CHAPTER 6

Random Structures

6.1. Graphons

6.1.1. Defining Graphons. We turn from the consideration of structures that, probabilistically, resemble a typical finite structure to structures that intrinsically campture an infinite random process. In that sense, we are following more closely the models of Section 5.1.1 that explicitly follow the construction of a graph. The standard treatments of these approaches are [**306, 139**].

DEFINITION 6.1.1. A random graph model is a sequence of probability distributions $\bar{\sigma} = (\sigma_k : k \in \mathbb{N})$, where σ_k is a distribution of graphs on vertices $\{1, 2, \ldots, k\}$, for each k.

At this point, we have not done much beyond the basic Erdős-Renyi model. Now, however, we would like to link these distributions together in such a way that the passage from k vertices to $k + 1$ vertices represents growth in a random graph process.

DEFINITION 6.1.2. We say that a random graph model $\bar{\sigma}$ is *consistent* if and only if, for each K, the distribution σ_k, interpreted on $\{1, 2, \ldots, k-1\}$, is equidistributed with σ_{k-1}.

Graphs are about adjacency. In most situations where graphs are the appropriate model of a mathematical process, the adjacency between vertices x and y should not depend on the adjacency between w and z, provided that all four are distinct vertices. We capture this with the following definition.

DEFINITION 6.1.3. We say that a random graph model $\bar{\sigma}$ is *local* if and only if, for any disjoint sets $V, W \subseteq \{1, \ldots, k\}$ of vertices, the induced subgraphs on V and W under σ_k are independent.

We first observe that the various models described in Section 5.1.1 are, in fact, consistent local random graph models. Consider first the Erdős-Renyi graph $\mathcal{G}(n, p)$ with constant p. The sequence $\bar{\gamma}_p = \{\mathcal{G}(n, p) : n \in \mathbb{N}\}$ is explicitly a random graph model. $\mathcal{G}(n-1, p)$ is equidistributed with $\mathcal{G}(n, p) \restriction_{\{1,\ldots,n-1\}}$. Moreover, subgraphs on disjoint sets of vertices are independent.

A more subtle example is the preferential attachment model. It is not hard to believe that the description given in Section 5.1.1 gives a probability distribution on each set of vertices. Consistency arises exactly from the step-by-step construction. It is not as easy to see that the model is local, since attachment is dependent on the global degree distribution. We will ultimately show that this model is local, but in an indirect way.

A second reasonable way to model the development of a random graph is given by countable random graph models.

DEFINITION 6.1.4. Let Ω be a locally finite tree of finite graphs up to identity with the empty graph at the root, and where each graph G on vertices $\{1, \ldots, k\}$ is an immediate successor of its induced subgraph on $\{1, \ldots, k-1\}$. Now a *countable random graph model* is a probability measure on the set of paths through Ω which is invariant under permutations of the positive integers.

In this formalism, the consistency is built in via the tree structure. We do still have to specify the localism.

DEFINITION 6.1.5. A countable random graph model is said to be *local* if the induced subgraphs on any two disjoint finite set of vertices are independent.

We observe that these two approaches are interchangeable, in the following sense.

PROPOSITION 6.1.6. *There are transformations Ψ_1 from consistent local random graph models to local countable random graph models and Ψ_2 from local countable random graph models to consistent local random graph models such that $\Psi_1 \circ \Psi_2$ and $\Psi_2 \circ \Psi_1$ are both the identity transformations.*

PROOF. For Ψ_1, we interpret σ_k as a distribution on the kth level of Ω. The Kolmogorov Extension Theorem (Section II.3 of [**407**]) shows that there is a unique probability measure on the set of paths through Ω that restricts to σ_k for each k.

For Ψ_2, we determine σ_k by setting the probability of a graph G on $\{1, \ldots, k\}$ as the probability of the set of paths through G. This random graph model is consistent by construction, and localism is preserved. □

Another natural approach to a graph process, althrough the randomness is less obvious, is to view an infinite graph process as a limit point of a sequence of graphs in a certain metric space whose points are graphs. This is the approach of [**77**]. One natural approach would be to use the "edit distance" between graphs — the number of adjacencies that must be added or removed to get from one to the other. The trouble with this approach is that if we take two independent samples of, for instance, $\mathcal{G}(n, \frac{1}{2})$, the edit distance is likely to be very large; while roughly the same proportion of vertices will be adjacent, the specific vertices which are adjacent can vary quite a lot. On the other hand, since these graphs will be isomorphic with high probability, we prefer a model in which they are likely to be close.

The appropriate metric was introduced in [**182**] and discussed in the present context by [**78**]. It arises by considering cuts in the graph, from the perspective of network flows. We generalize the standard Ford-Fulkerson framework slightly by considering arbitrary pairs of sets of vertices, and we count, in each finite graph G, the number $e_G(S, T)$ of edges joining a vertex in S to a vertex in T. We now define a metric on finite graphs.

DEFINITION 6.1.7. Given two graphs G_0, G_1, we define the cut distance $d_\square(G_0, G_1)$ as follows:

(1) If G_0, G_1 have the same vertex set V,
$$d_\square(G_0, G_1) = \max_{S,T \subseteq V} \frac{|e_{G_0}(S, T) - e_{G_1}(S, T)|}{|V|^2}.$$
(2) If G_0, G_1 have distinct sets of vertices, but with the same cardinality, then we set
$$d_\square(G_0, G_1) = \min_{(\sigma_0, \sigma_1)} d_\square(\sigma_0(G_0), \sigma_1(G_1))$$

where σ_i is a bijection of the vertices of G_i onto the set $\{1, \ldots, n\}$ and $\sigma_i(G_i)$ is interpreted as an isomorphic copy of G_i in the natural way.

(3) If G_i has n_i vertices, we construct, for each natural number m, the graph $G_i(m)$ by replacing each vertex of G_i with m vertices, where vertices of $G_i(m)$ are adjacent if and only if the original corresponding vertices in G_i were adjacent. We then define

$$d_\square(G_0, G_1) = \lim_{k\to\infty} d_\square\left(G_0(kn_1), G_1(kn_0)\right).$$

This cut distance is only a pseudometric. There are graphs $G_0 \neq G_1$ with $d_\square(G_0, G_1) = 0$. With this cut distance, though, we can make sense of limits in the space of finite graphs. In that sense, we can understand a random graph process as a limit point of a convergent sequence of finite graphs. Let $\mathcal{G}$ be the pseudometric space of finite graphs with $d_\square$, and let $\mathcal{G}_\square$ be its completion.

PROPOSITION 6.1.8. *There are transformations Ψ_1 from local countable random graph models to points in $\mathcal{G}_\square$ and Ψ_2 from points in $\mathcal{G}_\square$ to local countable random graph models such that $\Psi_1 \circ \Psi_2$ and $\Psi_2 \circ \Psi_1$ are both the identity transformations.*

PROOF. For Ψ_1, we begin by, for each k, independently generating a random graph G_k according to the distribution σ_k. We consider, for any fixed finite graph, the sequence $t_{ind}(F, G_k)$, where

$$t_{ind}(F, G) = .$$

We can show (making appropriate use of locality and consistency, as well as the Markov-Shebyshev inequality and the Borel-Cantelli Lemma) that this sequence converges almost surely for any fixed F. It follows from the definition of $d_\square$ that this sequence must then be Cauchy in $d_\square$, thus defining a point in the completion.

By contrast, a point in the completion is given by a Cauchy sequence, with two sequences equivalent if they merge to a Cauchy sequence. A Cauchy sequence of finite graphs can be replaced with an equivalent Cauchy sequence $(G_k : k \in \mathbb{N})$ where G_k has vertex set $\{1, \ldots, k\}$, and we can define a local countable random graph model that concentrates on this sequence. □

A fourth perspective on these processes is that of graph parameters. A *signed graph* is a structure in a language with two symmetric binary relations E_+ and E_-, representing the adjacencies and the nonadjacencies, respectively. In this sense, every graph is equivalent to a signed graph.

DEFINITION 6.1.9. A normalized multiplicative graph parameter is a function $f : \mathcal{G} \to \mathbb{R}$, where $\mathcal{G}$ denotes the set of finite signed graphs, satisfying the following additional properties:

(1) If $G_0 \cong G_1$, then $f(G_0) = f(G_1)$.
(2) For any disjoint graphs G_0, G_1, we have $f(G_0 \dot\cup G_1) = f(G_0)f(G_1)$.
(3) If V denotes the graph with one vertex and no edges, then $f(V) = 1$.
(4) For any G, we have $f(G) \geq 0$.

A graph parameter of particular importance for the study of random graphs is the *homomorphism density* defined by

$$t_{hom}(G_0, G_1) = \frac{|\{f : G_0 \to G_1\}|}{|G_1|^{|G_0|}},$$

that is, the probability that a random map from G_0 to G_1 is a homomorphism.

The final approach, that of [**307**] is that of measurable functions. Let $W : [0,1]^2 \to [0,1]$ be a symmetric measurable function. Certainly if the range of W is $\{0,1\}$, there is a natural interpretation of W as a graph structure on the vertex set $[0,1]$. We can interpret symmetric measurable functions more broadly as weighted graphs: If we partition the unit interval by measurable sets $V_1, \ldots, V_k$ in such a way that W is constant on $V_i \times V_j$ for every pair i, j, we can then interpret W as assigning a weight to an edge from vertex i to vertex j.

We could imagine these weights corresponding to probabilities of adjacency, and could imagine a sequence $(W_i : i \in \mathbb{N})$ of these step functions converging to some more general function — or, as in a standard presentation of the Lebesgue integral, an arbitrary symmetric measurable W being approximated by a sequence of step functions. In this sense, a symmetric measurable function $W : [0,1]^2 \to [0,1]$ represents a limit of graphs.

We can define certain graph-theoretic parameters in a natural way on these functions. Importantly, we can define the degree of a vertex x by

$$d(x) = \int_0^1 W(x,y)dy.$$

Similarly, we can define a quantity analogous to homomorphism density. Let G be a finite graph with vertex set V, and denote by dV the form $\prod_{i \in V} dx_i$. We then define the homomorphism density $t_{hom}(G, W)$ by

$$\int_{[0,1]^V} \prod_{iGj} W(x_i, x_j)dV.$$

In the case where W is a $\{0,1\}$-valued step function, this exactly matches the usual definition.

A natural difficulty arises in when to consider two of these measurable functions equivalent. As functions, they are highly sensitive to the labeling of vertices. This can be resolved by the use of measure-preserving transformations, but such transformations are not necessarily invertible. Instead, we define two symmetric measurable functions $W_0, W_1 : [0,1]^2 \to [0,1]$ to be *weakly isomorphic* if for any finite graph G we have $t_{hom}(G, W_0) = t_{hom}(G, W_1)$.

EXAMPLE 6.1.10. The infinite Erdős-Rényi random graph generated as the limit of $\mathcal{G}(n,p)$ can be represented as the function $W : (x,y) \mapsto p$.

We are now ready to consider a foundational result linking all of these perspectives.

THEOREM 6.1.11 (Lovász). *The following are equivalent, in the sense that one may construct, from an item of each class, an item of each other class in such a way that a cycle through these constructions will generate an object equivalent to the starting point:*

(1) *A consistent and local random graph model*
(2) *A local random countable graph model*
(3) *A point in the completion of the metric space of finite graphs*
(4) *A nonnegative normalized multiplicative graph parameter on signed graphs*
(5) *A measurable function $W : [0,1]^2 \to [0,1]$.*

PROOF. We have already established some parts of the proof. The equivalence of parts a and b is Proposition 6.1.6, and the equivalence of parts b and c is Proposition 6.1.8. It remains only to prove the equivalence of a, d, and e.

Given a consistent and local random graph model $\bar{\sigma}$, we construct a measurable function $W_{\bar{\sigma}}$ in this way. We begin by, for each k, independently generating a random graph G_k according to the distribution σ_k. We consider, for any fixed finite graph, the sequence $t_{ind}(F, G_k)$, where

$$t_{ind}(F, G) = .$$

We can show (making appropriate use of locality and consistency, as well as the Markov-Shebyshev inequality and the Borel-Cantelli Lemma) that this sequence converges almost surely for any fixed F. Now we let V_k, E_k be, respectively, the vertex set and the edge relation of G_k, and take an ultrafilter $\mathcal{F}$ on $\mathbb{N}$. We then take the ultraproduct $\prod_{\mathcal{F}}(V_k \times V_k)$. We can identify this product with $\left(\prod_{\mathcal{F}} V_k\right)^2$ in the natural way. We construct a measure by taking μ_k to be the uniform measure on V_k, with σ-algebra $\mathcal{M}_k = \mathcal{P}(V_k)$, and let $\mathcal{M} = \prod_{\mathcal{F}} M_k$, with the product measure μ. We can also construct an appropriate "edge" object on V by defining $E = \prod_{\mathcal{F}} E_n$, and define W_0 to be the conditional expectation of the characteristic function of E, given a pair in $\mathcal{M}^2$. Now for any finite graph F, the ultralimit $\lim_{\mathcal{F}} t_{hom}(F, G_n)$ gives exactly $t(F, W)$, as described above. This measurable function W_0 can be transformed to a measurable function $W : [0,1]^2 \to [0,1]$ with the same property [**155**].

Given a measurable function $W : [0,1]^2 \to [0,1]$ we construct a graph parameter meant to generalize the homomorphism density. We consider the unit interval as the set of vertices in the limit graph represented by W, and we consider $W(x, y)$ as a weight for an edge from x to y in a weighted graph. In that sense, allowing x_i to range over all possible "vertices" of W to which the vertex i might be mapped, the homomorphism density can be defined, for a graph $G = (V, E)$ as

$$t_{hom}(G, W) := \int_{[0,1]^V} \prod_{(i,j)\in E} W(x_i, x_j) \prod_{i\in V} dx_i.$$

The analogy in which W has finite domain and values from $\{0, 1\}$, and in which integration is taken with respect to the uniform measure, gives precisely the same graph parameter as the discrete case. To account for signed graphs, we modify t_{hom} to reflect the preservation of non-adjacencies:

$$t_{ind}(G, W) := \int_{[0,1]^V} \left(\left(\prod_{(i,j)\in E_+} W(x_i, x_j) \right) \left(\prod_{(i,j)\in E_-} (1 - W(x_i, x_j)) \right) \right) \prod_{i\in V} dx_i.$$

Observe that $t_{ind}(\cdot, W)$ is a normalized multiplicative graph parameter.

Given any normalized multiplicative graph parameter f on signed graphs, we will now produce a consistent local random graph model. Given a graph G on k vertices, let $\widehat{G}$ be the signed graph equivalent to the graph G. Now as G ranges over all graphs on k vertices, the map $\sigma_k : G \mapsto f(\widehat{G})$ gives a probability distribution on these graphs. We can calculate that $\sigma_k(G) = \sum \sigma_{k+1}(H)$, where H ranges over all one-vertex extensions of G, so that $\bar{\sigma}$ is consistent. Moreover, since f is

multiplicative (for disjoint S, T, we have $f(S \cup T) = f(S)f(T)$), we also see that $\bar{\sigma}$ is local. □

In view of this theorem, any of the five equivalent concepts has equal claim to the name, but the standard formulation in the literature is this:

DEFINITION 6.1.12. A *graphon* is a measurable function $W : [0,1]^2 \to [0,1]$.

One could also wish to infer, from a graph observed, the random process that gave rise to it. The literature in this direction is young, but a recent paper gives an algorithm to estimate a graphon W from a single graph sampled from it [**87**].

An initial test of the potential of graphons to capture the model-theoretic randomness of structures is their interaction with finite fields. In view of Theorems 5.3.14 and 5.3.17, it is interesting to explore the graphs definable in asymptotic classes of finite fields, and recent work has followed this lead.

THEOREM 6.1.13 ([**150**]). *Let $\Gamma = (U, V, E)$ be a definable bipartite graph over a definable set S_0 on an asymptotic class of finite structures $\mathcal{C}$. Then there exists a definable set S over S_0 and a definable step function W over S such that for every $\epsilon > 0$ there is $M > 0$ such that for every $F \in \mathcal{C}$ with $|F| \geq M$, every $x \in S(F)$ mapping onto $x_0 \in S_0(F)$, we have*

$$d_{\square}\left(\Gamma_{x_0}(F), W_x(F)\right) \leq \epsilon.$$

This result will be seen in Section 8.7.1 to be a version of Szemeredi's Regularity Lemma, which will be discussed in more detail there. Džamonja and Tomašić prove this reult using Theorem 5.3.9.

6.1.2. Invariant Measures. An interesting recent approach to random structures in recent years starts with the matter of measures on the space of countable structures. One sometimes says that a construction is *exchangeable* if it is invariant under permutations of the elements. This is certainly true, for instance, of the random graphs $\mathcal{G}(n, M)$ and $\mathcal{G}(n, p(n))$ from the previous chapter.

Consider a countable language L, and the space $\mathbb{M}_L$ of all L-structures with universe $\mathbb{N}$. For each quantifier-free formula $\varphi(x_1, \ldots, x_k)$ of L and each k-tuple $(n_1, \ldots, n_k)$ of natural numbers, we have a set $B_{\varphi(\bar{n})}$ of structures defined by

$$B_{\varphi(\bar{n})} := \{\mathcal{M} : \mathcal{M} \models \varphi(\bar{n})\}.$$

Now $\mathbb{M}_L$ is naturally a topological space in this so-called "logic topology," in which the sets $B_{\varphi(\bar{n})}$ are taken as the subbasic open sets. The *logic action* of S_∞ on $\mathbb{M}_L$ is the action in which an element of S_∞ acts in the natural way as a permutation of the universe, and the orbit of a structure under this action is exactly its isomorphism type.

DEFINITION 6.1.14. Let μ be a Borel probability measure on $\mathbb{M}_L$.

(1) We say that μ is *invariant* when it is invariant under the logic action.
(2) We say that μ *concentrates* on a set X when $\mu(X) = 1$.
(3) We say that μ *concentrates* on a structure $\mathcal{M}$ when it concentrates on the orbit of $\mathcal{M}$ under the logic action.

The uniform measure on the set of graphs with universe (that is, vertex set) $\mathbb{N}$ concentrates on the uniform countable random graph, as we saw in the previous chapter. We also saw that the uniform measure on triangle free graphs does not

concentrate on the Henson graph (the Fraïssé limit of triangle free graphs), although there is a known invariant measure that does. A significant recent body of work by Ackerman, Freer, Patel, and others investigates conditions on a structure $\mathcal{M}$ under which there exists an invariant probability measure that concentrates on $\mathcal{M}$.

We begin by defining some conditions that will be useful in the characterization.

DEFINITION 6.1.15. Let $\mathcal{M}$ be a countable L-structure.

(1) Let $\bar{a}$ be a finite subset of $\mathcal{M}$. The $L_{\omega_1\omega}$ *definable closure* of $\bar{a}$ in $\mathcal{M}$ (denoted $dcl(\bar{a})$)is the set of $b \in \mathcal{M}$ such that for any automorphism g of $\mathcal{M}$ with $g(a) = a$ for all $a \in \bar{a}$, we have $g(b) = b$.
(2) We say that $\mathcal{M}$ has *trivial definable closure* if and only if for every finite $\bar{a} \subseteq \mathcal{M}$, we have $dcl(\bar{a}) = \bar{a}$.

In the present section we oftentimes do not write that the definable closure in question is $L_{\omega_1\omega}$ definable closure, but we will have this meaning in mind throughout the present section.

THEOREM 6.1.16 ([**5**]). *Let L be a countable language and $\mathcal{M}$ a countable infinite L-structure. Then the following are equivalent:*

(1) *There is an invariant probability measure on $\mathbb{M}_L$ that concentrates on $\mathcal{M}$.*
(2) *The structure $\mathcal{M}$ has trivial definable closure.*

PROOF. The key element of the proof is in constructing a Borel structure with appropriate properties. We say that an L-structure $\mathcal{W}$ is a Borel structure if its universe is $\mathbb{R}$ and if, in addition, for each n-ary relation symbol $R \in L$, the set $R^{\mathcal{U}}$ is a Borel set of $\mathbb{R}^n$ and for each n-ary function symbol $f \in L$, the graph of $f^{\mathcal{U}}$ is a Borel subset of $\mathbb{R}^{n+1}$. These structures have been described in [**421, 345**].

In such a structure, we might hope to sample, as we sample the unit interval with graphons, to produce a countable sequence of elements of $\mathcal{W}$ that constitute the universe for a countable L-structure. We define $Clo(\mathcal{W})$ to be the set of countable sequences $\tau \subseteq \mathcal{W}$ where τ, as a set, contains all constants of $\mathcal{W}$ and is closed under all functions of $\mathcal{W}$. We say that $\mathcal{W}$ is *samplable* if and only if $Clo(\mathcal{W})$ is exactly the set of all countable sequences of elements of $\mathcal{W}$.

The main idea is to pass from the structure $\mathcal{M}$ to a samplable Borel structure $\mathcal{W}$ that gives rise to the appopriate measure. A theory T is said to admit *duplication of quantifier-free types* when for every complete $L_{\omega_1\omega}$ type $p(x, \bar{z})$ of distinct elements consistent with T there exists a complete $L_{\omega_1\omega}$ type $q(x, y, \bar{z})$ of distinct elements that implies boty $p(x, \bar{z})$ and $p(y, \bar{z})$. If every countable model of T has trivial definable closure, then T has duplication of quantifier free types, since for any type over $\bar{z}$, the type p cannot define x, so that there is $y \neq x$ also satisfying $p(y, \bar{z})$.

We will pass from the structure $\mathcal{M}$ to another structure $\overline{\mathcal{M}}$ representing $\mathcal{M}$. Let $\sim_k$ be the relation on k-tuples of elements from $\mathcal{M}$ denoting that they are in the same automorphism class. We establish a language $L_{\overline{\mathcal{M}}}$ consisting of a k-ary relation symbol for each $\sim_k$ equivalence class for each k. The structure $\overline{\mathcal{M}}$ is an $L_{\overline{\mathcal{M}}}$ structure interpreting each relation symbol as the appropriate equivalence class. It is not difficult to believe that $\mathcal{M}$ and $\overline{\mathcal{M}}$ are $L_{\omega_1\omega}$-interdefinable, or that $\overline{\mathcal{M}}$ has trivial definable closure if and only if $\mathcal{M}$ does. It is also true, although less obvious, that there is an invariant measure concentrating on $\overline{\mathcal{M}}$ if and only if there is one for $\mathcal{M}$.

The advantage of $\overline{\mathcal{M}}$ is that there is a theory $T_{\overline{\mathcal{M}}}$ axiomatized in a highly regular way whose models are exactly the isomorphic copies of $\overline{\mathcal{M}}$. The necessary condition

is that the axioms for $T_{\overline{\mathcal{M}}}$ have the property that Ackerman, Freer, and Patel call *pithy*: They are (possibly infinitary) Π_2 sentences of the form $\forall\bar{x}\exists y\varphi(\bar{x},y)$, with φ quantifier free.

We then construct a samplable Borel $L_{\overline{\mathcal{M}}}$-structure $\mathcal{W}$ such that for every probability measure μ giving every nonempty open set positive measure, and for every sentence $\forall\bar{x}\exists y\varphi(\bar{x},y) \in T_{\overline{\mathcal{M}}}$ and every $\bar{a} \in \mathcal{W}$, either $\mathcal{W} \models \varphi(\bar{a},b)$ for some $b \in \bar{a}$ or $\mu\{b \in \mathcal{W} : \mathcal{W} \models \varphi(\bar{a},b)\} > 0$. This requires careful assignment of predicate values and tracking of intervals, but the key element is that whenever there is no i such that $\varphi(\bar{a},a_i)$, an interval of y is added, each element satisfying $\varphi(\bar{a},y)$.

Now to get a measure λ concentrated on $\overline{\mathcal{M}}$, we define a function F mapping $Clo(\mathcal{W})$ to the set of $L_{\overline{\mathcal{M}}}$-structures with universe $\mathbb{N}$ in the natural way: given an infinite sequence $\tau \in Clo(\mathcal{W})$, we define a structure $F(\tau)$ on the natural numbers with $R^{F(\tau)}(n_1,\ldots,n_j)$ if and only if $\mathcal{W} \models R(\tau(n_1),\ldots\tau(n_j))$, and similarly for function and constant symbols. Given a probability measure μ on $\mathbb{R}$, we define a measure $\mu_{\mathcal{W}}$ on the set of countable $L_{\overline{M}}$-structures by composing F^{-1} with the product measure on $\mathbb{R}^\omega$. This measure is invariant.

Let $n_1,\ldots,n_k \in \mathbb{N}$, and let $\forall\bar{x}\exists y\varphi(\bar{x},y)$ be an axiom of $T_{\overline{\mathcal{M}}}$, considering each n_i as $\tau(n_i)$ for some fixed $\tau \in Clo(\mathcal{W})$. Now either $\mathcal{W} \models \varphi(\bar{n},n_i)$ for some i or there is a positive measure set of y such that $\mathcal{W} \models \varphi(\bar{n},y)$. In either case, we can show that, $\mu_{\mathcal{W}}$-almost surely, we have $F(\tau) \models \exists y\varphi(\bar{n},y)$, so that almost surely, $F(\tau) \models T_{\overline{\mathcal{M}}}$.

It follows that $\mu_{\mathcal{W}}$ concentrates on the models of $T_{\overline{\mathcal{M}}}$, that is, on isomorphic copies of $\overline{\mathcal{M}}$. By interdefinability of $\overline{\mathcal{M}}$ and $\mathcal{M}$, we have a measure concentrating on $\mathcal{M}$.

For the reverse implication, suppose that there is an invariant measure μ concentrated on $\mathcal{M}$, suppose for contradiction that $b \neq \bar{a}$ is in $dcl(\bar{a})$. Let $p(\bar{x},y)$ be an $L_{\omega_1\omega}$ formula isolating the type of $\bar{a}b$, and note that μ must concentrate on the set of structures that satisfy $\exists\bar{x},y[p(\bar{x},y)]$. Since μ is countably additive, there must be some tuple $\bar{n},m$ of natural numbers that satisfies p in a positive-measure set of structures. We consider the family of sets of structures

$$\Gamma_j := \{\mathcal{M} : \mathcal{M} \models p(\bar{n},j)\}.$$

Since b is in the definable closure of a, these sets must be disjoint, and since μ is invariant, they must all have the same finite positive measure. The measure of their union, then, must be infinite, a contradiction. □

From the perspective of this result, long-known amalgamation results concerning the Henson graph prove that the generic triangle-free graph has trivial definable closure, so that this result generalizes the result of Petrov and Vershik showing that there is an invariant measure under which that graph is almost sure.

