann. Pure Appl. Logic **30** (1986), no. 1, 63–82.

[CD13]   S. Chatterjee and P. Diaconis, *Estimating and understanding exponential random graph models*, Ann. Statist. **41** (2013), no. 5, 2428–2461.

[CT06]   T. M. Cover and J. A. Thomas, *Elements of information theory*, 2nd ed., Wiley, Hoboken, NJ, 2006.

[CV11]   S. Chatterjee and S. R. S. Varadhan, *The large deviation principle for the Erdős-Rényi random graph*, European J. Combin. **32** (2011), no. 7, 1000–1017.

[DF81]   P. Diaconis and D. Freedman, *On the statistics of vision: the Julesz conjecture*, J. Math. Psych. **24** (1981), no. 2, 112–138.

[DJ08]   P. Diaconis and S. Janson, *Graph limits and exchangeable random graphs*, Rend. Mat. Appl. (7) **28** (2008), no. 1, 33–61.

[HJS18]   H. Hatami, S. Janson, and B. Szegedy, *Graph properties, graph limits, and entropy*, J. Graph Theory **87** (2018), no. 2, 208–229.

[HN13]   H. Hatami and S. Norine, *The entropy of random-free graphons and properties*, Combin. Probab. Comput. **22** (2013), no. 4, 517–526.

[Jan13]   S. Janson, *Graphons, cut norm and distance, couplings and rearrangements*, New York Journal of Mathematics (NYJM) Monographs, vol. 4, State University of New York, University at Albany, 2013.

[Kal99]   O. Kallenberg, *Multivariate sampling and the estimation problem for exchangeable arrays*, J. Theoret. Probab. **12** (1999), no. 3, 859–883.

[Kal02]   _____, *Foundations of modern probability*, 2nd ed., Probability and its Applications, Springer, New York, 2002.

[Kal05]   _____, *Probabilistic symmetries and invariance principles*, Probability and its Applications, Springer, New York, 2005.

[Kec95]   A. S. Kechris, *Classical descriptive set theory*, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, New York, 1995.

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

[LS10]   L. Lovász and B. Szegedy, *Regularity partitions and the topology of graphons*, An irregular mind, Bolyai Soc. Math. Stud., vol. 21, János Bolyai Math. Soc., Budapest, 2010, pp. 415–446.

[PV10]   F. Petrov and A. Vershik, *Uncountable graphs and invariant measures on the set of universal countable graphs*, Random Structures & Algorithms **37** (2010), no. 3, 389–406.

[RS13]   C. Radin and L. Sadun, *Phase transitions in a complex network*, J. Phys. A **46** (2013), no. 30, 305002.

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY, CAMBRIDGE, MA 02138
Email address: nate@math.harvard.edu

BORELIAN CORPORATION, CAMBRIDGE, MA 02139
Email address: freer@borelian.com

DEPARTMENT OF PHILOSOPHY, HARVARD UNIVERSITY, CAMBRIDGE, MA 02138
Email address: rpatel@fas.harvard.edu