Ideologically, of course, the idea of sampling from a samplable Borel structure to get a random structure is familiar from graphons, but it is not entirely straightforward to say what the connection is. Let $\mathcal{W}$ be a Borel graph defined on the unit interval. We define a grapon $W_{\mathcal{W}}$ by setting $W_{\mathcal{W}}(x,y) = 1$ if $\mathcal{W} \models xEy$, and $W_{\mathcal{W}}(x,y) = 0$ otherwise. Now the distribution of a random graph sampled from $W_{\mathcal{W}}$ is exactly $\mu_{\mathcal{W}}$ as in the proof above. Similarly, if W is a graphon taking values in $\{0,1\}$, we can define a Borel graph $\mathcal{W}_W$ corresponding to it.

The theorerm also shows that if a countable graph admits an invariant measure, then there is some samplable Borel structure from which that measure is the sampling measure. Consequently, if W is a graphon such that random graphs sampled from W concentrate on a single isomorphism type, then there is a $\{0,1\}$-valued graphon W' such that random graphs sampled from W' are also concentrated on that same type. This connection is explored in more detail in [**4**].

In [**4**], a condition is given similar to duplication of types under which an inverse limit construction on finite structures can construct a measure concentrated on models of a single theory, but assigning measure zero to each isomorphism type.

Recall from Chapter 2 that Gaifman proposed in 1964 an idea of measure models. In this context, he asked for conditions under which it is possible for a theory to have a measure model with strict equality and invariant under finite permutations of the domain [**189**]. This problem has recently been solved. Note that while Gaifman's original statement of the problem specified invariance under finite permutations, by the Kolmogorov Extension Theorem (Theorem II.3.4 of [**407**]), this is equivalent to invariance under the full logic action.

Definition 6.1.17. Let G act on X, and let μ be an invariant probability measure. Then the system (G, X, μ) is said to be *ergodic* if and only if for every $V \subseteq S$ with $g(x) \in V$ for every $g \in G$ and μ-almost every $x \in V$, we have $\mu(V) \in \{0,1\}$

Note that this definition is concordant with the standard definition in [**456**] of an ergodic measure-preserving transformaiton. For Walters, it is the transformation that is ergodic, and it is important that the invariant set have measure 1 or 0. From that perspective, the "almost every" quantifier is of no importance, since a difference of a measure 0 set will not change the provision that V has measure 1 or 0. Similarly, it does not matter whether we consider the forward or inverse image of elements of V: we specified that G is a group. For brevity, one usually says that either the group action or the measure is ergodic, although, of course, it is a property of the pair.

Theorem 6.1.18 ([**6**]). *Let Σ be a countable set of sentences in $L_{\omega_1\omega}$. Then the following are equivalent.*

1. *There is an invariant probability measure concentrated on models of Σ*
2. *There is an ergodic invariant probability measure concentrated on the models of Σ.*
3. *There is a countable fragment A of $L_{\omega_1\omega}$ and a consistent theory $T \subseteq A$ complete for A and extending Σ such that T has trivial A-definable closure.*
4. *For all countable fragments A of $L_{\omega_1\omega}$ containing Σ, there is a consistent theory $T \subseteq A$ complete for A and extending Σ such that T has trivial A-definable closure.*

Proof Outline. We explain only the proof that the existence of an invariant measure implies the existence of an ergodic invariant measure. Let Σ be a countable set of sentences such that there is an invariant measure μ concentrated on models of Σ, and suppose that there is no ergodic measure concentrated on Σ.

Now we can decompose μ as a mixture of ergodic measures. This decomposition is standard (see, for instance, [**456, 256**]), but may not be familiar to many readers of the present volume, so we give some idea of it. Now there exists an invariant set

I of positive measure in the sense of μ, and we have

$$\mu(A) = \mu(I)\mu(A|I) + \mu(\neg I)\mu(A|\neg I).$$

Now these conditional measures are invariant, so that μ is a nontrivial convex combination of invariant measures. We want, in the end, to have these measures ergodic. By Choquet's Theorem [**366**] there is a decomposition into ergodic measures.

Since none of the measures in the decomposition is concentrated on the models of Σ, and since they are ergodic, they must all be concentrated away from the models of Σ. But then μ is concentrated away from the models of Σ, a contradiction. □

It is reasonable to ask about the number of distinct probability measures invariant under the logic action. In connection with the decomposition just described, it is a standard result that the set of probability measures on a fixed space X invariant under a fixed group action constitute a compact convex subset of the full space of probability measures on X, and that the extremal points of this convex set are exactly the ergodic measures. Naturally, then, if there is a unique invariant measure, it must be ergodic.

More is true, though. By the Birkhoff Ergodic Theorem, for any ergodic system (G, X, μ) there is a measure 1 set $Y \subseteq X$ such that for all $x \in Y$, all $\sigma \in G$, and all continuous real-valued functions f on X, we have

$$\frac{1}{n}\sum_{i=0}^{n-1} f\left(\sigma^i(x)\right) \to \int_X f d\mu.$$

In the case of a unique invariant measure, this convergence is uniform, and holds for all $x \in X$, not merely almost everywhere.

From this perspective, it is important to understand the diversity of the class of invariant ergodic measures, since they compose a convex basis for the class of all invariant measures. The following result not only revines the previous results on invariant measures, but also classifies the number of invariant measures concentrating on a structure. We first define a condition used in the theorem.

Definition 6.1.19. Let $\mathcal{M}$ be a structure. We say that $\mathcal{M}$ is t-homgeneous if and only if $Aut(\mathcal{M})$ acts transitively on the set of t-element subsets of $\mathcal{M}$. We say that $\mathcal{M}$ is *highly homogeneous* if and only if it is t-homogeneous for all t.

An older result of Cameron classifies the highly homogeneous structures up to interdefinability [**93**]. There are five.

Theorem 6.1.20 ([**3**]). *Let $\mathcal{M}$ be a countable structure in a countable language. Then exactly one of the following holds:*

1. *$\mathcal{M}$ has nontrivial $L_{\omega_1\omega}$-definable closure*
2. *$\mathcal{M}$ is highly homogeneous and there is a unique invariant measure concentrating on $\mathcal{M}$.*
3. *There are $2^{\aleph_0}$ distinct ergodic invariant probability measures concentrating on $\mathcal{M}$.*

Proof. We already know that existence of an invariant measure concentrating on $\mathcal{M}$ is equivalent to the failure of 1. The examples of highly homogeneous structures are all known to have trivial definable closure.

In the highly homogenous case, uniqueness follows by considering the number α_n of distinct quantifier-free types of n-tuples realized in $\mathcal{M}$. There are less than n

such types, since $\mathcal{M}$ is highly homogeneous, so that the set of structures in which a fixed tuple realizes each type has measure $\frac{1}{\alpha_n}$. Since these sets generate the Borel subsets of the orbit of $\mathcal{M}$, this uniquely determines the measure.

It remains to show that a structure which is not highly homogeneous but has trivial definable closure has continuum many measures. Certainly there is an invariant measure that concentrates on $\mathcal{M}$. The structure $\overline{\mathcal{M}}$ constructed in the proof of Theorem 6.1.16 is ultrahomogeneous (that is, every isomorphism of finitely generated substructures extends to a full automorphism), so it suffices to consider $\mathcal{M}$ ultrahomogeneous. We consider the structure $\mathcal{W}$ constructed in the proof of Theorem 6.1.16, and modify the measure $\mu_{\mathcal{W}}$. Consider A_0, A_1, non-isomorphic n-element substructures of $\mathcal{M}$, and let φ_i be a quantifier-free formula satisfied by A_i but not in A_{1-i}.

Now let $\mathcal{I} = (U_1, \ldots, U_k)$, where each U_i is a union of finitely many half-open intervals, and let $u : \mathcal{I} \to \mathbb{R}^{>0}$ be a function satisfying $\sum_{i=1}^{k} u(U_i) = 1$. We can then modify $\mu_{\mathcal{W}}$ by setting, for each Borel set $B \subseteq \mathbb{R}$,

$$\mu_{\mathcal{W}}^{\mathcal{I}}(B) = \sum_{i=1}^{k} u(U_i)\mu(B|U_i).$$

This new measure $\mu_{\mathcal{W}}^{\mathcal{I}}$ is an invariant measure concentrated on the orbit of $\mathcal{M}$. It is possible to achieve continuum many distinct measures in this way as $(\mathcal{I}, u)$ varies. □

6.1.3. Entropy Methods for Graphons and Invariant Measures. The language of statistical mechanics has long been a part of the study of large random graphs. In physical terms, this area studies the rise of macroscopic (thermal) properties of systems from the states of their microscopic constituent particles and subsystems. Of course, there are limits to the analogy of, for instance, the large finite number of molecules in a macroscopic sample of a gas on one hand, and the truly infinite number of elements of a countable structure on the other. However, the way in which this analysis is made is generally by treating the molecules *as if* there are infinitely many, for instance, by approximating a summation by an integral. Certainly this approximation must have a role in the program Anil Nerode has talked about since at least 2011 in the formulation of controls for statistical mechanical and quantum systems.

On a more directly falsifiable level, a clear entry of physical language into the study of random structures came with the observation of what Chung and Lu call "the rise of the giant component." Stated simply, if $p(n) \in o\left(\frac{1}{n}\right)$, so that $np > 1$, a random graph drawn from $\mathcal{G}(n,p)$ will, with high probability, have a component whose size is a constant fraction of n. If $np < 1$, on the other hand, all components have size $O(\log n)$. The case where $np = 1$ is a boundary case, at which the macroscopic property of the existence of a giant component changes abruptly. Such situations are well-known in the physical context, and are known as *phase transitions*. Erdős and Rényi did not use this language, but did describe "threshold functions," with essentially similar meaning. Certainly by 1970, Stepanov makes reference to this phenomenon as a phase transition [**423**].

Certainly in the 1970s, entropy bounds were used in probabilistic method estimates to construct expander graphs [**368, 110**]. We have already seen a dynamical

treatment of graphons, in the sense that a graphon constitutes a size continuum Borel structure from which we sample to get a countable structure, with a probability distribution invariant under the logic action. In this context, entropy has well-established meaning. Since entropy is not often presented in a clear elementary way, particularly to logicians, we explain a little here exactly what is being talked about.

Let (X, μ) be a probability space and $\mathcal{P} = \{A_i : i \leq n\}$ a partition of X, with $p_i := \mu(A_i)$. The entropy of $\mathcal{P}$, denoted $H(\mathcal{P})$ is intended to be a measure of the information gained by performing an experiment and observing which A_i occurs, and really depends only on the probabilities p_i, for which reason we frequently write $H(p_1, \ldots, p_n)$. Naturally, if there is i with $p_i = 1$, no information is gained by such an observation. It is also reasonable to require that any such measurement of information should be invariant under permutation of the indices and the addition of a probability zero part, and that it should be continuous in each p_i.

There are two other conditions that are less obvious. If we understand entropy as the information gained by making the observation, it is at least plausible that the maximum information would be in the case as far from concentration as possible — the case where $p_i = p_j$ for all i, j.

Finally, if $\mathcal{P}_0 = \{A_{0,i} : i \leq n_0\}$ and $\mathcal{P}_1 = \{A_{1,i} : i \leq n_1\}$, with $p_{i,j} := \mu(A_{i,j})$, there is a natural partition $\mathcal{P}_0 \vee \mathcal{P}_1$ consisting of the $n_0 n_1$ intersections $A_{0,i} \cap A_{1,j}$. We define, as an intermediate notion, the *conditional entropy* $H(\mathcal{P}_1|\mathcal{P}_0)$ by the follwoing relation:

$$H(\mathcal{P}_1|\mathcal{P}_0) = \sum_{j=1}^{n_2} \left(p_{0,j} H\left(\frac{\mu(A_{1,1} \cap A_{0,j})}{p_{0,j}}, \frac{\mu(A_{1,2} \cap A_{0,j})}{p_{0,j}}, \cdots \frac{\mu(A_{1,n_1} \cap A_{0,j})}{p_{0,j}} \right) \right).$$

In view of our interpretation of H as information gain, we will require that the information gain from the joint observation $\mathcal{P}_0 \vee \mathcal{P}_1$ should be the informaiton gain from $\mathcal{P}_0$, plus the information gain from $\mathcal{P}_1$ given the result of $\mathcal{P}_0$. Formulaically, we require

$$H(\mathcal{P}_0 \vee \mathcal{P}_1) = H(\mathcal{P}_0) + H(\mathcal{P}_1|\mathcal{P}_0).$$

The following theorem is standard, and is found, for instance, in [**275**].

THEOREM 6.1.21. *Let H be a real-valued function defined on all finite strings of nonnegative reals $(p_1, \ldots, p_n)$ such that $\sum_{i=1}^{n} p_i = 1$, and let H satisfy the following conditions:*

1. *For any sequence $\bar{p}$, we have $H(\bar{p}, 0) = H(\bar{p})$.*
2. *H is continuous in all its arguments.*
3. *$H(\bar{p})$ takes its maximum value where $p_i = p_j$ for all i, j.*
4. *Again recalling that H is defined on partitions by virtue of depending only on the associated probabilities, for any partitions $\mathcal{P}_0, \mathcal{P}_1$, we have*

$$H(\mathcal{P}_0 \vee \mathcal{P}_1) = H(\mathcal{P}_0) + H(\mathcal{P}_1|\mathcal{P}_0).$$

Then there is some positive real λ such that $H(p_1, \ldots, p_n) = -\lambda \sum_{i=1}^{n} p_i \log p_i$.

PROOF. We define $H^n_{max} := H\left(\frac{1}{n}, \ldots, \frac{1}{n}\right)$, and will first establish that $H^n_{max} = \lambda \log n$. Indeed, since H^{n+1}_{max} is, by assumption, maximal for $(n+1)$-tuples, $H^n_{max} = H\left(\frac{1}{n}, \cdots, \frac{1}{n}, 0\right) \leq H^{n+1}_{max}$ for any n, so that H^n_{max} is increasing in n.

Now if we select partitions $\mathcal{P}_1, \ldots \mathcal{P}_m$ which are pairwise independent in the obvious sense, we apply condition 4 to see that $H\left(\mathcal{P}_1 \vee \mathcal{P}_2 \vee \cdots \vee \mathcal{P}_m\right) = \sum_{i=1}^{m} H(\mathcal{P}_i)$, and if each of $\mathcal{P}_i$ has k parts with uniform distribution as in condition 3, we then have $H\left(\mathcal{P}_1 \vee \mathcal{P}_2 \vee \cdots \vee \mathcal{P}_m\right) = mH_{max}^k$. However, $\mathcal{P}_1 \vee \mathcal{P}_2 \vee \cdots \vee \mathcal{P}_m$ consists of k^m parts with uniform distribution, so that $H_{max}^{k^m} = mH_{max}^k$.

Now consider integers k_0, k_1, m, n such that

$$k_0^m \leq k_1^n \leq k_0^{m+1}.$$

In this case, we will have

$$m \log k_0 \leq n \log k_1 \leq (m+1) \log k_0,$$

so that

$$\frac{m}{n} \leq \frac{\log k_1}{\log k_0} \leq \frac{m+1}{n}.$$

A similar argument, from the observation that H_{max}^n is increasing, allows us to observe that

$$\frac{m}{n} \leq \frac{H_{max}^{k_1}}{H_{max}^{k_0}} \leq \frac{m+1}{n}.$$

This shows that, for arbitrary values of n, we have

$$\left| \frac{H_{max}^{k_1}}{H_{max}^{k_0}} - \frac{\log k_1}{\log k_0} \right| \leq \frac{1}{n},$$

leaving $H_{max}^k = \lambda \log k$, as desired.

The problem of the non-uniform case remains. We show that H is determined on rational arguments, and reason from continuity to the full theorem. Given $\mathcal{P} = (A_1, \ldots, A_n)$, with $\mu(A_i) = p_i = \frac{\ell_i}{q}$. We refine $\mathcal{P}$ to $\mathcal{P}' = (A_{i,j} : i \leq n, j \leq \ell_i)$, where $\mu(A_{i,j}) = \frac{1}{q}$ for all i, j. Now

$$H(\mathcal{P}'|\mathcal{P}) = \left(\sum_{i=1}^{n} p_i H_{max}^{\ell_i} \right) = \left(\sum_{i=1}^{n} p_i \lambda \log \ell_i \right).$$

Using product-to-sum properties of logarithms, we convert this to

$$H(\mathcal{P}'|\mathcal{P}) = \lambda \left(\log q + \sum_{i=1}^{n} p_i \log p_i \right).$$

Since $\mathcal{P}'$ is a refinement of $\mathcal{P}$, we know that $H(\mathcal{P} \vee \mathcal{P}') = H(\mathcal{P}') = H_{max}^q$. On the other hand, by condition 4, we have

$$H(\mathcal{P} \vee \mathcal{P}') = H(\mathcal{P}) + H(\mathcal{P}'|\mathcal{P}).$$

We can then derive

$$H(\mathcal{P}) = -\lambda \sum_{i=1}^{n} p_i \log p_i$$

as desired. □

Similarly, from a physical perspective, if we consider a system of k particles (with some very large k), each with six degrees of freedom (three each of position and momentum), the energy of the system is a function of all $6k$ degrees of freedom, and its values are probabilistic, with distribution $\rho(E)$ and mean $\overline{E}$. We define a

quantity $\Delta\Gamma$ as the volume of the $6k$-dimensional phase space such that $\rho\left(\overline{E}\right)\Delta\Gamma = 1$. We then define $H = \log\Delta\Gamma$.

Now since $\log\Delta\Gamma = -\log\rho\left(\overline{E}\right)$, we can expand this expression as

$$H = -\sum \rho(E_n)\log\rho(E_n),$$

exactly mathcing the previous argument. Also, we have

$$dH = \frac{d\Delta\Gamma}{\Delta\Gamma},$$

another familiar presentation in thermodynamics.

Now let (X,μ) be a measure space and $T : X \to X$ a measure-preserving transformation. Let $\mathcal{A}$ be a σ-algebra of measurable sets. We denote by $T^{-n}(\mathcal{A})$ the σ-algebra $\{T^{-n}(A) : A \in \mathcal{A}\}$. Now if we have a partition $\mathcal{P} = \{A_0,\ldots,A_k\}$ of X, we can also partition X, for each n, by $\mathcal{P}_n^T$, consisting of all sets of the form

$$\bigcap i = 0^{n-1} T^{-i}(A_{i_j}).$$

That is, two elements a_0, a_i are in the same part of $\mathcal{P}_n^T$ if and only if a_0 and a_1 are in the same part of $\mathcal{P}$, and also $T^{-1}(a_0)$ and $T^{-1}(a_1)$ are in the same part of $\mathcal{P}$, and so on. We define

$$h(T,\mathcal{P}) := \lim_{n\to\infty} H(\mathcal{P}_n^T),$$

and define $h(t)$ to be the supremum over all partitions $\mathcal{P}$. It is not obvious, but true, that the limit $h(T,\mathcal{P})$ must always exist, so that this quantity is well-defined.

Now in Pinsker's original proof of the existence of expander graphs, and in later work in this tradition, for instance, by Chung, entropy seems to be used primarily to express the tradeoffs available in the parameters for expander construction, without much explicit mention of its information-theoretic content. For instance, there is α such that for sufficiently large N, a graph selected at random from all D-regular bipartite graphs with N vertices in each part will have the property that for all sets S of at most αN vertices, the neighborhood of that set is of size at least A times the size of S with high probability if and only if

$$D > \frac{h_2(\alpha) + h_2(\alpha A)}{h_2(\alpha) - \alpha A h_2(1/A)},$$

where $h_2(z) = z\log z + (1-z)\log(1-z)$, with the convention that $0\log 0 = 0$ (which is true, in the limit), as is described in [**446**].

More recently, however, entropy has received more center-stage attention in the study of random graphs. Given a graphon $W : [0,1]^2 \to [0,1]$, we define the entropy of W as

$$H(W) = \int_0^1 \int_0^1 h_2(W(x,y))dxdy.$$

Janson proved that this definition, in the limit, matches the entropy

$$\frac{H\left(G(n,W)\right)}{\binom{n}{2}},$$

in the sense of taking $G(n,W)$ as a random graph drawn from W with finitely many outcomes.

THEOREM 6.1.22 ([**250**]). *Let W be a graphon. Then $ran(W) = \{0,1\}$ if and only if $H(W) = 0$.*

PROOF. The integral defining $H(W)$ is zero if and only if $h_2(W(x,y)) = 0$ for almost every (x,y). Now $h_2(W(x,y)) = 0$ if and only if $W(x,y) \in \{0,1\}$. □

This result shows that the entropy $H\left(G(n,W)\right)$ of a $\{0,1\}$-valued graphon is $o(n^2)$. Hatami and Norine showed that this bound can be made tight, in the sense that there are $\{0,1\}$-valued graphons with asymptotic entropy arbitrarily close to n^2. However, some conditions push this asymptotic entropy lower. If Q is a set of graphs, closed under isomorphism, such that every convergent sequence of graphs in Q converges to a graphon with range $\{0,1\}$. Then $H\left(G(n,W)\right)$ is $O(n \log n)$ [**224**].

One emerging concern about random graph models has been their bias as estimators. In many applications, actual simulation of the random process of generating a large random graph has become a practical part of science. For instance, to perform a hypothesis test on network observations, one could simulate many random graphs with the appropriate properties, and then observe the distribution of their properties. Of course, the quality of such a simulation matters.

One statistical approach to the quality of simulations is through the notion of bias. If we take a sample with measurements $\bar{a}$ and use the function δ evaluated on $\bar{a}$ to estimate the value of a parameter θ, then an important property of δ as an estimator of θ is the difference between the expected value of $\delta(\bar{a})$ as $\bar{a}$ is drawn at random from the true value of θ. This difference is known as the *bias* of δ, and δ is said to be *unbiased* if the bias is zero. A standard example is the sample variance as an unbiased estimator of population variance.

Some have suggested that a random graph model with maximum entropy relative to the prescribed properties would be unbiased estimators [**449**], although I have not seen a proof of this property, or even a precise formulation of what it should mean in terms of the parameter being estimated.

Radin and Sadun have shown that graphons exhibiting minimal entropy subject to appropriate constraints also optimize for small triangle density [**373**].

In view of the characterization in Theorem 6.1.22 and the characterization of the graphons arising as sampling from Borel structures, studying the asymptotic growth of the entropy function is relevant to the general study of random structures. As we have seen, in the case of graphs (in a language with a relation of arity 2), the growth is approximately n^2. More recent work has generalized that insight [**7**].

We first note that the investigation of entropy functions on random structures can be restricted to a special case.

We say that an L-structure $\mathcal{A}$ is *non-redundant* when for any relation symbol $R \in L$, the structure satisfies R only on tuples of distinct elements. We say that an invariant measure is non-redundant when it concentrates on non-redundant structures. Since any theory T is quantifier-free interdefinable with a non-redundant relational theory T', and any invariant measure concentrating on T gives rise in a natural way to an invariant measure concentrating on T', it suffices to consider that case.

DEFINITION 6.1.23. Let L be a relational signature. An *extended L-hypergraphon* is a measurable function

$$W : [0,1]^{[\{1,\ldots,k\}]^{<k}} \to \mathcal{M},$$

where $\mathcal{M}$ is the space of probability measures on the set of non-redundant quantifier-free L-types, such that for any sequence $\bar{x} \in [0,1]^{[\{1,\ldots,k\}]^{<k}}$ and any permutation

σ of the natural numbers, the measure $W\left(x_{\sigma(i_1)}, \ldots x_{\sigma(i_j)}\right)$ is equal to the measure $W\left(x_{i_1}, \ldots, x_{i_j}\right)$, translated by σ.

If L is the signature for graphs, an extended L-hypergraphon is not exactly a graphon — it takes sample sequences from the unit interval to probability measures — but sampling from a graphon gives an extended hypergraphon. The reverse is not quite true. Some extended hypergraphons do not arise from sampling Borel structures.

For any extended L-hypergraphon $\mathcal{W}$ and any set S, we take an independent identically distributed sequence $\left(\zeta_D : D \in [\{1, \ldots, k\}]^{<k}\right)$ of reals from the unit interval, and represent $G(S, \mathcal{W})$ in the following way. For a set $J \in [\{1, \ldots, k\}]^k$, let $Z_J = (z_F : F \subseteq_{fin} J)$ be defined by $z_F = \zeta_F$. Now for each ζ_D, the quantifier-free type of the tuple $(n_1, \ldots, n_k)$ is distributed according to

$$E_J^W := \mathcal{W}\left(Z_J, \zeta_J\right).$$

THEOREM 6.1.24 ([**7**]). *Let L be a relational signature where k is the maximum arity of any relation symbol of L. Let $\mathcal{W}$ be an extended L-hypergraphon, and denote by $G(S, \mathcal{W})$ a structure sampled from $\mathcal{W}$ with universe S. Then*

$$\lim_{n\to\infty} \frac{H\left(G(\mathbb{N}, \mathcal{W})\right)(n)}{\binom{n}{k}} = \int H(W(t))dt.$$

PROOF. We prove only the less-technical side, that

$$\lim_{n\to\infty} \frac{H\left(G(\mathbb{N}, \mathcal{W})\right)(n)}{\binom{n}{k}} \geq \int H(W(t))dt.$$

By the independence assumptions of the model construction,

$$H\left(G(n, \mathcal{W}) \middle| \left(\zeta_I : I \in [\{1, \ldots, n\}]^{<k}\right)\right)$$

is equal to the sum $\sum_{J\in[\{1,\ldots,n\}]^k} H(\mathcal{W}(Z_J))$. Taking expectations, we can find that

$$H\left(G(n, \mathcal{W})\right) \geq \sum_{J\in[\{1,\ldots,n\}]^k} E(H(\mathcal{W}(Z_J))),$$

which, by identical distribution, is equal to

$$\left|[\{1, \ldots, n\}]^k\right| \int H(W(t))dt.$$

□

One particular outcome of this theorem is that for some constant C, we have $H\left(G(\mathbb{N}, \mathcal{W})\right)(n) = Cn^k + o(n^k)$, generalizing the result for graphs.

6.2. Keisler Randomizations

6.2.1. The Idea of Randomization Structures. There have been other ways beyond graphons to capture randomness in a single structure. The idea of Keisler was to start with a structure $\mathcal{M}$, and pass to another structure whose elements are random elements sampled from $\mathcal{M}$. After Keisler's initial formulation of these structures and their theories, an additional refinement was added by Ben Yaacov and Keisler by formulating this process in continuous logic. In any case, the phenomenon of random elements sampled from another structure is by now

familiar, paralleling, as it does, the sampling of graphons and of sampleable Borel structures.

We first consider the relationship of this sections with the previous sections. Let $\mathcal{W}$ be a structure — in the preceding sections it would have been a samplable Borel structure, but these restrictions are not critical for the current purpose. The countable sequences of elements from $\mathcal{W}$ can be parameterized as functions from $[0,1)$ to $\mathcal{W}$ with countable range. We call the set of such functions $\mathcal{W}^{[0,1)}$. We can describe a random structure sampled from $\mathcal{W}$ in this sense by talking about the probabilities that various sentences in the language of $\mathcal{M}$ hold on an element of $\mathcal{W}^{[0,1)}$. We will form a two-sorted metric structure in the following way: we take K to be the set $\mathcal{M}^{[0,1)}$, and B to be the set of Borel subsets of $[0,1)$, with the standard Lebesgue measure λ. For any formula $\varphi(\bar{x})$ in the language of $\mathcal{W}$, we include in our structure a function mapping tuples of elements from K into elements of B by setting, for any tuple of functions $\bar{f} \in \mathcal{M}^{[0,1)}$,

$$[\![\varphi(\bar{f})]\!] := \left\{t \in [0,1) : \mathcal{M} \models \varphi\left(\bar{f}(t)\right)\right\}.$$

We define the metrics $d_K(f,g) = \lambda\{t : f(t) \neq g(t)\}$ and $d_B(a,b) = \lambda(a \triangle b)$. This important example of the randomization structures of the present section was introduced in [**23**], where it was called the *Borel randomization of* $\mathcal{M}$.

This approach differs in an important way from the work of Gaifman we have already mentioned. Gaifman took the assignment of probability values (measures) to sentences to be primitive. In the present approach, we take the sampling to be primitive, and calculate probabilies according to the canonical probablilty theory of Kolmogorov.

The major initial insight of this line of research is that, given an initial structure $\mathcal{M}$, the "randomizations" of $\mathcal{M}$ are all models of a single complete theory. From this, it turns out that the passage from $Th(M)$ to the "randomization theory" preserves various important properties of first-order theories.

We begin with describing the appropriate signature for the randomizaiton structure.

Definition 6.2.1 (Ben Yaacov-Keisler). Let L be a first-order signature. Then L^R is the following continuous signature with a sort K of random elements and a sort B of events.

- For each first-order formula $\varphi(\bar{x})$ in n free variables, the signature L^R has the n-ary function symbol $[\![\varphi(\cdot)]\!] : K^n \to B$.
- A unary predicate a unary (continuous) predicate μ on B,
- The Boolean operations $\sqcup, \sqcap, \neg, \top$, and $\bot$ on B
- A metric d_K on K
- A metric d_B on B.

Now, given an L-structure $\mathcal{M}$, we begin to construct a randomization $\mathcal{M}^R$.

Definition 6.2.2. Let $\mathcal{M}$ be an L-structure. Then a *pre-randomization of* $\mathcal{M}$ is a 2-sorted L^R-structure with universe (K,B) such that

(1) $(B, \sqcup, \sqcap, \neg)$ is an algebra of sets.
(2) There is some pair (Ω, μ) such that (Ω, B, μ) is an atomless finitely additive probability space.
(3) K is a set of functions from Ω to $\mathcal{M}$.

(4) For any $\bar{f} \in K^n$, the set $[\![\varphi(\bar{f})]\!]$ is the element of B consisting of all w such that $\mathcal{M} \models \varphi\left(\bar{f}(w)\right)$

This much already captures K as a set of random elements of $\mathcal{M}$. A more complete treatment of the idea of a randomization structure would provide semantic meaning for the measure μ and the metrics on the two sorts, and restrict B to events with semantic meaning in the sense of $\mathcal{M}$.

DEFINITION 6.2.3. Let $\mathcal{M}$ be an L-structure, and $\mathcal{M}^R$ a pre-randomization of $\mathcal{M}$. We say that $\mathcal{M}^R$ is a randomization of $\mathcal{M}$ if and only if the following conditions hold.

(1) For each $b \in B$ and each $\epsilon > 0$, there are $f, g \in K$ such that $\mu\left(b \triangle [\![f = g]\!]\right) < \epsilon$
(2) For each formula $\varphi(x, \bar{y})$ of L, for each $\epsilon > 0$, and for each tuple $\bar{g}$ of elements from K, there is $f \in K$ such that $\mu\left([\![\varphi(f, \bar{g})]\!] \triangle [\![\exists x\ \varphi(x, \bar{g})]\!]\right) < \epsilon$.
(3) The pseudo-metric d_K on K is defined by $d_K(f, g) = \mu[\![f \neq g]\!]$.
(4) The pseudo-metric d_B on B is defined by $d_B(a, b) = \mu(a \triangle b)$.

The degree to which randomizations of $\mathcal{M}$ are canonical will be considered in the next section. For now, it suffices that we have them. It is at least relevant for completeness that [**57**] shows how to randomize a metric structure.

There is a natural notion of pseudofiniteness for metric structures, first described by Goldbring and Lopes. Indeed, there are two closely related notions, as Goldbring and Lopes note that compact metric structures are frequently the right metric analogy to finite first-order structures.

DEFINITION 6.2.4 ([**202**]). Let $\mathcal{M}$ be a metric structure.

(1) We say that $\mathcal{M}$ is *pseudofinite* if and only if for any sentence σ, if for all finite substructures $\mathcal{A} \subseteq \mathcal{M}$, we have $\sigma^{\mathcal{A}} = 0$, then we also have $\sigma^{\mathcal{M}} = 0$.
(2) We say that $\mathcal{M}$ is *strongly pseudofinite* if and only if for any sentence σ with $\sigma^{\mathcal{M}} = 0$ there is some finite substructure $\mathcal{A}$ such that $\sigma^{\mathcal{A}} = 0$.
(3) We say that $\mathcal{M}$ is *pseudocompact* if and only if for any sentence σ, if for all compact substructures $\mathcal{A} \subseteq \mathcal{M}$, we have $\sigma^{\mathcal{A}} = 0$, then we also have $\sigma^{\mathcal{M}} = 0$.
(4) We say that $\mathcal{M}$ is *strongly pseudocompact* if and only if for any sentence σ with $\sigma^{\mathcal{M}} = 0$ there is some compact substructure $\mathcal{A}$ such that $\sigma^{\mathcal{A}} = 0$.

One natural test for these definitions is whether they agree with the classical notion of pseudofiniteness on classical structures considered as discrete metric structures. This turns out to be the case.

THEOREM 6.2.5 ([**202**]). *Let $\mathcal{M}$ be a classical first-order structure. Then the following conditions are equivalent.*

(1) *$\mathcal{M}$ is pseudofinite (in the classical first-order sense).*
(2) *$\mathcal{M}$, considered as a discrete metric structure, is pseudofinite.*
(3) *$\mathcal{M}$, considered as a discrete metric structure, is strongly pseudofinite.*
(4) *$\mathcal{M}$, considered as a discrete metric structure, is pseudocompact.*
(5) *$\mathcal{M}$, considered as a discrete metric structure, is strongly pseudocompact.*

PROOF. Some of the implications are straightforward. Pseudofiniteness as a metric structure implies pseudofiniteness in the classical sense (the converse is more

subtle, since there are metric structures that are not discrete). Since every finite structure is compact, a strongly pseudofinite structure is strongly pseudocompact.

Strong pseudofiniteness implies pseudofiniteness and strong pseudocompactness implies pseudocompactness, much as these concepts follow in the classical case: If every sentence σ with $\sigma^{\mathcal{M}} = 0$ has a finite (or compact) model within $\mathcal{M}$, then $\mathcal{M}$ is elementarily equivalent to an ultraproduct of finite (respectively, compact) structures. In that case, if every sentence evaluates to zero on every finite (compact) structure, then it must evaluate to zero in the ultraproduct, so that $\mathcal{M}$ is pseudofinite.

Suppose that $\mathcal{M}$ is pseudofinite in the classical sense. Write $\mathcal{M}$, up to elementary equivalence, as a ultraproduct $\prod_{\mathcal{U}} A_i$ of finite structures. By Proposition 2.6.9, we have $\mathcal{M} \equiv_{CFO} \prod_{\mathcal{U}} A_i$. Then we observe that $\mathcal{M}$ is strongly pseudofinite.

It now suffices to prove that if $\mathcal{M}$ is pseudocompact, then it is pseudofinite in the classical sense. Recall from Definition 2.6.7 the natural translation of classical formulas to continuous formulas. Let $\mathcal{M}$ be pseudocompact, and φ a classical L-sentence with $\mathcal{M} \models \varphi$. For each predicate P occurring in φ (for convenience, we count equality as such a predicate for the present proof), since $\mathcal{M}$ is discrete, we know that if $P^{\mathcal{M}}(\bar{a}) \leq \frac{1}{2}$, then $P^{\mathcal{M}}(\bar{a}) = 0$. Consequently, for each such P, we form the sentence

$$\psi'_P(\bar{x}) = \left(P(\bar{x}) \dot{-} \left(P(\bar{x}) \dot{-} \frac{1}{2} \right) \right).$$

We note that whenever $P^{\mathcal{M}}(\bar{x}) \leq \frac{1}{2}$ we have $P^{\mathcal{M}}(\bar{x}) \dot{-} \frac{1}{2} = 0$, so that $(\psi'_P)^{\mathcal{M}}(\bar{x}) = P^{\mathcal{M}}(\bar{x}) \leq \frac{1}{2}$. On the other hand, if $P^{\mathcal{M}}(\bar{x}) > \frac{1}{2}$, then $(\psi'_P)^{\mathcal{M}}(\bar{x}) = \frac{1}{2}$. Consequently, we define $\psi_P = \sup_x \psi'_P(\bar{x})$, and note that $\psi_P^{\mathcal{M}} = 0$. Let

$$\Psi = \max\left(\tilde{\varphi}, \psi_P : P \text{ occurs in } \varphi\right).$$

Now there is a compact structure $\mathcal{A} \subseteq \mathcal{M}$ with $\Psi^{\mathcal{A}} \leq \frac{1}{4}$. We take a quotient of $\mathcal{A}$ by the relation $E_{1/4}$ that holds of (x, y) if and only if $d(x, y) \leq \frac{1}{4}$. This relation is transitive, since $\psi_{=}^{\mathcal{A}} \leq \frac{1}{4}$. Since $\mathcal{A}$ is compact, a covering by open balls of radius $\frac{1}{4}$ must have a finite subcover, so that this quotient must be finite. We note that every predicate of L is invariant on $E_{1/4}$, so that we can interpret L in a natural way on the quotient $\mathcal{B} = \mathcal{A}/E_{1/4}$. One can then see that $\mathcal{B}$ is a finite structure satisfying φ. □

The sturcture of definability in a randomization of $\mathcal{M}$ is closely related to definability in $\mathcal{M}$. Recall that we say that an element is definable if there is a formula satisfied by exactly that element.

THEOREM 6.2.6 ([**21**]). *Let $\mathcal{M}$ be a first-order L-structure, and (K, B) a randomization of $\mathcal{M}$, and let $A \subseteq K$.*

1. *An element $b \in B$ is definable with parameters from A if and only if it is in the d_B-closure of the set of elements e such that $e = [\![\varphi(\bar{a})]\!]$ for some first order formula φ and some tuple $\bar{a} \subseteq A$.*
2. *An element $f \in K$ is definable with parameters from A if and only if there exist pairwise disjoint elements $\{E_n : n \in \mathbb{N}\}$ of B such that*
   (a) $\sum_{n=0}^{\infty} \mu(E_n) = 1$,

(b) *For every n, the element E_n is definable with parameters from A, and*
(c) *For every n, there is some classical L-formula φ_n and some tuple $\bar{a} \subseteq A$ such that $f \restriction_{E_n}$ is equal to the set $[\![\varphi_n(\bar{a})]\!]$.*

6.2.2. The Randomization Theory. Keisler constructed a transformation from a classical first-order L-theory T to a *randomization theory* T^R whose models are the randomizations of models of T. This theory is expressed in the randomization signature L^R introduced in the previous section. We axiomatize that the true formulas of T have value 0, that the Boolean connectives in type B match their first-order counterparts, and that the definitive criteria of randomizations of a model are satisfied.

DEFINITION 6.2.7. Let T be a first-order L-theory. Then the randomization of T, denoted T^R, is the theory axiomatized by the following schema:

(1) For any L-formula $\varphi(\bar{x})$ for which $\forall \bar{x} \varphi(\bar{x})$ is logically valid, the formula $\sup_{\bar{x}} d_B([\![\varphi(\bar{x})]\!], \top)$
(2) For any $\varphi \in T$, the formula $d_B([\![\varphi]\!], \top)$
(3) Standard Boolean axioms in type B
(4) $\sup_{\bar{x}} d_B([\![\neg\varphi(\bar{x})]\!], \neg[\![\varphi(\bar{x})]\!])$
(5) $\sup_{\bar{x}} d_B([\![\varphi_1 \vee \varphi_2(\bar{x})]\!], [\![\varphi_1(\bar{x})]\!] \sqcup [\![\varphi_2(\bar{x})]\!])$
(6) $\sup_{\bar{x}} d_B([\![\varphi_1 \wedge \varphi_2(\bar{x})]\!], [\![\varphi_1(\bar{x})]\!] \sqcap [\![\varphi_2(\bar{x})]\!])$
(7) $\sup_{x,y} d_K(x, y) = 1 - \mu[\![x = y]\!]$
(8) $\sup_{a,b} d_B(a, b) = \mu(a \triangle b)$
(9) $\sup_{\bar{y}} \inf_{x} d_B([\![\varphi(x, \bar{y})]\!], [\![\exists x\ \varphi(x, \bar{y})]\!])$
(10) $\sup_{b} \inf_{x,y} d_B(b, [\![x = y]\!])$
(11) $\mu(\top) = 1$
(12) $\mu(\bot) = 0$
(13) $\sup_{a,b} (\mu(a \sqcup b) - \mu(a) - \mu(b) + \mu(a \sqcap b))$
(14) $\sup_{a} \inf_{b} \mu(a \sqcap b) = \frac{\mu(a)}{2}$

This theory T^R is reasonably well-behaved even without any assumptions on T or its models.

THEOREM 6.2.8 ([**272**]). *The randomization theory T^R admits quantifier elimination.*

PROOF. We show first that all quantifiers of sort K can be eliminated, and then those of sort B. We consider the case

$$\psi = \inf f \in K \psi_1(f, \bar{x})$$

where ψ_1 is quantifier-free. Consider the set of first-order formulas $\varphi_i(f, \bar{x})$ such that the term $[\![\varphi_x(f, \bar{x})]\!]$ occurs in ψ_1. We can now express ψ_1 as a quantifier free formula ψ_2 in which f occurs only in terms of the form $[\![\theta(f, \bar{x})]\!]$, where θ is a conjunction of literals in the φ_i. Now there exists a formula ψ_3 equivalent to ψ_2 in which each such θ is replaced by a new variable of sort B, with only existential B-quantifiers.

By the standard induction, it now suffices to prove quantifier elimination for the case of a formula ψ with only sort B quantifiers.

Again we focus on a single quantifer of type B. Let

$$\psi_1 = \inf_{b \in B} \psi(b, \bar{x}),$$

where ψ_1 is quantifier-free. We let $\{\alpha_i : i \in I\}$ be a finite set of terms such that $\{b\} \cup \{\alpha_i : i \in I\}$ generates a boolean algebra containing all terms of sort B occurring in ψ_1. At this point, any term of sort B occurring in ψ_1 can be expressed in such a way that the only terms involving b are the terms of the form $b \sqcap \alpha_i$. By axiom 14, these can be replaced by terms not involving b at all. Consequently, ψ is equivalent to a quantifier-free formula. □

The completeness of the theory T^R follows from quantifier elimination.

THEOREM 6.2.9. *If T is complete, then T^R is complete.*

PROOF. Given a sentence ψ, we assume without loss of generality that ψ is quantifier-free. It can, then, have no variables. Since there are no constant symbols of sort K and only two constant symbols of sort B, the constant terms must be boolean combinations of $\bot, \top$, and terms of the form $[\![\varphi]\!]$, where φ is a sentence of L. By axioms 2, 11, and 12, either $T^R \vdash \varphi$ or $T^R \vdash \neg\varphi$. □

So we do, at least, have a theory of something. It remains to show that we actually have a theory of the randomization structures.

THEOREM 6.2.10. *If T is the complete theory of an L-structure $\mathcal{M}$, then, up to pre-isomorphism, the randomizations of $\mathcal{M}$ are exactly the models of T^R.*

PROOF. Let $\mathcal{N} = (K, B)$ be a randomization of $\mathcal{M}$. If φ is a sentence true in $\mathcal{M}$, then $d_B^{\mathcal{N}}([\![\varphi]\!], \top)$ is the measure of the symmetric difference of the set of elements satisfying φ and the universe $\top$ — that is, the measure of the emptyset. Now every element of B is approximated by elements of the form $[\![\varphi]\!]$, so by approximation all of the boolean axioms (1–7, 9, and 11–13) hold. Axiom 8 holds by definition of a randomization. Axiom 10 is a direct interpretation of the first point of Definition 6.2.3. Axiom 14 arises from the fact that (Ω, B, μ) is atomless. We conclude that $\mathcal{N} \models T^R$.

Now suppose that $\mathcal{N} = (K, B)$ is a model of T^R. Let Ω be the set of functions from K to $\mathcal{M}$ with finite range. For $k \in K$, define $X_k : \Omega \to M$ by $X_k : \omega \mapsto \omega(k)$ We let $K' := \{X_k : k \in K\}$ and $B' = \left\{[\![\varphi\left(\overline{X_k}\right)]\!] : k \in K, \varphi \in L\right\}$. The natural maps from K to K' and from B to B' are a pre-isomorphism, and (K', B') is a randomization of $\mathcal{M}$. □

In connection with the interpretation of both randomizations and pseudofinite structures as being, in different senses, "random," we return briefly to pseudofinite metric structures in light of Theorem 6.2.10.

COROLLARY 6.2.11 ([**202**]). *If $\mathcal{M}$ is a pseudofinite classical first-order structure, then any of its randomizations is a pseudofinite metric structure.*

PROOF. In metric structures, as in classical structures, pseudofiniteness is equivalent to being elementarily equivalent to an ultraproduct of finite structures. By Theorem 6.2.10, it suffices to prove that some model of $(Th(\mathcal{M}))^R$ is an ultraproduct of finite structures. View $\mathcal{M}$ as an ultraproduct $\prod_{\mathcal{U}} \mathcal{M}_n$ of finite structures.

Let K_n be the set of functions from $\{1,\dots,n\}$ to $\mathcal{M}_n$, and let $B_n = \mathcal{P}\{1,\dots,n\}$, with counting measure. We set $\mathcal{N} = \prod_{\mathcal{U}}(K_n, B_n)$, defining d_B and d_K in the obvious ways. Axioms 3–8 and 11–13 hold by construction on $\mathcal{N}$. Axioms 1, 9, and 10 are true in each of the factor structures, so are true in $\mathcal{N}$, as well. Axiom 2 also transfers from the factor structures, only with the provision that a sentence may only be true in almost all of them. It remains to verify axiom 14.

Let $a = (a_n : n \in \mathbb{N})$ be an element of $\prod_{\mathcal{U}} B_n$. If $\mu(a_n) = 0$ for a set of n which is not $\mathcal{U}$-null, then $\mu(a) = 0$ and there is nothing to prove. Suppose, then, that $\mu(a_n) > 0$ for all $n \in \mathcal{U}$. Since the measure on B_n is counting measure, that puts the cardinality of the event a_n greater than 1 for almost all n. For any n where this condition holds, we define some subset $b_n \subseteq a_n$ with $|b_n| \leq \frac{|a_n|}{2}$. Now Axiom 14 is verified. □

This result links together the results of the previous chapter with those of this chapter, and establishes that randomization structures generalize pseudofinite structures, as well as matching the sampling model of an important class of graphons and of the entire class of samplable structures.

Much of the model-theoretic work around randomizations has followed the program of transfering stability characteristics from T to T^R, and there seems to be a robust transfer. A natural place to start is with the existence of prime models.

THEOREM 6.2.12 ([**23**]). *Let T be a first-order L-theory. Then T has a prime model if and only if T^R has a prime model. Moreover, a model of T^R is prime if and only if it is isomorphic to the Borel randomization of a prime model of T.*

PROOF. Let $\mathcal{M}$ be a prime model of T. Consider the Borel randomization $(\mathcal{M}^{[0,1)}, B)$ described at the outset of Section 6.2.1. We will show that this model of T^R is prime. Let $\bar{f}$ be a tuple in $\mathcal{M}^{[0,1]}$ of complete type p. Let $ran(\bar{f})$ be enumerated by $\{\bar{a}_i : i \in I\}$, denote by p_i the complete L-type of $\bar{a}_i$, and let $\alpha_i = \mu\left\{x : \bar{f}(x) = \bar{a}_i\right\}$. Since $\mathcal{M}$ is prime, each p_i is isolated, so that it must be realized in every model of T^R, so that $(\mathcal{M}^{[0,1]}, B)$ is prime, as desired.

On the other hand, suppose T^R has a prime model $\mathcal{N} = (K, B)$. If T does not have a prime model, then there is a formula $\varphi(\bar{x})$ consistent with T but not in a principal type. Because $\varphi(\bar{x})$ is consistent with T, we have $T^R \models [\![\exists \bar{x}\varphi(\bar{x})]\!] \doteq \top$, so that there is a tuple $\bar{f}$ such that $\left(d_B([\![\varphi(\bar{f})]\!)\right)^{\mathcal{N}} = 0$ (this is a nontrivial theorem of [**59**]). We denote the type of $\bar{f}$ by $p_{\bar{f}}$. Since $\mathcal{N}$ is prime and $p_{\bar{f}}$ is realized in $\mathcal{N}$, it follows that $p_{\bar{f}}$ is principal. On the other hand, then, $p_{\bar{f}}$ must be realized in the Borel randomization of an arbitrary countable model $\mathcal{N}_0$, and can be decomposed as in the previous part into complete L-types p_i with measures α_i. If $\alpha_i > 0$, then p_i is realized in $\mathcal{N}_0$, which, we recall, is arbitrary. Thus, p_i is principal, and $\varphi(\bar{x})$ belongs to a principal type p_i. Thus, T has a prime model. □

In view of Section 7.4, the transfer of stability properties has potential importance to the general theory of logic and probability. We summarize several important results here.

THEOREM 6.2.13 ([**59, 56, 57**]). *Let T be a first-order theory and T^R its randomization.*

- *T is $\aleph_0$-categorical if and only if T^R is $\aleph_0$-categorical.*

- *T is $\aleph_0$-stable if and only if T^R is $\aleph_0$-stable.*
- *If T is stable, then T^R is stable.*
- *T is NIP if and only if T^R is NIP.*
- *If T is simple unstable, then T^R is not simple.*

6.3. Algorithmically Random Structures

6.3.1. Algorithmic Randomness and Selection Randomness. Readers of Chapter 3 may have already wondered about another method of constructing "random" structures that has not yet explicitly been discussed. We could order the unordered pairs of distinct natural numbers lexicographically (or in whatever other way we like) by $\{\{a_e, b_e\} : e \in \mathbb{N}\}$, let $\sigma \in 2^\omega$ be Martin-Löf random (or algorithmically random in whatever other way we like), and construct a graph with vertex set $\mathbb{N}$ by putting an edge between a_e and b_e if and only if $\sigma(e) = 1$.

A related technique would be to start with some fixed graph $\mathcal{G}$ with vertex set $\mathbb{N}$, let $S \subseteq \mathbb{N}$ be an algorithmically random set, and let $\mathcal{G}^r$ be the induced subgraph of $\mathcal{G}$ with vertex set S.

There is certainly a relationship of these techniques to the others described in this chapter and the previous one, and we will see that under the right conditions (at least in the obvious cases) they lead to similar ends. It is appropriate, though, before taking up these approaches to random structures, to note some sensitivities and some connections.

The reader will recall from Chapter 3 that the measure-theoretic definitions of algorithmic randomness arise naturally from the Kolmogorov model of probability described in Section 2.1.2: we begin with a measure space of total measure 1, interpret random variables as functions defined on that space, and interpret sets as events. An element of the codomain of these random variables is defined to be random in various senses, according to its membership in sets which effectively have large measure or its nonmembership in sets which effectively have small measure. All of this makes sense, in principle, whether the codomain of the random variables is the set of reals, the set of graphs, or the set of semisimple Lie algebras[1].

There are, however, significant sensitivities in this study. What is the appropriate measure to use. In the case of graphs, the construction given in the introduction to this paragraph seems natural enough, but it is not at all obvious that the graph constructed (even up to isomorphism) should be independent of the ordering of pairs, and even then the Lebesgue measure on reals still seems foreign to the structure of the random graphs in a sense that none of the previous models have.

We constructed random graphs as models of an almost sure theory, and while the precise structure may have required some suspension of disbelief, the selection of the first order theory was at least clearly from the world of graphs (and we were fortunate to find that theory $\aleph_0$-categorical when we got there. Perhaps the main result of Section 6.1 is that several different approaches — many of them transparently arising from the intrinsic nature of graphs — led to interchangeable

[1]To the best of my knowledge, there has been limited work to date on the effective structure theory of Lie algebras of any kind, in spite of the rich algebraic literature on their structure and their relationship with several algebraic groups which have been somewhat more studied in this field, e.g. in [**231, 19**].

definitions. Invariant measures, sampleable Borel structures, and Keisler randomizations have an even more direct connection to the graphs they represent. The choice of a measure in which to construct *algorithmically random* structures requires care and then careful defense.

Several approaches have been attempted, which deal with these issues differently, but the sensitivity around whether the randomness is really of an "intrinsic" nature does seem to be a real issue. It is, of course, a standard (but often tacit) concern in mathematics, to, as Hilbert put it, "use in the proof of a theorem as far as possible only those auxiliary means that are required by the content of the theorem," [**228**], with the standard problematic example presented by the proof of the Prime Number Theorem using methods of complex analysis. Something of the philosophical literature on this point can be found in [**26, 138**]. Philosophical and cultural concerns notwithstanding, it is difficult to accept a mathematical definition as definitive unless it either shows tight tethering to the object being modeled or has some proven sense of canonicity, and both seem hard to come by here.

6.3.2. Defining Algorithmically Random Structures. An early approach to algorithmically random structures was given by Khoussainov.

DEFINITION 6.3.1 ([**276**]). Let $L = (P_i : i \leq k)$ be a finite relational language, where P_i has arity n_i, and let $c_n : \omega \to \omega^n$ be a bijection for each $n \in \mathbb{N}$. Also, denote, for each $i \in \mathbb{N}$, by $q(i), r(i)$ the unique natural numbers such that $i = kq(i) + r(i)$, with $r(i) \in \{0, \ldots, k-1\}$. We say that a structure $\mathcal{A}$ with univers $\mathbb{N}$ is *Martin-Löf string-random* if and only if the function $\alpha_{\mathcal{A}} \in 2^\omega$ given by

$$\alpha_{\mathcal{A}} : i \mapsto P_{r(i)}(c_{n_{r(i)}}(q(i)))$$

is Martin-Löf random.

For a graph, we have exactly one binary relation, and (subject to the ordering given by c_n) corresponds exactly with the construction given at the outset of the previous section. Khoussainov and Nies independently proved the following result.

PROPOSITION 6.3.2. *Let $\mathcal{G}$ be a Martin-Löf string-random graph. Then $\mathcal{G}$ is isomorphic to the random graph; that is, to the unique countable model of the almost sure theory of finite graphs.*

PROOF. Suppose that $\mathcal{G}$ is Martin-Löf string random. Let $x_1, \ldots, x_n, y_1, \ldots, y_m$ be a sequence of distinct elements. It suffices to show that there is some z adjacent to every x_i and not adjacent to any y_i. We define $f : \omega \to \omega^{n+m}$ in the following way:

$$f : z \mapsto \left(c_2^{-1}(x_1, z), \ldots, c_2^{-1}(x_n, z), c_2^{-1}(y_1, z), \ldots, c_2^{-1}(y_m, z)\right).$$

Now for any z, the string $f(z)$ is a substring of $\alpha_{\mathcal{A}}$, and if $1^n0^m \notin ran(f)$, then we have a betting strategy by which, in any of the infinitely many times we come close to seeing 1^n0^m we can bet against its completion (see Proposition 3.2.20). So we have $1^n0^m \in ran(f)$, so that there is some z with the desired properties. □

Khoussainov, very reasonably, found this result counterintuitive, since a Martin-Löf random string can never be computable, but a Martin-Löf random graph must be (at least, up to isomorphism).

To extend this definition to structures with functions, Khoussainov defined a notion of *branching classes*, on which a similar definition could hold.

DEFINITION 6.3.3 ([**277, 223**]). Let K be a decidable class of structures, closed under isomorphism.

(1) A height function on K is a function $h : K/\cong \to \mathbb{N}$ with the following properties:
   (a) If $\emptyset \in K$, then $h(\emptyset) = 0$.
   (b) For each $i \in \mathbb{N}$, the set of isomorphism types of structures $h^{-1}(i)$ is finite,
   (c) For each $\mathcal{A} \in K$ with $h(\mathcal{A}) = i > 0$, there is a substructure $\mathcal{A}[i-1] \in K$ with $h(\mathcal{A}[i-1]) = i-1$, with the property that for any substructure $\mathcal{B} \subseteq \mathcal{A}$ with $h(\mathcal{B}) < i$, we have $\mathcal{B} \subseteq \mathcal{A}[i-1]$, and
   (d) For each $\mathcal{A}, \mathcal{B} \in K$ with $h(\mathcal{A}) = i$ and $\mathcal{A}[i-1] \subset \mathcal{B} \subseteq \mathcal{A}$ such that $\mathcal{A}[i-1] = \mathcal{B}[i-1]$, we have $h(\mathcal{B}) = i$.
(2) We say that a height function h on K is *effective* if and only if it is computable as a function $h : K \to \mathbb{N}$ and there is a computable function to determine, given i, the cardinality of $h^{-1}(i)$.
(3) We say that K, with height function h, is a *branching class* if for any $\mathcal{A} \in K$ with $h(\mathcal{A}) = i$, there exist distinct structures $\mathcal{B}_0, \mathcal{B}_1$ such that $h(\mathcal{B}_0) = h(\mathcal{B}_1) > h(\mathcal{A})$, and $\mathcal{B}_0[i] = \mathcal{B}_1[i] = \mathcal{A}$.
(4) We say that K is an *effective branching class* if and only if it is a branching class with an effective height function.

Examples of effective branching classes do not seem as abundant as one might wish. Nevertheless, some are available.

EXAMPLE 6.3.4. Let L be a finite language consisting only of function and constant symbols. A c-generated L-algebra is an L-structure in which every element is the interpretation of some ground term, and a c-generated partial L-algebra $\mathcal{A}$ is a set A along with interpretations of constant symbols of L as distinguished elements of A and interpretations of function symbols of L as partial functions. We define the height of a ground term by induction, with $h(c) = 0$ for every constant symbol c, and $h(t) = \max\{h(t_1), \ldots, h(t_n)\} + 1$ where f is an n-ary function and $t = f(t_1, \ldots, t_n)$. We define the height of an element of a c-generated algebra as the minimum height of a ground term representing that element. We now consider the class of proper partial algebras, that is, partial algebras that arise by taking an infinite c-generated L-algebra and restricting the interpretation of functions to be defined exactly on elements of height less than n for some n. Certainly, the axioms of an effective height function are satisfied here. Moreover, the partial functions can be extended in many ways, so that the class of proper partial algebras forms a branching class.

One would like, of course, in light of the previous chapter, to have graphs and fields as up-front examples. This is difficult, though. In the case of fields, only the prime fields are c-generated, so those would require a very different approach. Some restricted classes of graphs, generally with additional structure, can be shown to be branching classes, but the full class of finite graphs does not seem obvious; choices for the height function satisfying axioms 3 and 4 seem elusive.

The goal here, of course, is to give many options for extension, so that an infinite path representing all the necessary choices will construct an infinite structure. The reader will do well to compare the following construction to Definition 6.1.4.

DEFINITION 6.3.5. Let K be a branching class. We simultaneously define a tree $\mathcal{T}(K)$ and a map $\gamma : \mathcal{T}(K) \to K \cup \{\emptyset\}$, (by convention, we will assign height -1 to $\emptyset$), in the following way.

(1) $\mathcal{T}(K)$ has a root r, with $\gamma(r) := \emptyset$.
(2) Let $n \geq 1$, and let $\sigma \in \mathcal{T}(K)$ be a node at level $n-1$. The successors of σ at level n are given by $\{\sigma 1, \ldots, \sigma k\}$, and $\gamma(\sigma i)$ is, for each i, a distinct structure of height n such that $\gamma(\sigma i)[n-1] = \gamma(\sigma)$, in such a way that the successors of σ exhaust such structures.

LEMMA 6.3.6. *The function γ induces a bijection γ^* from paths of $\mathcal{T}(K)$ to countable structures which arise as direct limits of systems of the form $((\mathcal{A}_i, f_i) : i \in \mathbb{N})$ where for each i, the structure $\mathcal{A}_i$ is an element of K and $f_i : \mathcal{A}_i \to \mathcal{A}_{i+1}$ is an embedding.*

PROOF. Given a path $\sigma \in [\mathcal{T}(K)]$, we define $\gamma^*(\sigma)$ to be the direct limit of the system $((\gamma(\sigma \upharpoonright_i, f_i) : i \in \mathbb{N})$, where f_i is the natural inclusion. The function γ^* is a bijection. □

The goal is now to identify the random branches in $\mathcal{T}(K)$, and identify the limits of those systems as the random structures.

DEFINITION 6.3.7. Let K be a branching class with height function h.

(1) For each $\mathcal{B} \in K$ with $h(\mathcal{B}) = n$, we define
$$Cone(\mathcal{B}) = \{\mathcal{A} \in [\mathcal{T}(K)] : \mathcal{A}[n] = \mathcal{B}\}$$
and form a topology with sets of the form $Cone(\mathcal{B})$ as a basis.
(2) We define a measure μ on $[\mathcal{T}(K)]$ in the following way:
   (a) $\mu(Cone(\emptyset)) = 1$
   (b) Let $\mathcal{A}$ be a label on $\sigma \in \mathcal{T}(K)$ with e immediate successors labeled with $\mathcal{A}_1, \ldots, \mathcal{A}_e$. Then for any i, we assign
$$\mu(Cone(\mathcal{A}_i)) = \frac{\mu(Cone(\mathcal{A}))}{e}.$$

Now, in light of the interpretation of $[\mathcal{T}(K)]$ as the class of infinite structures of interest and in light of Definition 3.2.12, the notion of randomness is natural: A structure is said to be *Martin-Löf Random* if and only if it is a μ-Martin-Löf random point of $[\mathcal{T}(K)]$.

Certainly the measure μ concentrates on the set of Martin-Löf random paths. Moreover, μ is invariant under the logic action of S_∞. Consequently, the results of Section 6.1.2 govern the circumstances under which, for instance, any two Martin-Löf random structures must be isomorphic.

PROPOSITION 6.3.8 ([**277**]). *Let K be a branching class. Then there exists a Martin-Löf random structure $\mathcal{A} \in [\mathcal{T}(K)]$ such that $\mathcal{A} \leq_T \emptyset'$.*

PROOF. Let $\{U_i : i \in \mathbb{N}\}$ be the universal Martin-Löf test for $[\mathcal{T}(K)]$. We will construct, by stages, a path σ outside the intersection of this test. Suppose that we have constructed $\sigma_s \upharpoonright_s$, and that, for all i, we have $\gamma(\sigma_s(i-1)) = \mathcal{B}_i$. If $Cone(\mathcal{B}_{s-1}) \subseteq U_s$, then we find (using a $\emptyset'$ oracle) the greatest $t < s$ such that $\sigma_s \upharpoonright_t$ has an extension, $\sigma_s^* \notin \bigcup_{i \leq s} U_i$, with $|\sigma_s^*| = s+1$. We define $\sigma_{s+1} \upharpoonright_{s+1} = \sigma_s^*$. We define $\sigma = \lim_s \sigma_s$, and we see that $\gamma^*(\sigma)$ is Martin-Löf random and $\emptyset'$-computable. □

Since the class of finite graphs does not appear to be an effective branching class in any routine way, it would be easy to suspect that the difficulty of having random structures isomorphic to computable structures might have been avoided. This is not the case.

PROPOSITION 6.3.9 ([**277, 223**]). *There is a branching class K such that some Martin-Löf random structure in $[\mathcal{T}(K)]$ has a computable copy.*

PROOF. Consider a structure whose elements are infinite binary strings, in a language with a constant symbol e interpreted as the empty string and two binary predicates $\nVdash, \nvdash$, where $\nVdash(\sigma, \tau)$ if and only if $\tau = \sigma 1$ and $\nvdash(\sigma, \tau)$ if and only if $\tau = \sigma 0$. The class K will consist of substructures of this structure.

In particular, letting $\leq$ denote the lexicographical ordering, we let K consist of the substructures of the form

$$\mathcal{P}_\sigma = \left(\{\tau : (\tau \leq \sigma) \wedge (|\tau| \leq |\sigma|)\}, e, \nVdash, \nvdash\right).$$

We define an effective height function $h(\mathcal{P}_\sigma) = |\sigma|$.

Now let $\{U_i : i \in \mathbb{N}\}$ be the universal Martin-Löf test for $[\mathcal{T}(K)]$, and consider the leftmost element σ of $[\mathcal{T}(K)]$ U_0. This path is both random and computably approximable from the left. Let $t_0 \leq t_1 \leq \cdots$ be such a computable approximation of strictly increasing length. The direct limit of $\mathcal{P}_{\bar{t}}$ is isomorphic to $\mathcal{P}_\sigma$, but is computable. □

There is a natural version of Post's problem here. We are guaranteed that in any branching class there will be a Δ^0_2 random structure, and computable structures sometimes occur. Given a branching class K, for what degrees $\mathbf{d}$ is there a 1-random element of $[\mathcal{T}(K)]$ of degree $\mathbf{d}$. Some partial answers are known.

In this connection, we make reference to the Turing degree of the isomorphism type of a structure. Given a countable structure $\mathcal{M}$, consider the set of Turing degrees of structures isomorphic to $\mathcal{M}$ with universe $\mathbb{N}$. It is well-known that this set may or may not have a least element under Turing reducibility. If it does, then that degree is called the *degree of the isomorphism type of* $\mathcal{M}$ [**378, 90**].

THEOREM 6.3.10 ([**223**]). *There is a branching class K such that $[\mathcal{T}(K)]$ has Martin-Löf random structures whose isomorphism types have degree $\emptyset$ and degree $\emptyset'$, but has no Martin-Löf random structures whose isomorhpism types have any other degree.*

PROOF. Again, we let $\leq$ denote the lexicographical ordering on nodes in a tree. Let K be an effectively branching class with the following properties:

(1) For every $\eta \in [\mathcal{T}(K)]$ and every sequence $\sigma_0 \leq \sigma_1 \leq \sigma_2 \leq \cdots$ with $\lim_{i\to\infty} \sigma_i = \eta$, we have $\gamma^*(\eta) \leq_T (\sigma_i : i \in \mathbb{N})$.
(2) For every path $\eta \in [\mathcal{T}(K)]$ and every $\mathcal{M} \cong \gamma^*(\eta)$, there is a sequence $\sigma_0 \leq \sigma_1 \leq \sigma_2 \leq \cdots$ with $\lim_{i\to\infty} \sigma_i = \eta$, such that $(\sigma_i : i \in \mathbb{N}) \leq_T \gamma^*(\eta)$.

We note that the class constructed in the previous proof has these properties. Such a class will always have the properties required by the conclusion of the theorem.

We first consider that, for a universal Martin-Löf test $(U_i : i \in \mathbb{N})$, the set of paths in $\mathcal{T}(K) - U_0$ has a leftmost element η that can be approximated by a computable sequence $\sigma_0 \leq \sigma_1 \leq \sigma_2 \leq \cdots$, so that $\gamma^*(\eta)$ has a computable copy. Similarly, if we let υ be the right-most path in $\mathcal{T}(K) - U_0$, then every copy of $\gamma^*(\upsilon)$ computes υ. It can be shown that υ has degree $\emptyset'$.

As usual, the proof that no other degrees are possible is more technical, but the essential element is that if $\gamma^*(\eta)$ has a degree, then η must be the limit of either a computable increasing sequence or a computable decreasing sequence, as in the prior two cases. □

Of the (false) intuition that a random structure cannot be computable, something does remain, though.

THEOREM 6.3.11 ([**223**]). *Let $\mathcal{A}$ be a computable structure. Then $\mathcal{A}$ is not 2-random in any branching class.*

6.3.3. Glasner-Weiss Measures and Their Random Structures. The sensitivity of the characterization of random structures to the choice of measure was already highlighted in Section 6.1.2, where unique ergodicity was already an issue. An approach by Fouché draws on topological dynamics to address this canonicity problem.

We consider again the logic action of S_∞ on the topological space $\mathbb{M}_L$ of all L-structures. The group S_∞, of course, enjoys a topological and measurable structure of its own; indeed, there is a left-invariant metric given by

$$d(\sigma,\tau) = \begin{cases} 2^{-n} & \text{if } n \text{ is the first place that } \sigma(n) \neq \tau(n) \\ 0 & \text{if } \forall n\,[\sigma(n) = \tau(n)] \end{cases},$$

but it is not complete. On the other hand, there is a complete metric on S_∞ given by

$$D(\sigma,\tau) = d(\sigma,\tau) + d(\sigma^{-1},\tau^{-1}).$$

Since functions of finite support (that is, those which fix all but finitely many elements of $\mathbb{N}$) are dense in this metric, we see that S_∞ is second countable.

Now the dense subgroup $S_\infty^{fin} < S_\infty$ of functions of finite support is the union of an increasing sequence of finite groups, and on each of these, the uniform probability measure is left-invariant, inducing a left-invariant finitely additive probability measure on S_∞.

DEFINITION 6.3.12. We say that a group is *amenable* if it admits a left-invariant finitely additive probability measure.

Thus, we say that S_∞ is amenable. Now in the context of Hausdorff topological groups acting on Hausdorff spaces, we have the following situation:

DEFINITION 6.3.13. Let G be a Hausdorff topological group and X a Hausdorff space.

(1) A G-flow on X, denoted $(G,\cdot,X)$ is a continuous action of G on X.
(2) A G-subflow of $(G,\cdot,X)$ is a G-flow $(G,\odot,Y)$ where $Y \subseteq X$ and $\cdot\restriction_{G\times Y} = \odot$.

The logic action of S_∞ on $\mathbb{M}_L$ is, of course, a flow, and if φ is an $L_{\omega_1\omega}$ sentence of L, then the restriction of the logic action to models of φ is a subflow. There is a natural interest in decomposing flows into their constituent parts, and the following result identifies those parts.

PROPOSITION 6.3.14. *Let $(G,\cdot,Y)$ be a subflow of $(G,\cdot,X)$. The following conditions are equivalent:*

(1) *$(G,\cdot,Y)$ has no proper subflows*
(2) *Every orbit of $(G,\cdot,Y)$ is dense.*

Subflows satisfying either of these two equivalent conditions are called *minimal subflows*. The minimal subflows are connected to one another by the following standard result.

THEOREM 6.3.15. *Let G be a Hausdorff topological group. Then there exists a unique G-flow $(G, \cdot, M)$ on a Hausdorff space such that for every G-flow $(G, \odot, X)$ and every minimal subflow $(G, \odot, Y)$ of $(G, \odot, X)$, there exists a homomorphism $\pi : (G, \cdot, M) \to (G, \odot, Y)$.*

This unique flow is called the *universal minimal flow* of G. The following result of Glasner and Weis identified the universal minimal flow of S_∞. Since this universal flow is uniquely determined by S_∞, it gives a canonicity to that flow among all S_∞ actions, including the logic action in any language.

THEOREM 6.3.16 ([**199**]). *The logic action of S_∞ on the set of linear orderings with universe $\mathbb{N}$ is the universal minimal flow of S_∞.*

PROOF. We consider the set Ω^k of maps from the set $\mathbb{N}^k$ to the set $\{0, 1\}$, noting that it represents the set of structures with universe $\mathbb{N}$ in a language with a single k-ary relation. S_∞ acts on this set via the logic action.

For $T : S_k \to \{0, 1\}$, we denote by H_T the set of all $\pi \in S_k$ such that for all $\sigma \in S_k$ we have $T(\pi^{-1}\sigma) = T(\sigma)$. For any $\omega \in \Omega^k$, we define $\hat{\omega} : \mathbb{N}^k \to \{0, 1\}^{S_k}$ by

$$\hat{\omega}(i_1, \ldots, i_k) : \sigma \mapsto \omega(i_{\sigma(1)}, \ldots, i_{\sigma(k)}).$$

Finally, we denote by $\Omega^k_{[H]}$ the set of $\omega \in \Omega^k$ such that for all $\bar{i} \in \mathbb{N}^k$ the set $H_{\hat{\omega}(\bar{i})}$ is conjugate to H.

Let Σ be a minimal subset of Ω_k. For each $\omega \in \Omega^k$, Ramsey's theorem gives an infinite set $J \subseteq \mathbb{N}$ which is monochromatic in that $\hat{\omega} \restriction_{J^k}$ has a constant value $T : S_k \to \{0, 1\}$. After possibly applying some elements of S_∞, we transform ω to an ω_0 where $\hat{\omega}_0$ is constant.

We will also see that there is a subgroup $H \subseteq S_k$ such that Σ is contained in $\Omega^k_{[H]}$. Indeed, if T is the constant value of $\hat{\omega}_0$, then we let $H = H_T$.

We are now prepared to define, for the logic action on $\mathcal{L}$, the space of linear orderings on $\mathbb{N}$, a map $\pi : \mathcal{L} \to \Sigma$. For a linear ordering $L = (\mathbb{N}, <_L)$, and for each $\bar{u} \in \mathbb{N}^k$, we select $\sigma_{L,\bar{u}} \in S_k$ so that

$$u_{\sigma^{-1}_{L,\bar{u}}(1)} <_L u_{\sigma^{-1}_{L,\bar{u}}(2)} <_L \cdots <_L u_{\sigma^{-1}_{L,\bar{u}}(k)}$$

and define π by

$$\pi(L) : \bar{u} \mapsto T(\sigma_{L,\bar{u}}).$$

To show that π communites with the logic action on $\mathcal{L}$, let $\alpha \in S_\infty$. Consider a linear ordering L and its image $M = \alpha L$. In that case,

$$\pi(M)(i_1, \ldots, i_k) = T(\sigma_{M,\bar{i}}),$$

and

$$\alpha^{-1} i_{\sigma^{-1}_{M,\bar{i}}(1)} <_L \alpha^{-1} i_{\sigma^{-1}_{M,\bar{i}}(2)} <_L \cdots <_L \alpha^{-1} i_{\sigma^{-1}_{M,\bar{i}}(k)}.$$

We will then compare $\sigma_{M,\bar{i}}$ with $\sigma_{M,\alpha^{-1}\bar{i}}$, since $\alpha\pi L(\bar{i}) = T(\sigma_{M,\alpha^{-1}\bar{i}})$. Showing that these two are identical, we have $\pi\alpha = \alpha\pi$, as required. □

At this point, we have canonicity, but no measure. Recall from Section 6.1.2 our discussion of the decomposition of an invariant measure as a mixture of ergodic measures, and the observation that the set of all invariant measures must be a

compact convex set whose extremal points are exactly the ergodic measures. We say that a flow is *uniquely ergodic* if and only if there is exactly one invariant measure (which must, of necessity, be ergodic).

THEOREM 6.3.17 ([**199**]). *The logic action of S_∞ on the space $\mathcal{L}$ of linear orderings with universe $\mathbb{N}$ is uniquely ergodic.*

PROOF. We let μ be an arbitrary S_∞-invariant probability measure on $\mathcal{L}$, and consider its values on sets of the form X_ℓ, where ℓ is a finite linear ordering and X_ℓ is the set of all extensions of ℓ to elements of $\mathcal{L}$. Since, for every finite ℓ, every element of $\mathcal{L}$ extends $\sigma\ell$ for some $\sigma in S_n$ for some n, the space $\mathcal{L}$ must be covered by S_∞-translates of X_ℓ, and by invariance, each must have measure $\frac{1}{|\ell|}$. This uniquely determines the value of μ on all Borel subsets of $\mathcal{L}$. □

We follow here the usage of Fouché in calling this unique measure on the universal minimal flow the *Glasner-Weiss measure*, and denote it by μ_{GW}.

LEMMA 6.3.18 ([**174**]). *The Glasner-Weiss measure is computable.*

PROOF. We say that a set S of orders is a *simple* set if there are sets P and N, each of finite linear orderings, such that S consists of exactly the linear orderings extending those in P which do not extend those in N. It suffices to compute the measure on simple sets. This can be done by the inclusion-exclusion principle, using the calculation in the proof of Theorem 6.3.17. □

THEOREM 6.3.19 ([**174**]). *Let $(L, <)$ be a μ_{GW}-Martin-Löf random linear ordering. Then $(L, <)$ is dense and without endpoints.*

PROOF. The statement that L is dense and the statement that L is without endpoints can each be expressed as infinite conjuctions of Σ_1 conditions. It suffices to show that each of these Σ_1 conditions has measure 1. Really, the two are similar, so we show here the argument respecting density.

Let $n \neq m \in \mathbb{N}$. For density to fail on n, m — that is, for $\{n, m\}$ to be an unordered successor pair — we must have, for each $N \in \mathbb{N}$,

$$\forall j \leq N \ [j \leq \min\{n, m\} \vee j \geq \max\{n, m\}] .$$

For a fixed N, the quantifier-free part of this condition has measure $\frac{M}{N}$ where M is the number of finite linear orderings on $\{1, \ldots, N\}$ such that $\{n, m\}$ is an unordered successor pair in that finite ordering. Since there are $2(N-1)(N-2)$ such orderings, we have

$$\frac{M}{N} = \frac{2(N-1)(N-2)}{N} = \frac{2}{N}.$$

In the limit, this condition has measure zero, so that a random linear ordering cannot satisfy it. □

We denote, as usual, by $ML_{\mu_{GW}}$ the set of μ_{GW}-Martin-Löf randoms. Now for any linear ordering $\mathcal{L} = (L, <)$, it is natural to consider the set

$$S_{\mu_{GW}}(\mathcal{L}) = \{\sigma \in S_\infty : \sigma\mathcal{L} \in ML_{\mu_{GW}}\}.$$

Naturally, this set is empty if $\mathcal{L}$ is not a dense linear ordering without endpoints. In particular, an interesting approach to Khoussainov's paradox that algorithmically random structures could be isomorphic to a computable structure goes through the consideration of $S_{\mu_{GW}}(\mathcal{L})$ where $\mathcal{L}$ is computable.

We approach the problem in this way. The set of *randomizers* of a computable linear ordering $\mathcal{L}$ — that is, the set of permutations that transform a computable $\mathcal{L}$ to a μ_{GW}-random linear ordering — can be expressed as

$$\mathcal{R}(\mathcal{L}) := \bigcup_{\sigma \in S_\infty^c} S_{\mu_{GW}}(\sigma\mathcal{L}).$$

While the problem of identifying the randomizers of a computable linear ordering remains open, Fouché has described some progress on this problem.

DEFINITION 6.3.20. ([**269**]) Let K be an age of L-structures.

(1) We say that K is *Ramsey* if and only if for all $\pi, A \in K$ for which there is an embedding $\varphi_0 : \pi \to A$, and for all positive $r \in \mathbb{N}$, there is some $B \in K$ such that for every r-coloring χ of embeddings of $\varphi_1 : \pi \to B$, there exists an embedding $\alpha : A \to B$ such that χ is constant on the range of $\alpha_* : \psi \mapsto \alpha \circ \psi$, which maps embeddings $\pi \to A$ to embeddings $\pi \to B$.
(2) We say that K is an *order age* if and only if there is a binary predicate $< \in L$ which is interpreted as a linear ordering in every element of K.

It is not hard to believe that the Fraïssé limit of an order age should be a linear ordering. Let K be an order age, and let K_0 be the class of reducts of elements of K to the language that omits the predicate $<$.

DEFINITION 6.3.21. Let K be an order age in a signature $L = L_0 \cup \{<\}$. Let K_0 be the L_0 reduct of K. We say that K is *reasonable* if and only if for every embedding $\pi : A_0 \to B_0$ of elements of K_0, every linear ordering on A_0 extends, via π, to a linear ordering $\preceq$ on B_0 so that $(B_0, \preceq) \in K$.

It was observed in [**269**] that if K is a reasonable order age, then the Fraïssé limit $\mathcal{F}_0$ of K_0 is the reduct of the Fraïssé limit $\mathcal{F}$ of K, say, $\mathcal{F} = (\mathcal{F}_0, <^{\mathcal{F}})$. We now consider the action of $Aut(\mathcal{F}_0) < S_\infty$ on the class of linear orderings on $\mathcal{F}_0$. In this action, we consider the orbit of $\mathcal{F}$, and denote the topological closure of this orbit by $X_\mathcal{K}$. We call X_K the *discerning flow* of K.

THEOREM 6.3.22 ([**175**]). *Let K be a computable age (order class) in signature $L = L_0 \cup \{\leq\}$ which is Ramsey and has the ordering property. Let K_0 denote the L_0 reduct of K. Let X_K be its discerning flow, and $\mathcal{F} = (\mathcal{F}_0, <)$ its Fraïssé limit. If X_K contains a μ_{GW}-Martin-Löf random structure, then $Aut(\mathcal{F})$ is amenable.*

PROOF. A linear ordering A belongs to X_K if and only if for any $B \in K_0$ with an ordering $<_B$, the ordering $<_B$ is the restriction of A. Since K is computable, the relation $A \in X_K$ is Π^0_1, and X_K either contains a random element or has positive measure. The latter case makes $Aut(\mathcal{F})$ amenable. □

Returning to the question of randomizers for a computable linear ordering, we see that if $\mathcal{L}$ is a computable linear ordering in X_K, then if there is a randomizer σ of $\mathcal{L}$ in $Aut(\mathcal{F}_0)$, then $Aut(\mathcal{F}_0)$ must be amenable.

The other issue, of course, is to consider μ_{GW}-random structures which are not linear orderings. The key is to observe that the logic action on the set of linear orderings is the *universal* minimal flow of S_∞.

Consider the logic action of S_∞ on the set $\mathbb{M}_\varphi$ of models of $\varphi \in L_{\omega_1\omega}$. This is a minimal flow if and only if φ is complete. Since the logic action on the set $\mathbb{M}_{LO}$ of linear orderings is the universal minimal flow, we have a homomorphism

$\pi : (S_\infty, \cdot, \mathbb{M}_{LO}) \to (S_\infty, \cdot, \mathbb{M}_\varphi)$. Now π must respect orbits, so we can define a random element of $\mathbb{M}_\varphi$ to be $\pi(\mathcal{A})$ for some μ_{GW}-random linear ordering $\mathcal{A}$. Another approach is to only define random elements of $\mathbb{M}_\varphi$ where the logic action on $\mathbb{M}_\varphi$ is uniquely ergodic. This will either refer us back to Section 6.1.2 or forward to Section 8.5.

6.3.4. Haar Measure and Haar Compatible Measure. Some classes of structures come with standard measures, but those measures are not always transparently measures on the class of structures involved. For instance, in considering algebraic fields, Haar measure is perhaps the measure most natural to the situation. Unfortunately, Haar measure is a measure on the absolute Galois group, and not on the class of algebraic fields itself.

Let $K \supseteq F$ be a Galois extension of fields. Then $Gal(K/F)$ is profinite, and so locally compact. Any locally compact group admits a unique translation-invariant measure, known as Haar measure. In the case of Galois groups, this measure is characterized by the following well-known result.

LEMMA 6.3.23. *Let $K \supseteq F$ be a Galois extension of fields, and let μ_H be the Haar measure on $Gal(K/F)$. If $J \subseteq K$ is a finite Galois extension of F, then $\mu_H\left(Gal(K/J)\right) = \frac{1}{[J:F]}$.*

It would be interesting to have a well-established notion of an algorithmically random field. In view of Theorems **??** and 5.3.6, we expect that a random field should be pseudofinite. At the very least, they should be pseudo algebraically closed; the intuition of the introduction to [**180**] is that, if we select at random a finite sequence $\bar{\sigma}$ of automorphisms of $\bar{\mathbb{Q}}$, then the fixed field of $\bar{\sigma}$ will be pseudo algebraically closed with Haar probability 1. At least two recent efforts have attempted to use Haar measure in different ways to define algorithmically random fields.

In addition to the usual addition, multiplication, and identity elements, we frequently consider fields in the language $L^* = (+, \cdot, 0, 1, (R_n : n \in \mathbb{N}))$, where for each n, the "root predicate" R_n is n-ary, and $R_n(a_0, \ldots, a_{n-1})$ holds in the field $\mathcal{F}$ exactly when the polynomial $x^n + \sum_{i=0} n - 1 a_i x^i$ has a solution in $\mathcal{F}$. In keeping with Khoussainov's definitions from section **??** — and, indeed, with standard usage in effective structure theory — we identify a field with its atomic diagram, so that it can be regarded as an element of 2^ω. We note that in this treatment, exactly matching the usual logic topology, a basic open set of algebraic fields is given by a finite string $\sigma \in 2^n$, giving a finite initial segment of the atomic diagram.

PROPOSITION 6.3.24 ([**341**]). *The class of algebraic fields of characteristic 0 up to isomorphism is computably homeomorphic to 2^ω under its usual topology.*

PROOF. Enumerate the monic elements of $\bar{\mathbb{Q}}[X]$ by $(f_i : i \in \mathbb{N})$. Fir each $\sigma \in 2^{<\omega}$, we define a polynomial g_σ and a subextension $\mathbb{Q} \subseteq F_\sigma \subseteq \bar{\mathbb{Q}}$ by recursion. We set $F_\emptyset = \mathbb{Q}$.

For each $\sigma \in 2^{<\omega}$, we find the least n with the following properties:

(1) $f_n \in F_\sigma[X]$,
(2) f_n has prime degree,
(3) f_n is irreducible in $F_\sigma[X]$, and
(4) For the least root x of f_n in $\bar{\mathbb{Q}}$, and for every $\tau \in 2^{<\omega}$ with $\tau 0 \subseteq \sigma$, there is no root of f_τ in $F_\sigma(x)$.

Since F_σ is finitely generated (by induction), there must be some n such that these conditions are satisfied. We set $g_\sigma = f_n$. Now we set $F_{\sigma 0} = F_\sigma$ and $F_{\sigma 1} = F_\sigma(\alpha)$ where α is the least root of g_σ in $\overline{\mathbb{Q}}$.

This assignment $\sigma \to F_\sigma$ defines a homeomorphism on a basis for 2^ω, and so uniquely extends to 2^ω. The resulting homeomorphism can be seen to be computable. □

This homeomorphism would, of course, justify using the uniform Lebesgue measure on 2^ω, exactly as Khoussainov and his collaborators did. However, this choice of measure is not obvious, and even gives some counterintuitive results. For instance, 29 is prime, so the class of fields containing a 29th root of 2 has measure $\frac{1}{2}$. On the other hand, the fields containing a 28th or 30th root of 2 have much smaller measure. To address this incongruity, Miller suggested the following measure.

Definition 6.3.25. We define the Haar-compatible measure H on the set of algebraically closed fields of characteristic up to isomorphism as follows. Let $\sigma \in 2^{<\omega}$. Let U_σ denote the basic open set of fields whose atomic diagram has initial segment σ. Then we set $H(U_\emptyset) = 1$, and

$$H(\sigma i) = \begin{cases} \frac{1}{d} H(U_\sigma) & \text{if } i = 1 \\ \frac{d-1}{d} H(U_\sigma) & \text{if } i = 0 \end{cases},$$

where d is the degree of the polynomial f_σ, as described in the proof of Proposition 6.3.24.

This measure can, of course, be extended to arbitrary sets in the usual way. It is evident that this measure has some significant relationship to Haar measure, although its *prima facie* definition still depends on the choice of f_σ. The following result both solidifies the connection to Haar measure and sharply limits the sensitivity of the measure to the choice of f_σ.

Lemma 6.3.26 ([**341**]). *Let $K \subseteq \mathbb{Q}$ be a finite Galois extension, and let $\mathcal{V}_K$ be the set of algebraic fields up to isomorphism containing K. Let G_K be the stabilizer of K in $G = Gal(\overline{\mathbb{Q}}/\mathbb{Q})$. Then $H(\mathcal{V}_K) = \mu_H(G_K)$.*

Proof. If $K = \mathbb{Q}$, the statement is obvious. Let $\sigma \in 2^\omega$, and let $C_{K,\sigma} = \{n | F_{\sigma\restriction_n} \cap K \subsetneq F_{\sigma\restriction_n 1} \cap K\}$. Note that these are exactly the branchings relevant for whether $h \in \mathcal{V}_K$. For each $n \in C_{K,\sigma}$, we have already designated a polynomial $f_{\sigma\restriction_n}$ of degree d_n, and the case where $F_\sigma \in \mathcal{V}_K$ has probability given by $H(\mathcal{U}_{\sigma\restriction_n 1})$, which is equal to $\frac{1}{d_n} H(\mathcal{U}_{\sigma\restriction_n})$. After exhausting all the prime factors of $[K : \mathbb{Q}]$, counting with multiplicity, we will have $F_\sigma \in \mathcal{V}_K$ with the probability we have calculated by induction. This measure will match the Haar measure of G_K, as required. □

While this measure is also not without its difficulties, Russell Miller and Johanna Franklin have recently claimed the following result, for which a proof has not yet been published.

Theorem 6.3.27. *For every H-Schnorr-random real $h \in 2^\omega$, the corresponding field F_f is relatively computably categorical in the language L^*. Moreover, there is a Kurtz-random $h \in 2^\omega$ such that F_h is not relatively computably categorical in L^*.*

Another approach is possible for developing a notion of Haar-random algebraic fields. The present author, working with Harizanov and Shlapentokh, has proposed the following approach.

DEFINITION 6.3.28. Let $S \subseteq Gal(\overline{\mathbb{Q}}/\mathbb{Q})$. We say that S is computable if and only if for each n there is a computable function $I_n : \overline{\mathbb{Q}}^n \to \mathbb{N} \times \overline{\mathbb{Q}}^n$ such that $I(\vec{\alpha}) = (n, T_S)$ if and only if T_S is a set of size n, and is exactly the set of images of $\vec{\alpha}$ under elements of S.

We note that Haar measure is a computable measure on $Gal(\overline{\mathbb{Q}}/\mathbb{Q})$, so it is clear what a μ_H-Martin-Löf random element of $Gal(\overline{\mathbb{Q}}/\mathbb{Q})$ should be.

DEFINITION 6.3.29. Let F be an algebraic field. We say that F is Haar random if and only if its absolute Galois group is of measure zero and contains a μ_H-Martin-Löf random element.

This seems to us closer to what Fried and Jarden had in mind in their description referenced above of the "typical" behavior of fields. While the full characterization of the Haar random fields is a work in progress, we do have the following initial results.

PROPOSITION 6.3.30 ([**91**]). *No computable field is Haar random. Moreover, there exists a non-computable algebraic extension of* $\mathbb{Q}$ *which is not random.*

PROOF. No algebraic number field is Haar random, as its absolute Galois group has positive measure. Let K be a computable infinite algebraic extension of $\mathbb{Q}$. Then the sequence of subsets of $Gal(\bar{\mathbb{Q}}/\mathbb{Q})$ that fix increasing subextensions of K constitutes a μ_H-Martin-Löf test, which can contain no random members.

On the other hand, if K is a computable field of infinite degree and M a non-computable field of infinite degree, then KM is not computable but every element of $Gal(\bar{\mathbb{Q}}/KM)$ is an element of $Gal(\bar{\mathbb{Q}}/K)$ which is computable and of measure 0. □

6.4. Invariant Random Subgroups

6.4.1. Defining Invariant Random Subgroups. In recent years, a notion of random structure has arisen among group theorists that bears striking similarities to some of the other notions introduced in the present chapter. Of course, Gromov described random groups some time ago in [**209**], but his approach does not seem to have led yet to extensive work in logic. The closest the logic literature seems to come to Gromov's random groups is its use as inspiration for generic computability, explored in Section 4.5.

Another approach, though, has thus far seemed more fruitful. Consider a locally compact group G, and consider the set $Sub(G)$ of closed subgroups of G. This set has a natural topology, first described in [**99**], with basis elements

$$\{O(S) : S \subseteq G \text{ open}\} \cup \{K(S) : S \subseteq G \text{ compact}\},$$

where

$$\begin{aligned} O(S) &= \{H : H \cap S \neq \emptyset\} \\ K(S) &= \{H : H \cap S = \emptyset\}. \end{aligned}$$

It is a well-known result that if G is locally compact, then $Sub(G)$ must be compact in this topology.

As an example of the kind of quesiton that can be asked in this space, consider the subgroups known as *lattices*. A lattice is a discrete subgroup $\Gamma \leq G$ such that the quotient of G by Γ has finite Haar measure — a generalization of the role that

$\mathbb{Z}^2$ plays in $\mathbb{R}^2$. The study of lattices in Lie groups is subtle [**321**], but we might, for instance, ask for which subgroups $H \leq SL_3(\mathbb{R})$ there is a sequence $(L_i : i \in \mathbb{N})$ such that $H = \lim_{i\to\infty} L_i$, and this topology on $Sub\,(SL_3(\mathbb{R}))$ gives a context in which such a question can be asked.

Consider the action of G on $Sub(G)$ by conjugation:

$$g : H \mapsto gHg^{-1}.$$

We are now prepared to state the definition of invariant random subgroups.

Definition 6.4.1. An invariant random subgroup is a Borel probability measure on $Sub(G)$ which is invariant under the action of G by conjugation.

As a first example, of course, an atomic measure concentrating on a single normal subgroup is an invariant random subgroup. We could, on the other hand, construct a probability measure that concentrates on the conjugacy class of a lattice: if $\Gamma \leq G$ is a lattice, then the action of G on $Sub(G)$ factors through G/Γ. In this sense, invariant random subgroups generalize both normal subgroups and lattices.

This feature of generalizing these important concepts was part of the early motivation for the study of invariant random subgroups. A classical theorem of Kesten states that a normal subgroup N of a finitely generated group G is amenable if and only if the spectral radius of the Cayley graph of G/N was equal to that of G. This result appeared to generalize beyond normal subgroups (with appropriate replacement of the Cayley graph), but was certainly not true for arbitrary subgroups. The result, in approrpiate form, does hold for invariant random subgroups [**2**].

Invariant random subgroups have also been used to strengthen the Kazhdan-Margulis Theorem. Let G be a connected linear semisimple Lie group with no compact factor, and $\Gamma \leq G$ discrete. A long-open question had asked whether there was a positive lower bound on the G-Haar measure of G/Γ, and another had asked whether, if G/Γ had finite measure but was not compact, whether Γ must contain a non-identity unipotent element. Kazhdan and Margulis [**266**] resolved both questions with the following result.

Theorem 6.4.2 ([**266**]). *If G is a semisimple Lie group with no compact factors, then there is an open neighborhood $U \subseteq G$ containing 1_G such that for every lattice $\Gamma \leq G$ there is some $g \in G$ with $g\Gamma g^{-1} \cap U = \{1_G\}$.*

This can be strengthened to the following result.

Theorem 6.4.3 ([**196**]). *Let G be a connected center-free semisimple Lie group with no compact factors. Then for every $\epsilon > 0$ there exists a neighborhood $U_\epsilon \subseteq G$ containing 1_G such that for any invariant random subgroup μ concentrating on discrete subgroups has the property that*

$$\mu\left\{\Gamma \leq G : \Gamma \cap U_\epsilon \neq \{1_G\}\right\} < \epsilon.$$

Gelander claims that this result, applied with $\epsilon < 1$ and the invariant random subgroups concentrating on lattices, clearly leads to the Kazhdan-Margulis result. Even without working out that reasoning, though, it is clear that Gelander's result gives considerably more contextual information, in the same way that Theorem 4.4.5 gives context to the Boone-Novikov result on the unsolvability of the word problem, except in the opposite direction: the Boone-Novikov Theorem is highly sensitive to the deterministic statement of the theorem; the Kazhdan-Margulis Theorem is not.

Classifying invariant random subgroups is, of course, an interesting problem, and some results are known. Borel proved, and Furstenberg later gave a new proof of, the following result.

THEOREM 6.4.4 ([**75, 186**]). *If G is a semisimple Lie group and H is a closed subgroup such that G/H has finite measure, then for any finite-dimensional representation of G, each H-invariant subspace is G-invariant. In particular, if G is simple and H is not dense in G, then the closure of H is discrete.*

This gives reason to hope that, at least in the right context, invariant random subgroups might concentrate on discrete groups. This turns out to be the case.

THEOREM 6.4.5 ([**1**]). *Let G be a simple Lie group with trivial center, and let μ be a non-atomic invariant random subgroup on G. Then μ is supported on discrete Zariski dense subgroups of G.*

PROOF. Let G and μ be as in the hypotheses of the theorem. We define two measurable functions from $Sub(G)$ to the Grassmanian of the Lie algebra $\mathfrak{g}$ of G: one mapping a closed subgroup H to the Lie algebra of the identity component of H, and the other map, and the other mapping H to the Lie algebra of the identity component of the Zariski closure of H. Via these two functions, μ induces two measures (μ_1 and μ_2, respectively) on the Grassmanian. These measures must be suported on $\{\{0\}, \mathfrak{g}\}$.

The measure μ_1 must assign measure 0 to $\mathfrak{g}$, since μ was nonatomic. Consequently, μ_1 must concentrate on the Lie algebra $\{0\}$. That is, for μ-almost all subgroups H, the identity component is trivial — in other words, H is discrete. We can also argue that μ_2 must concentrate on $\mathfrak{g}$, so that μ is concentrated on Zariski dense subgroups. □

Margulis proved a series of "Arithmeticity Theorems," closely related to the superrigidity results of Section 8.4.1, classifying the discrete subgroups of semisimple Lie groups (under very broad hypotheses) as the integer points of algebraic groups [**321**]. The exact statement, to say nothing of the proof, of Margulis's results is far beyond the scope of the present work, but it matters that this territory is understood.

Another characterization of invariant random subgroups tells less, but applies more broadly.

THEOREM 6.4.6 ([**1**]). *Let G be a locally compact second countable group, and μ an invariant random subgroup on G. Then there is a measure-preserving action (G, X, m) on a probability space such that μ is induced by the stabilizer map $x \mapsto G_x$.*

We frequently consider, for various purposes, the space $IRS(G)$ of invariant random subgroups of a fixed group G, in the weak-$*$ topology. The following result follows from Alaoglu's therem.

PROPOSITION 6.4.7. *If G is locally compact, then $IRS(G)$ is compact.*

There is, in any case, a considerable literature applying the space of invariant random subgroups to study the special points of this space representing lattices, as well as the manifolds arising from those lattices. A more extensive survey can be found in [**195**].

6.4.2. Classifying Invariant Random Subgroups. As with any space of measures, the ergodic measures play a leading role, and have been instrumental in classifying the invariant random subgroups in some important cases. A countably infinite group G is said to be *strongly simple* if its only ergodic invariant random subgroups are those concentrated on G and $\{1_G\}$. There are, as we will see, several known examples of this kind. However, the following question remains open.

PROBLEM 6.4.8 ([**195**]). Does every non-discrete locally compact group admit a non-trivial invariant random subgroup?

Some partial answers are known. The exposition of the next two results owes much to some unpublished lecture slides of Simon Thomas.

THEOREM 6.4.9 ([**426, 2**]). *Let G be a simple real Lie group of rank at least 2.*

1. *Let μ be an invariant random subgroup of G. Then either μ is uniform or there is a lattice $\Gamma \leq G$ such that μ concentrates on conjugates of Γ and its restriction to them is uniform.*
2. *If $\Gamma \leq G$ is a lattice and ν is an ergodic invariant random subgroup of Γ, then ν is either trivial or there is some $H \leq \Gamma$ of finite index such that ν concentrates on conjugates of H and its restriction to them is uniform.*

PROOF. Suppose that μ is a nontrivial ergodic invariant random subgroup of G. We have seen that μ is induced by the stabilizer map of an ergodic action of G. Now this action must either be *essentially transitive* (i.e. there is an orbit of measure 1) or *properly ergodic* (ergodic but not essentially transitive). Stuck and Zimmer showed that in a properly ergodic action, almost every point would have trivial stabilizer. That situation clearly does not obtain here, so there is a single orbit Gz of full measure, and ν concentrates uniformly on the conjugacy classes of the stabilizer G_z. The proof of the second part is similar. □

Vershik's classification of the invariant random subgroups of the group $S_{\infty,0}$ of finitely-supported permutations of $\mathbb{N}$ is more complex, in that it requires more definitions to state precisely.

Let $\alpha : \mathbb{N} \to [0,1]$ with $\sum i \in \mathbb{N}\alpha(i) = 1$. We can interpret α as an (atomic) probability measure on $\mathbb{N}$, and can construct a product measure on $\mathbb{N}^{\mathbb{N}}$, denoting it by μ_α. The action of $S_{\infty,0}$ on $\mathbb{N}^{\mathbb{N}}$ given by

$$\pi : \xi \to \xi \circ \pi^{-1}$$

is ergodic with respect to μ_α.

Now for any $\xi \in \mathbb{N}^{\mathbb{N}}$, we define a partition η_ξ of $\mathbb{N}$ as follows. The part B_i^ξ is given by $\{n \in \mathbb{N} : \xi(n) = i\}$, and $\eta_\xi = \{B_i^\xi : i \in \mathbb{N}\}$. Under this correspondance, the measure μ_α gives rise to a measure ν_α on partitions of $\mathbb{N}$, and this measure is invariant under the action of $S_{0,\infty}$. An older result of Kingman (a variant of de Finetti's theorem) tells us that these are all of the invariant measures on the space of partitions of $\mathbb{N}$.

Let $S_{0,T}$ be the finitely supported permutations of the set T. We set P_α to be a direct sum of one copy of $\mathbb{Z}_2$ for each i where $\alpha(i) \neq 0$. For any $A \leq P_\alpha$, we can define a map

$$s_\xi : \bigoplus \alpha(i) \neq 0 S_{0,B_i^\xi} \to P_\alpha$$

by mapping each coordinate to its sign as a permutation. Then, for each $A \leq P_\alpha$, we have a map

$$f_A : \xi \to s_\xi^{-1}(A),$$

which, under mild conditions on ξ, transfers μ_α to an ergodic random measure ν_α^A on $S_{0,\infty}$.

THEOREM 6.4.10 ([**451**]). *The ergodic invariant random subgroups of $S_{\infty,0}$ are exactly the measures that arise as ν_α^A for some choices of α and A.*

PROOF. Now if $H \leq S_{\infty,0}$ has no nontrivial invariant partitions, then H must either be the full group or an alternating group. If H does have nontrivial invariant partitions but is transitive, all of the nontrivial minimal blocks must have the same finite length. A measure giving positive probability to such groups must induce a probability measure on partitions of $\mathbb{N}$ with countably many blocks of a fixed finite length. We can show that there are no such measures.

Let Π be the set of partitions of $\mathbb{N}$, and let $P : Sub(S_{\infty,0}) \to \Pi$ map a subgroup to the partition according to its orbits. Now for any ergodic invariant random subgroup ν, the map P induces from ν an ergodic random invariant partition, which must be the image of μ_α for some α. There is a ν-measure 1 set Ξ of subgroups H such that there exists a $\xi \in \mathbb{N}^\mathbb{N}$ so that the partition $P(H)$ is exactly the partition $\{B_i^\xi : i \in \mathbb{N}\}$.

It is possible, but not elementary, to show that there is a fixed subgroup $A \leq \bigoplus_{i\in\mathbb{N}} \mathbb{Z}_2$ such that ν concentrates on the same set of subgroups as ν_α^A. But the action of $S_{\infty,0}$ on this set is now uniquely ergodic, so $\nu = \nu_\alpha^A$. □

Moving toward somewhat larger classes of groups, Thomas and Tucker-Drob classified invariant random subgroups of groups they called *strictly diagonal limits of finite symmetric groups*. Let S_n and $S_{\ell n}$ be finite symmetric groups, and $h : S_n \to S_{\ell n}$. We say that h is a diagonal embedding if its image acts on each of its orbits on $\{1, \ldots, \ell n\}$ via its permutation representation. A countable locally finite group G is said to be a strictly diagonal limit of finite symmetric groups (an *SDS*-group) if and only if it is the union of an increasing chain of finite symmetric groups under diagonal embeddings.

To classify random subgroups of such a group, we consider the action of $\bigcup_{i\in\mathbb{N}} S_{n_i}$ on

$$X = \prod_{i\in\mathbb{N}} \{1, \ldots, n_i\}.$$

If we take μ to be the product of the uniform measures on the factors of X, we have a measure-preserving action of G on (X, μ). The diagonal action of G on X^r with the product measure is ergodic for any $r \in \mathbb{N}$, so that measure ν_r induced by the stabilizer map of this action is an invariant random measure. Now this group G is simple if and only if n_i is odd for all but finitely many i. If this condition fails, then $A(G) = \bigcup_{i\in\mathbb{N}} A_{n_i}$ is a subgroup, and we have a map $X^r \to Sub(G)$ by $\bar{x} \to G_{\bar{x}} \cap A(G)$. This map induces (from the product measure on X^r another invariant random subgroup λ_r on G.

THEOREM 6.4.11 ([**439**]). *If G is an SDS-group, then the nontrivial ergodic invariant random subgroups are exactly the following:*

(1) *If G is simple, $\{\nu_r : r \in \mathbb{N}\}$*

(2) *If G is not simple,* $\{\nu_r, \lambda_r : r \in \mathbb{N}\}$.

PROOF OUTLINE. We begin with characters. For finite groups, characters are neither more nor less than homomorphisms of the group into the multiplicative group of $\mathbb{C}$, or perhaps some other field. In the mid-1960's, Thoma argued that the naive translation of this concept was not the right notion of character for countable infinite groups, but that the right analogy was with positive-definite class functions [**437, 436**]. For this reason, those functions are often called *characters*, including in the paper by Thomas and Tucker-Drob, and we will follow this usage here.

For a countable discrete group Γ, a character of Γ is a map $\chi : \Gamma \to \mathbb{C}$ such that

(1) χ is conjugation invariant,
(2) $\chi(1) = 1$, and
(3) For any $\vec{\lambda} \in \mathbb{C}^n$ and any $\vec{g} \in \Gamma^n$, we have

$$\sum_{i,j=1}^{n} \lambda_i \overline{\lambda_j} \chi(g_j^{-1} g_i) \geq 0.$$

We let G be an SDS-group with the associated notation as in the discusssion preceding the statement of the theorem, and let ν be an ergodic random subgroup, induced by the stabilizer map of the ergodic action (G, Z, μ). We define χ to map an element $g \in G$ to the measure of its set of fixed points. We note that this map is, indeed a character.

Now by applying the Pointwise Ergodic Theorem (see Section 3.2.2) to the characteristic function of the set of fixed points F_g^Z of a particular $g \in G$, and denoting the G_i orbit in Z of $z \in Z$ by $\Omega_i(z)$, we have, for μ-almost every $z \in Z$ and for all $g \in G$,

$$\mu(F_g^Z) = \lim_{i \to \infty} \frac{\left| F_g^{\Omega_i(z)} \right|}{|\Omega_i(z)|}.$$

Let z satisfy the above equation, and let $H \leq G$ be its stabilizer. Denote by H_i the group $H \cap G_i$, so that

$$\chi(g) = \lim_{i \to \infty} \frac{\left| \left\{ s \in G_i | sgs^{-1} \in H_i \right\} \right|}{|G_i|}.$$

It is true, but not obvious, that there are only finitely many i such that H_i acts transitively on $\{1, \ldots, n_i\}$, and we let I be the set of $i \in \mathbb{N}$ such that this does *not* occur. For each i, we set r_i to be the maximum cardinality of an H_i-invariant subset $U \subseteq \{1, \ldots, n_i\}$ with $|U| \leq \frac{n_i}{2}$, and we set $r = \liminf r_i$, and let I_r be the set of $i \in I$ such that $r_i = r$, such that $i > r+1$, and such that $i > \max\{j \in I : r_j < r\}$. We note that I_r is closed upwards.

Now for each $i \in I_r$, we can prove that there is a unique H_i invariant subset $U_i \subseteq \{1, \ldots, i\}$ of cardinality r such that H_n acts transitively on the remainder of $\{1, \ldots, i\}$. We let B_H be the set of all $x \in x$ such that for any $i \in I_r$, we have $x \restriction_i \in U_i$.

At this point we begin to peek out of the business of defining one subgroup or stabilizer after another, as we notice that B_H is exactly the set of $x \in X$ with finite H-orbit. The map $H \mapsto B_H$ respects the action of G, so that it induces a G-invariant probability measure on the set of r-element subsets of X. If there is some $i \in I_r$ such that H_n acts nontrivially on U_i, then there are $x \neq y \in B_H$ where

xE_0y, with E_0 representing the Vitali relation of equality on all but finitely many entries. Since there is not a G-invariant Borel probability measure on

$$\{F \in [X]^r : E_0 \restriction_F \text{ is not identity}\},$$

the measure ν must concentrate on the H such that H_i acts trivially on U_i.

Knowing that H_i acts trivially on U_i, we can then see that if G is simple, then H_i is the symmetric group on $\{1, \ldots, i\} - U_i$ for all $i \in I_r$. Otherwise, either that condition obtains, or H_i is the alternating group on the same set for all but finitely many $i \in I_r$. It follows, again using the Pointwise Ergodic Theorem, that ν_r and (in the non-simple case) λ_r are the only possible invariant random subgroups. □

Thomas and Tucker-Drob use a similar method to handle another special case arising from the alternating groups. We say that G is an $L(Alt)$ group if $G = \bigcup_{i \in \mathbb{N}} A_{n_i}$. For each $i, j \in \mathbb{N}$, we denote by $s_{i,j}$ the number of orbits of G_i on $\{1, \ldots, j\}$, and by $f_{i,j}$ the number of trivial orbits of G_i on $\{1, \ldots, j\}$. We further write

$$e_i = n_i - (s_{(i-1),i} n_{i-1} + f_{(i-1),i}).$$

With this notation, we say that an $L(Alt)$ group G is an *almost diagonal limit* if and only if the following two conditions obtain:

(1) For all $i \in \mathbb{N}$, we have $s_{(i-1),i} > 0$, and
(2) $\sum_{i \in \mathbb{N}} \frac{e_i}{s_{0,i}} < \infty$

THEOREM 6.4.12 ([**440**]). *If G is an $L(Alt)$ group, then G has a nontrivial ergodic invariant random subgroup if and only if G can be expressed as an almost diagonal limit of finite alternating groups.*

PROOF OUTLINE. We say that an $L(Alt)$ group G as above has *linear natural orbit growth* if and only if $a_i = \lim_{j \to \infty} \frac{s_{i,j}}{n_j}$ is positive, and sublinear growth otherise. In any case, the limit is guaranteed to exist, and the condition is independent of the representation of G as a union of finite alternating groups. Further, a group with linear natural orbit growth is an almost diagonal limit.

By the Pointwise Ergodic Theorem, a G-invariant ergodic probability measure only exists in the case of linear natural orbit growth, and in that case the action is uniquely ergodic. □

The classification of the invariant random subgroups in this case is similar to the previous case.

On the other hand, some groups have obstructions to such a classificaiton. Of course, it is a traditionally subtle point to state that something is an obstruction to classification, in the sense that we normally do not have a precise mathematical definition of what would count as a classification (see a detailed discussion in [**239, 234, 92**]). Nevertheless, there are some features that seem, at least in a relative sense, to convince us that we will not have a simple classification.

We have seen that the set of invariant probability measures on a space constitutes a simplex with the ergodic measures as extremal points. A *Poulsen simplex* is a simplex in which the extremal points are dense in the simplex — a situation that at least suggests a high level of complexity in the space of measures.

THEOREM 6.4.13 ([**80**]). *If G is a nonabelian free group, then the space of invariant random subgroups of G concentrating on subgroups of infinite index constitute a Poulsen subsimplex.*

PROOF. Let F_r be the free group on r generators (denoted $S = \{s_1, s_2, \ldots, s_r\}$), and let ν be an invariant random subgroup concentrating on subgroups of infinite index. Then it suffices to construct a set of ergodic measures $(\nu_p : p \in (0,1))$ with $\lim_{p\to 0} \nu_p = \nu$.

For each $p \in (0,1)$, we construct a random directed graph Γ_p by sampling a subgroup $K \leq F_r$ according to ν, and constructing the Schreier coset graph $\Gamma_{p,0} = Sch(K,S)$ of K with generating set S, in the following way. The vertices of $\Gamma_{p,0}$ are exactly the right cosets of K in F_r, and there is an edge from Kg_1 to Kg_2 if and only if $g_2 = g_1 s_i$ for some i.

Now at stage n, with $\Gamma_{p,n}$ constructed, we designate a random subset $X_{p,n}$ of the vertices of $\Gamma_{p,n} - \Gamma_{p,n-1}$ (using $\Gamma_{p,-1} = \emptyset$) by including each vertex independently with probability p — that is, a Bernoulli p-percolation of $\Gamma_{p,n}$; see [**313**]. We now designate independently, for each $x \in X_{p,n}$, a ν-random $K_{p,x,n} \leq F_r$, and form $\Gamma_{p,n+1}$ as the disjoint union of $\Gamma_{p,n}$ and the disjoint union of the Schreier coset graphs of $K_{p,x,n}$ for all n, with an undirected edge from each $x \in X_{p,n}$ to $K_{p,x,n}1$. We define $\Gamma_p = \bigcup_{n\in\mathbb{N}} \Gamma_{p,n}$.

We form a new graph $\phi(\Gamma_p)$ as follows. For an edge (v,w) joining a vertex in $\Gamma_{p,n}$ to a vertex in $\Gamma_{p,n+1} - \Gamma_{p,n}$ for any n, we remove the edge, as well as the edges (v, vs_1) and (w, ws_1), and we add the edge (v, ws_1) and (w, vs_1). Now $\phi(\Gamma_p)$ is $Sch(H_p, S)$ for some $H_p \leq G$, and the random mapping $p \mapsto H_p$ induces an invariant random subgroup ν_p on G.

To show that $\nu_p \to \nu$, let $\epsilon > 0$ be given, and take p smal enough that with probability at least $1 - \epsilon$ the graph-metric ball of radius r centered at $K \in \phi(\Gamma_p)$ has trivial intersection with $X_{p,0}$. If that intersection is trivial, then this ball is isomorphic to the analogous ball in $Sch(K,S)$. Then for any subset $T \subseteq F_r$, the ν_p proabability that a subgroup intersects the ball of radius r about the identity in exactly F cannot differ from the ν-probability of this event by more than 2ϵ.

We could show that ν_p is ergodic by showing that any measurable invariant function is constant. We actually prove ergodicity in this way in a slightly different space, and then transfer the result.

For a bounded function f mapping rooted graphs with labels from S and degree bounded by $2r+1$, we use the Bernoulli condition in the construction of Γ_p to show that f factors through the map π taking a rooted labeled graph to the subgroup whose Schrieier coset is the same as the S-labeled component of the root. Say $f = g \circ \pi$. By the symmetry in the induction step of the construction of Γ_p, we have $g(K_{p,x}) = g(K_p)$, so that g is constant ν-almost everywhere. This property pulls back, so that ν_p has the same property. □

As we conclude the technical core of this two-chapter sequence on random structures, it is fitting that we do so on this note. We began with a consideration of random graphs, and we end here, constructing invariant random subgroups to order using an intermediate construction of (appropriately labeled) random graphs.

6.5. Probabilistic Boolean Networks

6.5.1. An Application: Gene Regulatory Networks. At this point, we engage in a brief digression from the pure mathematics that tends to characterize mathematical logic and from the idealized models of computing that tend to characterize complexity theory, and consider an application problem on which the material of the present chapter has significant promise. Of course, logic is always closely involved in systems of specification and control, even if that interpretation is rarely emphasized by mathematicians. On the other hand, while domain-specific logics for such applications are common currency in certain parts of software engineering and electrical engineering, there seem to have been few explicit applications of classical, infinitary, or continuous model theory — not to say computability or the existence of invariant measures — to this area.

The problem of sequencing an archetypal human genome was declared solved by biologists in 2003, although other organisms had already been sequenced. Both in principle and in fact, this is a major accomplishment of science, but the sequencing of a genome has a fundamentally Linnean character: We get a list, perhaps well organized, with a good deal of additional data, of what genes there are. Of course, there is interpretation of this data which is literally Darwinian, but the question of how genes interact to carry out the life processes of a cell remains a broad area of active research. Questions of which genes are expressed in what strength under what conditions seem to be of special importance. For a gene to be "expressed" means roughly that the protein it encodes is produced in the cell, and while each individual step of transcription and translation — the whole process from gene to protein — happens or does not happen as a binary state, this process can happen quantitatively more or less often.

For the present purposes, we might first represent the genome as a directed graph in which each vertex represents a gene. Some genes produce code "promoter" proteins which cause other genes to be expressed more. We might start by drawing an edge from gene a to gene b if and only if a promotes b. There are also "inhibitors" which decrease the expression of another gene. This would give us a second "color" of edge in the directed graph. The topology of the graph can be intricate. Positive and negative feed-back and feed-forward loops are not uncommon. Of course, the strength of these actions can also vary quantitatively. [**13**]

A full understanding of the gene regulatory network of an organism is a strong understanding of its biology, and is a difficult target even for a relatively simple species. It still does not represent complete understanding, though. One might still want to include the interaction of this system with the environment. In many plant species, for example, there is a characteristic subnetwork of the interactome that activates under conditions of cold stress, and another characteristic subnetwork that activates under conditions of drought stress. How does each of these subnetworks improve the organism's survival? Perhaps more interesting from a mathematical standpoint, is the intersection of these two subnetworks what we might find by randomly intersecting two sets of vertices, or does this intersection reflect a biological function common to the two responses?

There are sensible ways to model these systems, and many of them have the potential to be reflected in the random structures described in the present chapter.

6.5.2. Defining Probabilistic Boolean Networks. An initial attempt at modeling gene regulatory networks would be to use a vertex for each gene with an edge between each pair of genes, weighting each edge according to the correlation coefficient of the expressions of the two genes. We might, to simplify the model and avoid overfitting, choose a threshold τ, and discard all edges of absolute weight less than τ. It is entirely possible, of course, that two genes may be weakly correlated as a pair, but still in the same connected component in this graph.

This representation is certainly adequate for some biological purposes, but it also has shortcomings. In paticular, it does not represent the dynamic nature of gene expression, and it is, in the end, a purely empirical model, showing little direct reflection of the phenomena under study.

A second approach to modeling is to again use genes as vertices, but to use a directed edge between genes to represent the action of one gene to up- or down-regulate another. We can add structure by giving each vertex a Boolean state, which may change over time, to reflect the expression (or lack thereof) of the respective gene. We might specify that a gene v takes on, in stage $t+1$, the value of a given Boolean function, called the *coupling function* of the stage t values of its predecessors.

There remains the problem of specifying the Boolean functions that determine the evolution of the network. By analogy with hypothesis tests in which we generate a probability model reflecting the null hypothesis, we can define a model of the random variation of the networks by selecting a Boolean function at random. It is customary to first specify, for each vertex v, a parameter K_v, intending that the value of v at stage t will be determined by a Boolean function of at most K_v other vertices. Often, a global value K is set so that $K_v = K$ for all v.

From here, we can specify a probability distribution on the (finitely many) Boolean functions on the vertex set depending on at most K_v arguments. Different models for this abound, but the uniform distribution is commonly used. It is still at issue whether the coupling functions should be fixed for all time or should be re-selected at each step, and, in the latter case, whether these choices should be independent. We assume that all vertices update synchronously. This model was originally introduced in [**262, 261, 263, 264**], and a survey of the various choices possible in the model formulation can be found in [**14**].

This level of modeling can represent, for instance, stability properties. In many systems, for instance, a perturbation of the expression of a few genes will not change the long-term state of the model.

Of course, at this point, the Bayesian networks of Section 2.3.1 will come to mind. There, too, the value at one vertex depends, at each stage of evaluation, on the values of several inputs. We have, as yet, made no independence requirements on our Boolean networks. Moreover, aside from the choice of coupling functions — really, a choice of network topology — the dynamics of a Boolean network are completely deterministic.

A major challenge lies in inferring the structure of a Boolean network from data. In particular, in spite of even the more generous assumptions in the "annealed" models that allow coupling functions to be re-selected at every time step, the real biochemical processes being modeled are genuinely not deterministic. Molecules become more or less likely to bind with one another. Worse, available measurements are likely to represent a small sample for the purposes of model inference.

Probabilistic Boolean Networks, described in detail in [**408, 409**], represent an attempt to address these issues by fusing the concepts of Boolean networks and Bayesian networks.

DEFINITION 6.5.1. A Probabilistic Boolean Network consists of the following:

- A set $V = \{v_i : i \in I\}$ of vertices.
- For every $i \in I$, a set $F_i = \left\{f^(i)_j : j \in J_i\right\}$ of Boolean functions, each with domain some subset of V.
- A joint probability distribution P on the random variables F_i

A probabilisitc Boolean network is said to be independent if and only if the random variables $\{F_i : i \in I\}$ are independent. The dynamics of a network, independent or not, constitute a Markov chain whose transitions are determined by the distribution P.

An important feature of this model is that it incorporates transition of network topology as a random event, according as the same or different values of F_i are realized at each time step.

There remains the question of the relationship of probabilistic Boolean networks to Bayesian networks. An initial concern is the temporal structure. Natively, Bayesian networks only represent the joint probability distribution on the values at the vertices. This can be corrected by a straightforward adaptation, though.

DEFINITION 6.5.2. A dynamic Bayesian network $(B_1, B_\rightarrow)$ consists of the following:

- A Bayesian network B_0 (see Section 2.3.1) on the vertex set $\{v_{i,0} : i \in I\}$, defining the initial state $(v_{i,0} : i \in I)$ of the system, and
- A Bayesian network $B_\rightarrow$ on the vertex set $(v_{i,t} : i \in I; t \in \mathbb{N})$, where the predecessors of $v_{i,t}$ are all included in $(v_{j,t-1} : j \in I)$, defining the transition probabilities $P(v_{i,t}|v_{i,t-1})$.

It is worthwhile to note that dynamical Bayesian networks have both hidden Markov models and Kalman filters as special cases [**350**].

The principle connection between probabilistic Boolean networks and dynamic Bayesian networks is given by the following:

THEOREM 6.5.3 ([**293**]). *The following properties of a joint probability distribution P are equivalent:*

(1) *P can be represented by an independent probabilistic Boolean network.*
(2) *P can be represented by a binary-valued dynamic Boolean network.*

6.5.3. Some Problems on Probabilistic Boolean Networks. It is becoming a routine laboratory technique to gather data on expression of genes. A major biological goal in this area is to use this data on gene expression to infer the network that produces it. Once the network structure is known with high certainty, one can then attempt to design interventions to modify the state of the network, for instance, to treat cancer.

The inference problem is fundamentally one of machine learning, which is properly the subject of Chapter 7. However, some results on this particular problem may help to motivate the more general approach of that chapter.

The *state* of a Boolean network at a particular time is the assignment of truth values to its vertices. Perhaps the most basic piece of inference in Boolean networks

is the consistency problem: Given a particular vertex v and known examples of its value under some truth assignments to other vertices, is there a Boolean function that could be the coupling function of v in a (deterministic) Boolean network?

As long as the number of vertices is finite (obviously true in the case of genetic data), the search space of possible Boolean functions is also finite, so that this problem is obviously effectively solvable. One might still wish for a more efficient solution, and, in fact, if we have a bound on the number of predecessors of v, one exists. If T is the set of states of the predecessors (encoded as an element of 2^n) prescribed to be true, and F is the set of states prescribed to be false, the solution to the problem of existence of a common extension can be decided in the following way: we first sort both T and F, and then compare the two lists to see if they contain any common element. If they do not, there is a common extension. If $|T \cup F| = m$, this algorithm halts in time $O(m \log mp(n))$, where p is a polynomial.

Of course, we generally want more than to know that a solution exists. To find all consistent functions, we use the following procecdure. We seek to learn a boolean function $f : 2^n \to 2$. We begin with $f^{-1}(\bar{x})$ undefined for all $\bar{x}$. If, at any stage, we find an $\bar{x} \in T \cap F$, we halt and declare the data inconsistent. At stage s, we take the least $\bar{x}$ such that $\bar{x} \in T \cup F$, and set

$$f^s(\bar{x}) = \begin{cases} 0 & \text{if } \bar{x} \in F \\ 1 & \text{if } \bar{x} \in T \end{cases}$$

We set $f^- = \bigcup_s f^s$, and any extension of f^- to a total function on 2^n is then consistent. The time complexity to generate f^- is $O(mp(n))$. This procedure, with several added features, can be found in [**294**].

Those experienced with applied data analysis will find it no surprise that sometimes there is no consistent extension. Certainly there is measurement noise in the complex processes of gathering the biochemical data, and there are other sources of inconsistency, as well. To deal with this, we assign a nonnegative weight $w : T \cup F \to \mathbb{R}$, and attempt to find a Boolean function f minimizing the total weight of all states $\bar{x}$ that f misclassifies. This problem, too, is solvable in polynomial time.

Frequently samples are taken in a way that give only one time point of data (e.g. clinical samples from a human tumor). Since the dynamical system represented by either a Boolean or probabilisitc Boolean network constitutes an ergodic Markov chain, there must exist a stationary distribution — that is, a distribution which is fixed by the transition matrix. Under biologically reasonable hypotheses, most of the probability mass of the steady state distribution will be in the attractors of the dynamical system. Where only a single time point is available, the customary treatment is to assume that the sample represents a stationary state of the network, and thus, with high probability, comes from the attractor states.

Obviously, as we pass from deterministic Boolean networks to probabilisitc Boolean networks, these problems become much more difficult. The dynamics surrounding the steady state assumption give a clue to a reeasonable approach. The following algorithm, introduced in [**358**], generates a Boolean network, but can be leveraged to generate a probabilistic Boolean network.

We begin by generating, at random, a set of k attractor states and, for each vertex v, a set W_v of predictor vertices. We then check compatibility: is the assumption that the attractor states generated are attractors consistent with the

choice of W_v? If not, we generate new sets W_v. Otherwise, we use the attractors generated to determine sets T_v, F_v of known entries in the transition function f_v. The rest of the function f is generated at random from Boolean functions on W_v consistent with T_v and F_v. If the state transition diagram associated with the resulting network has cycles, we resample the functions f. Otherwise, we return the network generated.

This procedure generates a deterministic Boolean network consistent with the data. To generate a probabilistic Boolean network, we run this randomized algorithm many times, sampling from it to infer the distribution of the coupling functions.

Of course, the literature on inference of, and intervention in, probabilistic Boolean networks is large. The books [**408, 409**] are good entry points to the literature.

Bibliography

1. M. Abért, N. Bergeron, I. Biringer, T. Genander, N. Nikolov, J. Raimbault, and I. Samet, *On the grown tof L^2-invariants for sequences of lattices in Lie groups*, Annals of Mathematics **185** (2017), 711–790.
2. M. Abért, Y. Glasner, and B. Virág, *Kesten's theorem for invariant random subgroups*, Duke Mathematical Journal **163** (2014), 465–488.
3. N. Ackerman, C. Freer, A. Kwiatowska, and R. Patel, *A classification of orbits admitting a unique invariant measure*, Annals of Pure and Applied Logic **168** (2017), 19–36.
4. N. Ackerman, C. Freer, J. Nešetřil, and R. Patel, *Invariant measures via inverse limits of finite structures*, European Journal of Combinatorics **52** (2016), 248–289.
5. N. Ackerman, C. Freer, and R. Patel, *Invariant measures concentrated on countale structures*, Forum of Mathematics, Sigma **4** (2016), 1–59.
6. ———, *Countable infinitary theories admitting an invariant measure*, preprint, 2017.
7. ———, *The entropy function of an invariant measure*, Proceedings of the 14th and 15th Asian Logic Conferences, World Scientific Publishing, 2019, pp. 3–34.
8. N. Ackerman, C. Freer, and D. Roy, *On the computability of conditional probability*, Journal of the Association for Computing Machinery **66** (2019), 23:1–23:40.
9. ******E. W. Adams, *The logic of conditionals*, Reidel, 1975.
10. E. W. Adams, *A primer of probability logic*, CSLI Lecture Notes, no. 68, CSLI Publications, 1999.
11. S. Adams and A. S. Kechris, *Linear algebraic groups and countable Borel equivalence relations*, Journal of the American Mathematical Society **13** (2000), 909–943.
12. Scott Adams, *Dilbert*, (2001), October 25, `https://dilbert.com/strip/2001-20-25`.
13. B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular biology of the cell*, 5th ed., Garland, 2008.
14. M. Aldana, S. Coppersmith, and L. P. Kadanoff, *Boolean dynamics with random couplings*, Perspectives and Problems in Nonlinear Science, Springer, 2003, pp. 23–89.
15. K. Allen, L. Bienvenu, and T. A. Slaman, *On zeros of Martin-Löf random Brownian motion*, Journal of Logic and Analysis **6** (2014), 1–36.
16. J.-P. Allouche and J. Shallit, *Automatic sequences*, Cambridge, 2003.
17. N. Alon, R. A. Duke, H. Lefmann, V. Rödel, and R. Yuster, *The algorithmic aspects of the regularity lemma*, Journal of Algorithms **16** (1994), 80–109.
18. N. Alon and J. H. Spencer, *The probabilistic method*, third ed., Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley, 2008.
19. R. Alvir, W. Calvert, G. Goodman, V. Harizanov, J. Knight, A. Morozov, R. Miller, A. Soskova, and R. Weisshaar, *Interpreting a field in its Heisenberg group*, preprint, 2020.
20. J. J. Andrews and M. L. Curtis, *Free groups and handlebodies*, Proceedings of the American Mathematical Society **16** (1965), 192–195.
21. U. Andrews, I. Goldbring, and H. J. Keisler, *Definable closure in randomizations*, Annals of Pure and Applied Logic **166** (2015), 325–341.
22. ———, *Independence in randomizations*, Journal of Mathematical Logic **19** (2019), 1950005.
23. U. Andrews and H. J. Keisler, *Separable models of randomizations*, Journal of Symbolic Logic **80** (2015), 1149–1181.
24. U. Andrews, S. Lempp, J. S. Miller, K. M. Ng, L. San Mauro, and A. Sorbi, *Universal computably enumerable equivalence relations*, The Journal of Symbolic Logic **79** (2014), 60–88.
25. U. Andrews and A. Sorbi, *The complexity of index sets of classes of computably enumerable equivalence relations*, The Journal of Symbolic Logic **81** (2016), 1375–1395.

26. A. Arana, *Logical and semantic purity*, Protosociology **25** (2008), 36–48.
27. A. Arnould and P. Nicole, *The Port Royal Logic*, Gordon, 1861.
28. S. Arora and B. Barak, *Computational complexity*, Cambridge, 2009.
29. Е. А. Асарин and А. В. Покровский, Применение колмогоровской сложности к анализу динамики упровляемых систем, Автоматика и Телемеханика **1** (1986), 25–33.
30. M. Aschenbrenner, A. Dolich, D. Haskell, D. Macpherson, and S. Starchenko, *Vapnik-Chervonenkis density in some theories without the independence property, II*, Notre Dame Journal of Formal Logic **54** (2013), 311–363.
31. ———, *Vapnik-Chervonenkis density in some theories without the independence property, I*, Transactions of the American Mathematical Society **368** (2016), 5889–5949.
32. C. J. Ash and J. F. Knight, *Computable structures and the hyperarithmetical hierarchy*, Studies in Logic and the Foundations of Mathematics, vol. 144, Elsevier, 2000.
33. K. B. Athreya, J. M. Hitchcock, J. H. Lutz, and E. Mayordomo, *Effective strong dimension in algorithmic information and computational complexity*, SIAM Journal on Computing **37** (2007), 671–705.
34. J. Avigad, *Inverting the Furstenberg correspondence*, Discrete and Continuous Dynamical Systems **32** (2012), 3421–3431.
35. J. Avigad, J. Hözl, and L. Serafin, *A formally verified proof of the central limit theorem*, Journal of Automated Reasoning **59** (2017), 389–423.
36. J. Ax, *The elementary theory of finite fields*, Annals of Mathematics **85** (1968), 239–271.
37. L. Babai, *Trading group theory for randomness*, STOC '85: Proceedings of the seventeenth annual ACM symposium on Theory of Computing, 1985, pp. 421–429.
38. A. Baker, *Transcendental number theory*, Cambridge Mathematical Library, Cambridge, 1990.
39. J. T. Baldwin and S. Shelah, *Randomness and semigenericity*, Transactions of the American Mathematical Society **349** (1997), 1359–1376.
40. S. Banach, *Sur le problème de la measure*, Fundamenta Mathematicae **4** (1923), 7–33.
41. A.-L. Barabási and R. Albert, *Emergence of scaling in random networks*, Science **286** (1999), 509–512.
42. G. Barmpalias, D. Cenzer, and C. P. Porter, *The probability of a computable output from a random oracle*, preprint, 2016.
43. ———, *Random numbers as probabilities of machine behavior*, Theoretical Computer Science **673** (2017), 1–18.
44. George Barmpalias and Andrew Lewis-Pye, *Differences of halting probabilities*, preprint, 2016.
45. L. Barreira, *Dimension and recurrence in hyperbolic dynamics*, Progress in Mathematics, no. 272, Birkhäuser, 2008.
46. L. Bartholdi, *Counting paths in groups*, L'Enseignement Mathématique **45** (1999), 83–131.
47. N. A. Bazhenov and B. S. Kalmurzaev, *On dark computably enumerable equivalence relations*, Siberian Mathematical Journal **59** (2018), 22–30.
48. V. Becher, Y. Bugeaud, and T. A. Slaman, *On simply normal numbers to different bases*, Mathematische Annalen **364** (2016), 125–150.
49. V. Becher, O. Carton, and P. A. Heiber, *Normality and automata*, Journal of Computer and System Sciences **81** (2015), 1592–1613.
50. V. Becher and P. A. Heiber, *Normal numbers and finite automata*, Theoretical Computer Science **477** (2013), 109–116.
51. J. Beck, *An algorithmic approach to the Lovász local lemma I*, Random Structures and Algorithms **2** (1991), 343–365.
52. H. Becker and A. S. Kechris, *Borel actions of Polish groups*, Bulletin of the American Mathematical Society **28** (1993), 334–341.
53. O. Becker, A. Lubotzky, and A. Thom, *Stability and invariant random subgroups*, Duke Mathematical Journal **168** (2019), 2207–2234.
54. S. Ben-David, D. Pál, and S. Shalev-Shwartz, *Agnostic online learning*, Conference on Learning Theory (COLT), 2009.
55. I. Ben Yaacov, *Schrödinger's cat*, Israel Journal of Mathematics **153** (2006), 157–191.
56. ———, *Continuous and random Vapnik-Chervonenkis classes*, Israel Journal of Mathematics **173** (2009), 309–333.

57. ———, *On theories of random variables*, Israel Journal of Mathematics **194** (2013), 957–1012.
58. I. Ben Yaacov, A. Berenstein, C. W. Henson, and A. Usvyatsov, *Model theory for metric structures*, Model theory with applications to algebra and analysis, vol. 2, London Mathematical Socieity Lecture Note Series, no. 350, Cambridge, 2008, pp. 315–429.
59. I. Ben Yaacov and H. J. Keisler, *Randomizations of models as metric structures*, Confluentes Mathematici **1** (2009), 197–223.
60. I. Ben Yaacov and A. P. Pedersen, *A proof of completeness for continuous first-order logic*, Journal of Symbolic Logic (2010), 168–190.
61. I. Ben Yaacov and A. Usvyatsov, *Continuous first order logic and local stability*, Transactions of the American Mathematical Society **362** (2010), 5213–5259.
62. C. Bernardi and A. Sorbi, *Classifying positive equivalence relations*, The Journal of Symbolic Logic **48** (1983), 529–538.
63. A. Beros, *Learning theory in the arithmetic hierarchy*, Journal of Symbolic Logic **79** (2014), 908–927.
64. Ö. Beyarslan, *Random hypergraphs in pseudofinite fields*, Journal of the Institute of Mathematics of Jussieu **9** (2010), 29–47.
65. L. Bienvenu, *Game-theoretic approaches to randomness: unpredictability and stochasticity*, Ph.D. thesis, Université de Provence, 2008.
66. L. Bienvenu, A. Day, M. Hoyrup, I. Mezhirov, and A. Shen, *A constructive version of Birkhoff's ergodic theorem for Martin-Löf random points*, Information and Computation **210** (2012), 21–30.
67. C. M. Bishop, *Pattern recognition and machine learning*, Information Science and Statistics, Springer, 2006.
68. L. Blum and M. Blum, *Toward a mathematical theory of inductive inference*, Information and control **28** (1975), 125–155.
69. A. Blumer, A. Ehrenfeucht, D. Haussler, and M. K. Warmuth, *Learnability and the Vapnik-Chervonenkis dimension*, Journal of the ACM **36** (1989), 929–965.
70. B. Bollobás, *Random graphs*, 2nd ed., Cambridge Studies in Advanced Mathematics, no. 73, Cambridge, 2001.
71. G. Boole, *An investigation of the laws of thought, on which are founded the mathematical theories of logic and probabilities*, Macmillan, 1851.
72. W. W. Boone, *Certain simple, unsolvable problems of group theory V, VI*, Indagationes Mathematicae **60** (1957), 22–27, 227–232.
73. ———, *The word problem*, Proceedings of the National Academy of Sciences of the USA **44** (1958), 1061–1065.
74. M. Borda, *Fundamentals in information theory and coding*, Springer, 2011.
75. A. Borel, *Density properties for certain subgroups of semi-simple groups wihtout compact components*, Annals of Mathematics **72** (1960), 179–188.
76. M. E. Borel, *Les probabilités dénombrables et leurs applications arithmétiques*, Rendiconti del Circolo Matematico di Palermo (1884–1940) **27** (1909), 247–271.
77. C. Borgs, J. T. Chayes, L. Lov'asz, V. T. Sos, and K. Vesztergombi, *Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing*, Advances in Mathematics **219** (2008), 1801–1851.
78. C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi, *Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing*, Advances in Mathematics **219** (2008), 1801–1851.
79. K. H. Borgwardt, *The simplex method: A probabilistic analysis*, Algorithms and Combinatorics, no. 1, Springer, 1987.
80. L. Bowen, *Invariant random subgroups of the free group*, Groups, Geometry, and Dynamics **9** (2015), 891–916.
81. G. Boxall, *NIP for some pair-like theories*, Archive for Mathematical Logic **50** (2011), 353–359.
82. M. Braverman, *Parabolic Julia sets are polynomial time computable*, Nonlinearity **19** (2006), 1383–1401.
83. M. Braverman and M. Yampolsky, *Computability of julia sets*, Algorithms and Computation in Mathematics, no. 23, Springer, 2009.

84. J. Brody and M. C. Laskowski, *Rational limits of Shelah-Spencer graphs*, The Journal of Symbolic Logic **77** (2012), 580–592.
85. Y. Bugeaud, *Distribution modulo one and Diophantine approximation*, Cambridge Tracts in Mathematics, no. 193, Cambridge, 2012.
86. S. Buss and M. Minnes, *Probabilistic algorithmic randomness*, The Journal of Symbolic Logic **78** (2013), 579–601.
87. D. Cai, N. Ackerman, and C. Freer, *An iterative step-function estimator for graphons*, preprint, 2015.
88. W. Calvert, *Metric structures and probabilistic computation*, Theoretical Computer Science **412** (2011), 2766–2775.
89. ———, *PAC learning, VC dimension, and the arithmetic hierarchy*, Archive for Mathematical Logic **54** (2015), 871–883.
90. W. Calvert, V. Harizanov, and A. Shlapentokh, *Turing degrees of isomorphism types of algebraic objects*, The Journal of the London Mathematical Society **75** (2007), 273–286.
91. ———, *Random fields*, preprint, 2020.
92. W. Calvert and J. F. Knight, *Classification from a computable viewpoint*, Bulletin of Symbolic Logic **12** (2006), 191–219.
93. P. J. Cameron, *Transitivity of permutation groups on unordered sets*, Mathematische Zeitschrift **148** (1976), 127–139.
94. C. Carathéodory, *Über das lineare Mass von Punktmengen — eine Verallgemeinerung des Längenbegriffs*, Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-phisikalisch Klasse **1914** (1914), 404–426.
95. J. Case, S. Jain, and F. Stephan, *Effectivity questions for Kleene's recursion theorem*, Theoretical Computer Science **733** (2018), 55–70.
96. J. Case and C. Lynes, *Machine inductive inference and language identification*, International Colloquium on Automata, Languages, and Programming, 1982, pp. 107–115.
97. T. Ceccherini-Silberstein and M. Coornaert, *Cellular automata and groups*, Springer Monographs in Mathematics, Springer, 2010.
98. D. Cenzer, Π^0_1 *classes in computability theory*, Handbook of Computability, Studies in Logic and the Foundations of Mathematics, no. 140, Elsevier, 1999, pp. 37–85.
99. C. Chabauty, *Limite d'ensembles et géométrie des nombres*, Bulletin de la S. M. F. **78** (1950), 143–151.
100. G. J. Chaitin, *A theory of program size formally identical to information theory*, Journal of the Association for Computing Machinery **22** (1975), 329–340.
101. ———, *Algorithmic information theory*, Cambridge Tracts in Theoretical Computer Science, no. 1, Cambridge, 1987.
102. A. S. Charles, *Interpreting deep learning: The machine learning Rorschach test?*, preprint, 2018.
103. H. Chase and J. Freitag, *Model theory and machine learning*, preprint, 2018.
104. Z. Chatzidakis, *Théorie des modèles des corps valués*, lecture notes, 2008.
105. Z. Chatzidakis, L. van den Dries, and A. Macintyre, *Definable sets over finite fields*, Journal für die reine und angewandte Mathematik **427** (1992), 107–135.
106. G. Cherlin and E. Hrushovski, *Finite structures with few types*, Annals of Mathematics Studies, no. 152, Princeton University Press, 2003.
107. A. Chernikov and S. Starchenko, *Definable regularity lemmas for NIP hypergraphs*, Quarterly Journal of Mathematics **72** (2021), 1401–1433.
108. N. Chomsky, *Three models for the description of language*, IRE Transactions on Information Theory **2** (1956), 113–124.
109. ———, *Knowledge of language: Its nature, origin, and use*, Convergence, Praeger Scientific, 1986.
110. F. Chung, *On concentrators, superconcentrators, generalizers, and nonblocking networks*, The Bell System Technical Journal **58** (1978), 1765–1777.
111. F. Chung and L. Lu, *Complex graphs and networks*, CBMS Regional Conference Series in Mathematics, no. 107, American Mathematical Society, 2006.
112. F. Chung, L. Lu, T. G. Dewey, and D. J. Galas, *Duplication models for biological networks*, Journal of Computational Biology **10** (2003), 677–687.

113. B. Cisma, D. D. Dzhafarov, D. R. Hirschfeldt, C. G. Jockusch, R. Solomon, and L. B. Westrick, *The reverse mathematics of Hindman's theorem for sums of exactly two elements*, Computability **8** (2019), 253–263.
114. K. J. Compton, *Laws in logic and combinatorics*, Algorithms and Order, NATO ASI Series C, no. 255, Kluwer, 1989, pp. 353–383.
115. G. Conant and A. Pillay, *Pseudofinite groups and VC-dimension*, preprint, 2018.
116. A. Condon, *The complexity of stochastic games*, Information and Computation **96** (1992), 203–224.
117. C. T. Conley, A. S. Kechris, and B. D. Miller, *Stationary probability measures and topological realizations*, Israel Journal of Mathematics **198** (2013), 333–345.
118. C. T. Conley and B. D. Miller, *Measure reducibility of countable Borel equivalence relations*, Annals of Mathematics **185** (2017), 347–402.
119. D. Conlon and J. Fox, *Bounds for graph regularity and removal lemmas*, Geometric and Functional Analysis **22** (2012), 1191–1256.
120. A. Connes and B. Weiss, *Property T and asymptotically invariant sequences*, Israel Journal of Mathematics **37** (1980), 209–210.
121. S. D. Conte and C. de Boor, *Elementary numerical analysis*, 3rd ed., International Series in Pure and Applied Mathematics, McGraw-Hill, 1980.
122. O. Cooley, W. Fang, D. Del Giudice, and M. Kang, *Subcritical random hypergraphs, high-order components, and hypertrees*, 2019 Proceedings of the Sixteenth Workshop on Analytic Algorithmics and Combinatorics (ANALCO), 2019, pp. 111–118.
123. M. Coornaert, *Topoligcal dimension and dynamical systems*, Universitext, Springer, 2015.
124. T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms*, 3rd ed., MIT Press, 2009.
125. R. T. Cox, *Probability, frequency, and reasonable expectation*, American Journal of Physics **14** (1946), 1–13.
126. B. F. Csima, V. S. Harizanov, R. Miller, and A. Montalban, *Computability of Fraïssé limits*, Journal of Symbolic Logic **76** (2011), 66–93.
127. R. W. R. Darling and J. R. Norris, *Structure of large random hypergraphs*, The Annals of Applied Probability **15** (2005), 125–152.
128. S. Ben David, P. Hrubes, S. Moran, A. Shpilka, and A. Yehudayoff, *Learnability can be undecidable*, Nature Machine Intellignece **1** (2019), 44–48.
129. A. P. Dawid, *Probability, causality, and the empirical world: a Bayes-de Finetti-Popper-Borel synthesis*, Statistical Science **10** (2004), 44–57.
130. B. de Finetti, *Funzione caratteristica di un fenomeno aleatorio*, Memorie della R. Accademia dei Lincei **4** (1930), 86–133.
131. ———, *Foresight: its logical laws, its subjective sources*, Breakthroughs in Statisitcs, Springer, 1992, Translated by H. E. Kyberg, Jr.; Original published 1937, pp. 134–174.
132. ———, *A translation of 'The characteristic function of a random phenomenon' by Bruno de Finetti*, D. Alvarez-Melis and T. Broderick, translators. arXiv:1512.01229, 2015.
133. K. de Leeuw, E. F. Moore, C. E. Shannon, and N. Shapiro, *Computability by probabilistic machines*, Automata Studies, Annals of Mathematics Studies, no. 34, Princeton, 1956, pp. 183–212.
134. M. Dehn, *Über unendliche diskontinuierliche Gruppen*, Mathematische Annalen **71** (1911), 116–144.
135. ———, *Transformation der kurven auf zweiseitigen flächen*, Mathematische Annalen **72** (1912), 413–421.
136. A. P. Dempster, *Upper and lower probabilities induced by a multivalued mapping*, The Annals of Mathematical Statistics **38** (1967), 325–339.
137. O. Demuth, *On constructive pseudonumbers*, Commentationes Mathematicae Universitatis Carolinae **16** (1975), 315–331.
138. M. Detlefsen and A. Arana, *Purity of methods*, Philosophers' Imprint **11** (2011), 1–20.
139. P. Diaconis and S. Janson, *Graph limits and exchangeable random graphs*, Rendiconti di Matematica, Serie VII **28** (2008), 33–61.
140. A. Ditzen, *Definable equivalence relations on Polish spaces*, Ph.D. thesis, California Institute of Technology, 1992.
141. A. Dolich, D. Lippel, and J. Goodrick, *dp-minimal theories: basic facts and examples*, Notre Dame Journal of Formal Logic **52** (2011), 267–288.

142. M. D. Donsker, *An invariance principle for certain probability limit theorems*, Memoirs of the American Mathematical Society **6** (1951), 1–12.
143. R. Dougherty, S. Jackson, and A. S. Kechris, *The structure of hyperfinite Borel equivalence relations*, Transactions of the American Mathematical Society **341** (1994), 193–225.
144. R. G. Downey and E. J. Griffiths, *Schnorr randomness*, The Journal of Symbolic Logic **69** (2004), 533–554.
145. R. G. Downey and D. R. Hirschfeldt, *Algorithmic randomness and complexity*, Theory and Applications of Computability, Springer, 2010.
146. R. G. Downey, C. G. Jockusch Jr., and P. E. Schupp, *Asymptotic density and computably enumerable sets*, Journal of Mathematical Logic **13** (2013), 1350005.
147. A. Dudko and M. Yampolsky, *On computational complexity of Cremer Julia sets*, Fundamenta Mathematicae **252** (2021), 343–353.
148. R. M. Dudley, *Central limit theorems for empirical measures*, The annals of probability **6** (1978), 899–929.
149. J.-L. Duret, *Les corps faiblement algebriquement clos non separablement clos ont la properiete d'independance*, Model Theory of Algebra and Arithmetic, Lecture Notes in Mathematics, no. 834, Springer, 1980, pp. 136–162.
150. M. Džamonja and I. Tomašić, *Graphons arising from graphs definable over finite fields*, preprint, 2017.
151. P. D. Eastman, *Are you my mother?*, Random House, 1960.
152. G. Edgar, *Measure, topology, and fractal geometry*, second ed., Undergraduate Texts in Mathematics, Springer, 2008.
153. H. G. Eggleston, *Sets of fractional dimensions which occur in some problems of number theory*, Proceedings of the London Mathematical Society **54** (1952), 42–93.
154. K. Eickmeyer and M. Grohe, *Randomisation and derandomisation in descriptive complexity theory*, Logical Methods in Computer Science **7** (2011), 1–24.
155. G. Elek and B. Szegedy, *A measure-theoretica approach to the theory of dense hypergraphs*, Advances in Mathematics **231** (2012), 1731–1772.
156. R. Elwes, *Asymptotic classes of finite structures*, Journal of Symbolic Logic **72** (2007), 418–438.
157. H. B. Enderton, *A mathematical introduction to logic*, Academic Press, 1972.
158. I. Epstein, *Orbit inequivalent actions of non-amenable groups*, preprint, 2008.
159. P. Erdős, D. J. Kleitman, and B. L. Rothschild, *Asymptotic enumeration of K_n-free graphs*, Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973), vol. 2, Acad. Naz. Lincei, 1976, pp. 19–27.
160. P. Erdős and L. Lovász, *Problems and results on 3-chromatic hypergraphs and some related questions*, Infinite and Finite Sets, vol. II, Colloq. Math. Soc. János Bolyai, no. 10, North-Holland, 1975, pp. 609–627.
161. P. Erdős and A. Rényi, *On the evolution of random graphs*, Matematikai Kutató Intézet Közelményei **A** (1960), 17–60.
162. P. Erdős and A. Rényi, *On random graphs I*, Publicationes Mathematicae Debrecen **6** (1959), 290–297.
163. Ю. Л. Ершов, Позитивные Еквивалентности, Алгебра и Логика **10** (1971), 620–650.
164. R. Fagin, *Generalized first-order spectra and polynomial-time recognizable sets*, SIAM–AMS Proceedings, vol. 7, 1974.
165. ———, *Probabilities on finite models*, The Journal of Symbolic Logic **41** (1976), 50–58.
166. S. Fajardo and H. J. Keisler, *Model theory of stochastic processes*, Lecture Notes in Logic, no. 14, A K Peters, 2002.
167. K. Falconer, *Fractal geometry: Mathematical foundations and applications*, 2nd ed., Wiley, 2003.
168. U. Felgner, *Pseudo-endliche gruppen*, Jahrbuch der Kurt-Gödel-Gesellschaft **3** (1990), 94–108.
169. E. Fokina, V. Harizanov, and D. Turetsky, *Computability-theoretic categoricity and Scott families*, preprint, 2019.
170. G. B. Folland, *Real analysis*, 2nd ed., Pure and Applied Mathematics, Wiley, 1999.
171. M. Foreman, D. J. Rudolph, and B. Weiss, *The conjugacy problem in ergodic theory*, Annals of Mathematics **173** (2011), 1529–1586.

172. M. Foreman and B. Weiss, *An anti-classification theorem for ergodic measure presrving transormations*, Journal of the European Mathematical Society **6** (2004), 277–292.
173. W. Fouché, *Arithmetical representations of Brownian motion I*, The Journal of Symbolic Logic **65** (2000), 421–442.
174. ———, *Martin-Löf randomness, invariant measures and countable homogeneous structures*, Theory of Computing Systems **52** (2013), 65–79.
175. W. L. Fouché, *Algorithmic randomness and Ramsey properties of countable homogeneous structures*, Logic, language, information, and computation, Lecture Notes in Computer Science, no. 7456, Springer, 2012, pp. 246–256.
176. R. Fraïssé, *Sur l'extension aux relations de quelques propriétés des ordres*, Annales scientifiques de l'É.N.S. **71** (1954), 363–388.
177. J. N. Y. Franklin, N. Greenberg, J. S. Miller, and K. M. Ng, *Martin-Löf random points satisfy Birkhoff's ergodic theorem for effectively closed sets*, Proceedings of the American Mathematical Society **140** (2012), 3623–3628.
178. J. N. Y. Franklin and C. P. Porter (eds.), *Algorithmic randomness*, Lecture Notes in Logic, no. 50, Cambridge University Press, 2020.
179. C. Freer, *Computable de Finetti measures*, Annals of Pure and Applied Logic **163** (2012), 530–546.
180. M. D. Fried and M. Jarden, *Field arithmetic*, 2nd ed., Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 11, Springer, 2005.
181. H. Friedman and L. Stanley, *A Borel reducibility theory for classes of countable structures*, Journal of Symbolic Logic **54** (1989), 894–914.
182. A. Frieze and R. Kannan, *Quick approximation to matrices and applications*, Combinatorica **19** (1999), 175–220.
183. A. Furman, *What is a stationary measure?*, Notices of the American Mathematical Soceity **58** (2011), 1276–1277.
184. H. Furstenberg, *A Poisson formula for semi-simple Lie groups*, Annals of Mathematics **77** (1963), 335–386.
185. ———, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation*, Mathematical Systems Theory **1** (1967), 1–49.
186. ———, *A note on Borel's density theorem*, Proceedings of the American Mathematical Society **55** (1976), 209–212.
187. ———, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, Journal D'Analyse Mathématique **31** (1977), 204–256.
188. S. Gaal and L. Gál, *The discrepancy of the sequence* $\{(2^n x)\}$, Indagationes Mathematicae **26** (1964), 129–143.
189. H. Gaifman, *Concerning measures in first order calculi*, Israel Journal of Mathematics **2** (1964), 1–18.
190. S. Gao, *Invariant descriptive set theory*, Pure and Applied Mathematics, CRC Press, 2009.
191. S. Gao and P. Gerdes, *Computably enumerable equivalence relations*, Studia Logica **67** (2001), 27–59.
192. W. I. Gasarch, *The P=?NP poll*, ACM SIGACT News **33** (2002), 34–47.
193. H. Geffner and J. Pearl, *A framework for reasoning with defaults*, Tech. Report R-94, Cognitive Systems Laboratory, UCLA, 1987.
194. D. Geiger and J. Pearl, *Logical and algorithmic properties of conditional independence and graphical models*, Annals of Statistics **21** (1993), 2001–2021.
195. T. Gelander, *Lecture notes on invariant random subgroups and lattices in rank one and higher rank*, preprint, 2015.
196. ———, *Kazhdan-Margulis theorem for invariant random subgroups*, Advances in Mathematics **327** (2018), 47–51.
197. G. Ghoshal, V. Zlatić, G. Caldarelli, and M. E. J. Newman, *Random hypergraphs and their applications*, Physical Review E **79** (2009), 066118–1–066118–10.
198. J. Gill, *Computational complexity of probabilistic Turing machines*, SIAM Journal on Computing **6** (1977), 675–695.
199. E. Glasner and B. Wiess, *Minimal actions of the group* $\mathbb{S}(\mathbb{Z})$ *of permutations of the integers*, Geometric and Functional Analysis **12** (2002), 964–988.
200. Yu. V. Glebskii, D. I. Kogan, M. I. Kogonkii, and V. A. Talanov, *Volume and fraction of satisfiability of formulas of the lower predicate calculus*, Kibernetika (Kiev) (1969), 17–27.

201. E. Mark Gold, *Language identification in the limit*, Information and Control **10** (1967), 447–474.
202. I. Goldbring and V. C. Lopes, *Pseudofinite and pseudocompact metric structures*, Notre Dame Journal of Formal Logic **56** (2015), 493–510.
203. I. Goldbring and H. Towsner, *An approximate logic for measures*, Israel Journal of Mathematics **199** (2014), 867–913.
204. O. Goldreich, *A primer on pseudorandom generators*, University Lecture Series, no. 55, American Mathematical Society, 2010.
205. O. Goldreich, S. Micali, and A. Wigderson, *Proofs that yield nothing but their validity or all languages in NP have zero-knowledge proof systems*, Journal of the Association for Computing Machinery **38** (1991), 691–729.
206. S. Goldwasser, S. Micali, and C. Rackoff, *The knowledge complexity of interactive proof systems*, STOC '85: Proceedings of the seventeenth annual ACM symposium on Theory of Computing, 1985, pp. 291–304.
207. D. Gorenstein, *Classifying the finite simple groups*, Bulletin of the American Mathematical Society **14** (1986), 1–98.
208. N. Greenberg, J. S. Miller, A. Shen, and L. Brown Westrick, *Dimension 1 sequences are close to randoms*, Theoretical Computer Science **705** (2018), 99–112.
209. M. Gromov, *Random walk in random groups*, Geometric and Functional Analysis **13** (2003), 73–146.
210. Y. Guivarc'h, *Sur la loi des grands nombres et le rayon spectral d'une marche aléatoire*, Astérisque **74** (1980), 47–98.
211. A. Günaydin and P. Hieronymi, *Dependent pairs*, Journal of Symbolic Logic **76** (2011), 377–390.
212. Y. Gurevich and P. H. Schmitt, *The theory of ordered Abelian groups does not have the independence property*, Transactions of the American Mathematical Society **284** (1984), 171–182.
213. I. Hacking, *The emergence of probability*, 2nd ed., Cambridge, 2006.
214. R. Haenni and N. Lehmann, *Probabilistic artumentation systems: a new perspecive on the Dempster-Shafer theory*, International Journal of Intelligent Systems **18** (2003), 93–106.
215. R. Haenni, J.-W. Romeijn, G. Wheeler, and J. Williamson, *Probabilistic logics and probabilistic networks*, Synthese Library, vol. 350, Springer, 2011.
216. T. Hailperin, *Sentential probability logic*, Lehigh University Press, 1996.
217. J. Y. Halpern, *A counterexample to theorems of Cox and Fine*, Journal of Artificial Intelligence Research **10** (1999), 67–85.
218. ———, *Cox's theorem revisited*, Journal of Artifical Intelligence Research **11** (1999), 429–435.
219. ———, *Reasoning about uncertainty*, MIT Press, 2003.
220. V. S. Harizanov, *Inductive inference systems for learning classes of algorithmically generated sets and structures*, Induction, Algorithmic Learning Theory, and Philosophy (M. Friend, N. B. Goethe, and V. S. Harizanov, eds.), Logic, Epistemology, and the Unity of Science, no. 9, Springer, 2007, pp. 27–54.
221. V. S. Harizanov and F. Stephan, *On the learnability of vector spaces*, Journal of Computer and System Sciences **73** (2007), 109–122.
222. L. A. Harrington, A. S. Kechris, and A. Louveau, *A Glimm-Effros dichotomy for Borel equivalence relations*, Journal of the American Mathematical Society **3** (1990), 903–928.
223. M. Harrison-Trainor, B. Khoussainov, and D. Turetsky, *Effective aspects of algorithmically random structures*, Computability **8** (2019), 359–375.
224. H. Hatami and S. Norine, *The entropy of random-free graphons and properties*, Combinatorics, Probability, and Computing **22** (2013), 517–526.
225. F. Hausdorff, *Dimension und äußeres Maß*, Mathematische Annalen **79** (1918), 157–179.
226. C. W. Henson, *A family of countable homogeneous graphs*, Pacific Journal of Mathematics **38** (1971), 69–83.
227. P. Hieronymi and T. Nell, *Distal and non-distal pairs*, Journal of Symbolic Logic **82** (2017), 375–383.
228. D. Hilbert, *Lectures on the foundations of geometry, 1891–1902*, vol. 1, Springer, 2004, translation due to V. Pambuccian, Fragments of Euclidean and hyperbolic geometry, *Scientia mathematicae Japonicae* 53 (2001) pp. 361–400.

229. J. Hintikka and G. Sandu, *Game-theoretical semantics*, Handbook of Logic and Language, Amsterdam, 1997, pp. 361–410.
230. D. R. Hirschfeldt, C. G. Jockusch Jr., T. H. McNicholl, and P. E. Schupp, *Asymptotic density and the coarse computability bound*, Computability **5** (2016), 13–27.
231. D. R. Hirschfeldt, B. Khoussainov, R. A. Shore, and A. M. Slinko, *Degree spectra and computable dimensions in algebraic structures*, Annals of Pure and Applied Logic **115** (2002), 71–113.
232. J. M. Hitchcock, *Correspondence principles for effective dimensions*, Theory of Computing Systems **38** (2005), 559–571.
233. J. M. Hitchcock and J. H. Lutz, *Why computational complexity requires stricter martingales*, Theory of Computing Systems **39** (2006), 277–296.
234. G. Hjorth, *Classification and orbit equivalence relations*, Mathematical Surveys and Monographs, vol. 75, American Mathematical Society, 2000.
235. ———, *On invariants for measure preserving transformations*, Fundamenta Mathematicae **169** (2001), 51–84.
236. ———, *A converse to Dye's Theorem*, Transactions of the American Mathematical Society **357** (2005), 3083–3103.
237. ———, *Glimm-Effros for coanalytic equivalence relations*, Journal of Symbolic Logic **74** (2009), 402–422.
238. G. Hjorth and A. S. Kechris, *Analytic equivalence relations and Ulm-type classifications*, Journal of Symbolic Logic **60** (1995), 1273–1300.
239. W. Hodges, *What is a structure theory?*, Bulletin of the London Mathematical Society **19** (1987), 209–237.
240. ———, *Groups in pseudofinite fields*, Model theory of groups and automorphism groups, London Mathematical Society Lecture Note Series, no. 244, Cambridge University Press, 1997, pp. 90–109.
241. B. Host and B. Kra, *Nilpotent structures in ergodic theory*, Mathematical Surveys and Monographs, vol. 236, American Mathematical Society, 2018.
242. E. Hrushovski, *Pseudo-finite fields and related structures*, Model Theory and Applications, Quaderni di Matematica, no. 11, Aracne, 2002, pp. 151–212.
243. ———, *Stable group theory and approximate subgroups*, Journal of the American Mathematical Society **25** (2012), 189–243.
244. E. Hrushovski, Y. Peterzil, and A. Pillay, *Groups, measures, and the NIP*, Journal of the American Mathematical Society **21** (2008), 563–596.
245. E. Hrushovski and A. Pillay, *Groups definable in local fields and pseudo-finite fields*, Israel Journal of Mathematics **85** (1994), 203–262.
246. ———, *Definable subgroups of algebraic groups over finite fields*, Journal für die reine und angewandte Mathematik **462** (1995), 69–91.
247. T. W. Hungerford, *Algebra*, Graduate Texts in Mathematics, no. 73, Springer, 1974.
248. N. Immerman, *Descriptive complexity*, Graduate Texts in Computer Science, Springer, 1999.
249. S. Jain, D. Osherson, J. Royer, and A. Sharma, *Systems that learn: An introduction to learning theory*, 2nd ed., Learning, Development, and Conceptual Change, MIT Press, 1999.
250. S. Janson, *Graphons, cut norm, and distance, couplings and rearrangements*, NYJM Monographs, no. 4, New York Journal of Mathematics, 2013.
251. R. C. Jeffrey, *The logic of decision*, 2nd ed., University of Chicago Press, 1983.
252. C. G. Jockusch Jr and P. E. Schupp, *Generic computability, turing degrees, and asymptotic density*, Journalof the London Mathematical Society, 2nd Series **85** (2012), 472–490.
253. V. Kaimanovich, I. Kapovich, and P. Schupp, *The subadditive ergodic theorem and generic stretching factors for free group automorphisms*, Israel Journal of Mathematics **157** (2007), 1–46.
254. ———, *The subadditive ergodic theorem and generic stretching factors for free group automorphisms*, Israel Journal of Mathematics **157** (2007), 1–46.
255. O. Kallenberg, *Foundations of modern probability*, Probability and its Applications, Springer, 1997.
256. ———, *Probabilistic symmetries and invariance principles*, Probability and its Applications, Springer, 2005.
257. E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, and A. J. Hudspeth, *Principles of neural science*, fifth ed., McGraw-Hill, 2013.

258. I. Kaplansky, *Infinite Abelian groups*, revised ed., University of Michigan Press, 1969.
259. I. Kapovich, A. Myasnikov, P. Schupp, and V. Shpilrain, *Generic-case complexity, decision problems in group theory, and random walks*, Journal of Algebra **264** (2003), 665–694.
260. M. Karpinski and A. Macintyre, *Approximating volumes and integrals in o-minimal and p-minimal theories*, Connections between model theory and algeraic and analytic geometry, Quaderni di Matematica, no. 6, Arcane, 2000, pp. 149–177.
261. S. Kauffman, *Homeostasis and differentiation in random genetic control networks*, Nature **224** (1969), 177–178.
262. ———, *Metabolic stability and epigenesis in randomly constructed genetic nets*, Journal of Theoretical Biology **22** (1969), 437–467.
263. ———, *The large scale structure and dynamics of gene control circuits: an ensemble approach*, Journal of Theoretical Biology **44** (1974), 167–190.
264. ———, *The origins of order*, Oxford University Press, 1993.
265. S. M. Kautz, *Degrees of random sets*, Ph.D. thesis, Cornell University, 1991.
266. D. Kazhdan and G. Margulis, *A proof of Selberg's hypothesis*, Matematicheskii Sbornik **75** (1968), 163–168.
267. M. J. Kearns and U. V. Vazirani, *An introduction to computational learning theory*, MIT Press, 1994.
268. A. S. Kechris and B. D. Miller, *Topics in orbit equivalence*, Lecture Notes in Mathematics, no. 1852, Springer, 2004.
269. A. S. Kechris, V. G. Pestov, and S. Todorcevic, *Fraïssé limits, Ramsey theory, and topological dynamics of automorphism groups*, Geometric and Functional Analysis **15** (2005), 106–189.
270. ———, *Fraïssé limits, Ramsey theory, and topological dynamics of automorphism groups*, Geometric and Functional Analysis **15** (2005), 106–189.
271. A. S. Kechris and R. D. Tucker-Drob, *The complexity of classification results in ergodic theory*, Appalachian Set Theory 2006–2012, London Mathematical Society Lecture Note Series, no. 406, Cambridge, 2013, pp. 265–299.
272. H. J. Keisler, *Randomizing a model*, Advances in Mathematics **143** (1999), 124–158.
273. ———, *Probability quantifiers*, Model-Theoretic Logics, Perspectives in Logic, Cambridge University Press, 2016, Originally published 1985, pp. 509–556.
274. J. M. Keynes, *A treatise on probability*, MacMillan, 1921.
275. A. I. Khinchine, *Mathematical foundations of information theory*, Dover, 1957.
276. B. Khoussainov, *A quest for algorithmically random infinite structures*, Proceedings of the Joint Meeting of the Twenty-Third EACSL Annual Conference on Computer Science Logic (CSL) and the Twenty-Ninth Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), ACM, 2014, Article No. 56.
277. ———, *A quest for algorithmically random infinite structures, II*, Logical foundations of computer science, Lecture Notes in Computer Science, no. 9537, Springer, 2016, pp. 159–173.
278. H. Ki and T. Linton, *Normal numbers and subsets of* $\mathbb{N}$ *with given densities*, Fundamenta Mathematicae **144** (1994), 163–179.
279. S. Kiefer, R. Mayr, M. Shirmohammadi, and D. Wojtczak, *Strong determinacy of countable stochastic games*, Proceedings of the 32nd Annual ACM/IEEE Symposium on Logic in Computer Science, 2017.
280. J. H. Kim, O. Pikhurko, J. Spencer, and O. Verbitsky, *How complex are random graphs in first order logic?*, Random Structures & Algorithms **26** (2005), 119–145.
281. J. F. C. Kingman, *The ergodic theory of subadditive stochastic processes*, Journal of the Royal Statistical Soceity, Series B (Methodological **30** (1968), 499–510.
282. V. Klee and G. J. Minty, *How good is the simplex algorithm?*, Inequalities, III, Academic Press, 1972, pp. 159–175.
283. J. F. Knight, A. Pillay, and C. Steinhorn, *Definable sets in ordered structures II*, Transactions of the American Mathematical Society **295** (1986), 593–605.
284. D. E. Knuth, *The art of computer programming: Seminumerical algorithms*, 3rd ed., vol. 2, Pearson, 1998.
285. P. G. Kolaitis and M. Y. Vardi, *0–1 laws and decision problems for fragments of second-order logic*, Information and Computation **87** (1990), 302–338.
286. ———, *Infinitary logics and 0–1 laws*, Information and Computation **98** (1992), 258–294.
287. A. N. Kolmogorov, *Foundations of the theory of probability*, Chelsea, 1950.

288. А. Н. Колмогоров, Три подхода к определению понятия ⟨⟨ количество информатсии ⟩⟩, Проблемы передачи информации **1** (1965), 3–11.
289. B. Kra, *Commentary on 'Ergodic theory of amenable group actions': old and new*, Bulletin of the American Mathematical Society **55** (2018), 343–345.
290. P. N. Kryloff and N. Bogoliouboff, *La théorie générale de la mesure dans son application à l'étude des systémes dynamiques de la mécanique non linéaire*, Annals of Mathematics **38** (1937), 65–113.
291. M. H. Kutner, C. J. Nachtsheim, J. Neter, and W. Li, *Applied linear statistical models*, fifth ed., McGraw-Hill Irwin Series: Operations and Decision Sciences, McGraw-Hill, 2005.
292. A. Kučera and T. Slaman, *Randomnuess and recursive enumerability*, SIAM Journal on Computing **31** (2001), 199–211.
293. H. Lädesmäki, S. Hautaniemi, I. Shmulevich, and O. Yli-Harja, *Relationships between probabilistic Boolean networks and dynamic Bayesian networks as models of gene regulatory networks*, Signal Processing **86** (2006), 814–834.
294. H. Lädesmäki, I. Shmulevich, and O. Yli-Harja, *On learning gene regulatory networks under the Boolean network model*, Machine Learning **52** (2003), 147–167.
295. J. C. Lagarias, *The ultimate challenge: The* $3x+1$ *problem*, American Mathematical Society, 2010.
296. S. Lang and A. Weil, *Number of points of varieties in finite fields*, American Journal of Mthematics **76** (1954), 819–827.
297. M. C. Laskowski, *A simpler axiomatization of the Shelah-Spencer almost sure theories*, Israel Journal of Mathematics **161** (2007), 157–186.
298. S. L. Lauritzen and N. Wermuth, *Graphical models for associations between variables, some of which are qualitative and some quantitative*, The Annals of Statistics **17** (1989), 31–57.
299. Y. LeCun, Y. Bengio, and G. Hinton, *Deep learning*, Nature **521** (2015), 436–444.
300. M. Ledoux, *The concentration of measure*, Mathematical Surveys and Monographs, no. 89, American Mathematical Society, 2001.
301. Л. А. Левин, Законы сохранения (невозрастания) информации и вопросы обоснования теории вероятностей, Проблемы передачи информации **10** (1974), 30–35.
302. M. B. Levin, *On the discrepancy estimate of normal numbers*, Acta Arithmetica **88** (1999), 99–111.
303. M. Li, J. Tromp, and P. Vitányi, *Sharpening Occam's razor*, Information Processing Letters **85** (2003), 267–274.
304. Ming Li and Paul Vitányi, *An introduction to Kolmogorov complexity and its applications*, third ed., Texts in Computer Science, Springer, 2008.
305. N. Littlestone, *Learning quickly when irrelevant attributes abound: a new linear-threshold algorithm*, Machine Learning **2** (1988), 285–318.
306. L. Lovász, *Large networks and graph limits*, American Mathematical Society Colloquium Publications, vol. 60, American Mathematical Society, 2012.
307. L. Lovász and B. Szegedy, *Limits of dense graph sequences*, Journal of Combinatorial Theory, Series B **96** (2006), 933–957.
308. D. Loveland, *A new interpretation of the von Mises' concept of random sequence*, Zeitschrift fur mathematische Logik und Grundlagen dr Mathematik **12** (1966), 279–294.
309. A. Lubotzky and B. Weiss, *Groups and expanders*, Expanding Graphs, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, vol. 10, American Mathematical Society, 1993, pp. 95–109.
310. T. Luczak and J. Spencer, *When does the zero-one law hold?*, Journal of the American Mathematical Society **4** (1991), 451–468.
311. C. Lund, L. Fortnow, H. Karloff, and N. Nisan, *Algebraic methods for interactive proof systems*, Journal of the Association for Computing Machinery **39** (1992), 859–868.
312. J. H. Lutz, *Dimension in complexity classes*, SIAM Journal on Computing **32** (2003), 1236–1259.
313. R. Lyons and Y. Peres, *Probability on trees and networks*, Cambridge Series in Statistical and Probabilistic Mathematics, no. 42, Cambridge, 2016.
314. D. Macpherson and C. Steinhorn, *One-dimensional asymptotic classes of finite structures*, Transactions of the American Mathematical Society **380** (2008), 411–448.
315. ———, *Definability in classes of finite structures*, Finite and algorithmic model theory, London Mathematical Society Lecture Note Series, no. 379, Cambridge, 2011, pp. 140–176.

316. D. Macpherson and K. Tent, *Pseudofinite groups with NIP theory and definability in finite simple groups*, Groups and Model Theory, Contemporary Mathematics, no. 576, American Mathematical Society, 2012, pp. 255–267.
317. M. Malliaris and A. Pillay, *The stable regularity lemma revisited*, Proceedings of the American Mathematical Society **144** (2016), 1761–1765.
318. M. Malliaris and S. Shelah, *Regularity lemmas for stable graphs*, Transactions of the American Mathematical Society **366** (2014), 1551–1585.
319. ———, *Notes on the stable regularity lemma*, Bulletin of Symbolic Logic **27** (2021), 415–425.
320. V. W. Marek and M. Truszczyński, *Nonmonotonic logic*, Artificial Intelligence, Springer, 1993.
321. G. A. Margulis, *Discrete subgroups of semisimple Lie groups*, Ergebnisse der Mathematik un ihrer Grenzgebiete, 3. Folge, no. 17, Springer, 1991.
322. D. Marker, *Model theory*, Graduate Texts in Mathematics, no. 217, Springer, 2002.
323. A. Marks and S. Unger, *Baire measurable paradoxical decompositions via matchings*, Advances in Mathematics **289** (2016), 397–410.
324. R. Marshall, *Robust classes of finite structures*, Ph.D. thesis, University of Leeds, 2008.
325. P. Martin-Löf, *The definition of random sequences*, Information and Control **9** (1966), 602–619.
326. K. R. Matthews, *Generalized* $3x + 1$ *mappings: Markov cxhains and ergodic theory*, The Ultimate Challenge: The $3x + 1$ Problem (J. C. Lagarias, ed.), American Mathematical Society, 2010, pp. 79–.
327. K. R. Matthews and A. M. Watts, *A generalization of Hasse's generalization of the Syracuse algorithm*, Acta Arithmetica **43** (1984), 167–175.
328. P. Mattila, *Geometry of sets and measures in Euclidean spaces*, Cambridge Studies in Advanced Mathematics, no. 44, Cambridge University Press, 1995.
329. A. D. Matushkin, *Zero-one law for random uniform hypergraphs*, preprint, 2016.
330. A. D. Matushkin and M. E. Zhukovskii, *First order sentences about random graphs: small number of alternations*, Discrete Applied Mathematics **236** (2018), 329–346.
331. J. E. Maxfield, *Normal k-tuples*, Pacific Journal of Mathematics **3** (1953), 189–196.
332. E. Mayordomo, *A Kolmogorov complexity characterization of constructive Hausdorff dimension*, Information Processing Letters **84** (2002), 1–3.
333. W. Merkle, N. Mihalović, and T. A. Slaman, *Some results on effective randomness*, Theory of Computing Systems **39** (2006), 707–721.
334. J.-F. Mertens, *Stochastic games*, Handbook of Game Theory, vol. 3, Elsevier, 2002, pp. 1809–1832.
335. J.-F. Mertens and A. Neyman, *Stochastic games*, International Journal of Game Theory **10** (1981), 53–66.
336. P. Michel, *Busy beaver competition and Collatz-like problems*, Archive for Mathematical Logic **32** (1993), 351–367.
337. P. Michel and M. Margenstern, *Generalized* $3x+1$ *functions and the theory of computation*, The Ultimate Challenge: The $3x+1$ Problem (J. C. Lagarias, ed.), American Mathematical Society, 2010, pp. 105–.
338. C. F. Miller, III, *On group-theoretic decision problems and their classification*, Annals of Mathematics Studies, no. 68, Princeton University Press, 1971.
339. G. A. Miller and D. McNeill, *Psycholinguistics*, The Handbook of Social Psychology (G. Lindzey and E. Aronson, eds.), vol. 3, Addison-Wesley, 2nd ed., 1968, pp. 666–794.
340. G. L. Miller, *Riemann's hypothesis and tests for primality*, Journal of computer and system science **13** (1976), 300–317.
341. R. Miller, *Isomorphism and classification for countable structures*, Computability **8** (2019), 99–117.
342. J. Milnor, *Dynamics in one complex variable*, 3rd ed., Annals of Mathematics Studies, no. 160, Princeton University Press, 2006.
343. A. Montalbán, *Computable structure theory: Within the arithmetic*, preprint, 2019.
344. *****A. Montalban, *Computable structure theory, part ii*, not yet seen, 2021?
345. A. Montalbán and A. Nies, *Borel structures: a brief survey*, Effective Mathematics of the Uncountable, Lecture Notes in Logic, no. 41, Cambridge, 2013, pp. 124–134.
346. C. C. Moore, *Ergodicity of flows on homogeneous spaces*, American Journal of Mathematics **88** (1966), 154–178.

347. A. De Morgan, *Formal logic: or, the calculus of inference, necessary and probable*, Taylor and Walton, 1847.
348. R. A. Moser and G. Tardos, *A constructive proof of the general Lovász local lemma*, Journal of the Association for Computing Machinery **57** (2010), 1–15.
349. A. A. Muchnik, A. L. Semenov, and V. A. Uspensky, *Mathematical metaphysics of randomness*, Theoretical Computer Science **207** (1998), 263–317.
350. K. P. Murphy, *Dynamic Bayesian networks*, Ph.D. thesis, University of California, Berkeley, 2002.
351. T. Neary, *Small polynomial time universal Turing machines*, Proceedings of the 4th Irish Conference on the Mathematical Foundations of Computer Science and Information Technology (T. Hurley, et al., ed.), 2006, pp. 325–329.
352. A. Nies, *Computability and randomness*, Oxford Logic Guides, no. 51, Oxford, 2009.
353. N. Nisan and A. Wigderson, *Hardness vs. randomness*, Journal of Computer Systems and Sciences **49** (1994), 149–167.
354. P. S. Novikov, *On algorithmic unsolvability of the word problem in group theory*, Trudy Matimaticheskogo Instituta imeni V.A. Steklova, no. 44, Izdat. Akad. Nauk SSSR, 1955.
355. D. S. Ornstein and B. Weiss, *Ergodic theory of amenable group actions. I: The Rohlin Lemma*, Bulletin of the American Mathematical Society **2** (1980), 161–164.
356. D. N. Osherson and S. Weinstein, *Criteria of language learning*, Information and Control **52** (1982), 123–138.
357. D. Osin, *A topological zero-one law and elementary equivalence of finitely generated groups*, Annals of Pure and Applied Logic **172** (2021), 102915.
358. R. Pal, I. Ivanov, A. Datta, M. L. Bittner, and E. R. Dougherty, *Generating Boolean networks with a prescribed attractor structure*, Bioinformatics **21** (2005), 4021–4025.
359. J. Paris and A. Venkovská, *Pure inductive logic*, Perspectives in Logic, Cambridge, 2015.
360. J. Pearl, *Deciding consistency in inheritance networks*, Tech. Report 870053, Cognitive Systems Laboratory, UCLA, 1987.
361. ———, *Probabilistic reasoning in intelligent systems*, Morgan Kaufmann, 1988.
362. J. Pearl and A. Paz, *Graphoids: A graph-based logic for reasoning about relevance relations*, Tech. Report 850038, Cognitive Systems Laboratory, UCLA, 1985.
363. Y. B. Pesin, *Dimension theory in dynamical systems*, University of Chicago Press, 1997.
364. P. Petersen, *Riemannian geometry*, 3rd ed., Graduate Texts in Mathematics, no. 171, Springer, 2016.
365. F. Petrov and A. Vershik, *Uncountable graphs and invariant measures on the set of universal countable graphs*, Random Structures and Algorithms **37** (2010), 389–406.
366. R. R. Phelps, *Lectures on choquet's theorem*, Lecture Notes in Mathematics, no. 1757, Springer, 2001.
367. A. Pillay and C. Steinhorn, *Discrete o-minimal structures*, Annals of Pure and Applied Logic **34** (1987), 275–289.
368. M. S. Pinsker, *On the complexity of a concentrator*, 7th International Teletraffic Conference, 1973.
369. K. Popper, *The logic of scientific discovery*, Routledge, 2002, Original edition published in 1935; this edition dates from text of 1959.
370. E. L. Post, *Recursive unsolvability of a problem of Thue*, The Journal of Symbolic Logic **12** (1947), 1–11.
371. P. Potgieter, *Algorithmically random series and Brownian motion*, Annals of Pure and Applied Logic **169** (2018), 1210–1226.
372. M. O. Rabin, *Probabilistic algorithm for testing primality*, Journal of Number Theory **12** (1980), 128–138.
373. C. Radin and L. Sadun, *Phase transitions in a complex network*, Journal of Physics A **46** (2013), 305002.
374. F. P. Ramsey, *Truth and probability*, The Foundations of Mathematics and other Logical Essays (R. B. Braithwaite, ed.), Harcourt Brace, 1931, First published 1926, pp. 156–198.
375. J. Reimann, *Computability and fractal dimension*, Ph.D. thesis, Ruprecht-Karls-Universität Heidelberg, 2004.
376. R. Reiter, *A logic for default reasoning*, Artificial Intelligence **13** (1980), 81–132.
377. ———, *Nonmonotonic reasoning*, Annual Review of Computer Science **1987** (1987), 147–186.

378. L. J. Richter, *Degrees of structures*, The Journal of Symbolic Logic **46** (1981), 723–731.
379. A. Rivkind and O. Barak, *Local dynamics in trained recurrent neural networks*, Physical Review Letters **118** (2017), no. 258101, 1–5.
380. A. Robinson, *Complete theories*, Studies in logic and the foundations of mathematics, North-Holland, 1956.
381. F. Rosenblatt, *Principles of neurodynamics; perceptrons and the theory of brain mechansims*, Spartan, 1962.
382. S. Ross, *Introduction to probability models*, 12th ed., Academic Press, 2019.
383. J. J. Rotman, *The theory of groups*, Allyn and Bacon, 1965.
384. A. Rumyantsev, *Infinite computable version of Lovász local lemma*, preprint, 2010.
385. A. Rumyantsev and A. Shen, *Probabilistic constructions of computable objects and a computable version of Lovász local lemma*, Fundamental Informaticae **132** (2014), 1–14.
386. S. J. Russell and P. Norvig, *Artificial intelligence: A modern approach*, third ed., Prentice Hall Series in Artificial Intelligence, Prentice Hall, 2010.
387. J. Rute, *Algorithmic randomness and constructive/computable measure theory*, Algorithmic Randomness: Progress and Prospects (J. N. Y. Franklin and C. P. Porter, eds.), Lecture Notes in Logic, no. 50, Cambridge, 2020, pp. 58–114.
388. M. J. Ryten, *Model theory of finite difference fields and simple groups*, Ph.D. thesis, The University of Leeds, 2007.
389. N. Sauer, *On the density of families of sets*, Journal of Combinatorial Theory (A) **13** (1972), 145–147.
390. A. Saxe, Y. Bansal, J. Dapello, M. Advani, A. Kolchinsky, B. Tracey, and D. Cox, *On the information bottleneck theory of deep learning*, ICLR, 2018.
391. A.-M. Scheerer, *Computable absolutely normal numbers and discrepancies*, Mathematics of Computation (2017), 2911–2926.
392. J. Schmidhuber, *Deep learning in neural networks: An overview*, Neural Networks **61** (2015), 85–117.
393. K. Schmidt, *Asymptotically invariant sequences and an action of $SL(2, Z)$ on the 2-sphere*, Israel Journal of Mathematics **37** (1980), 193–208.
394. W. M. Schmidt, *Über die Normalität von Zahlen zu verschiedenen Basen*, Acta Arithmetica **VII** (1962), 299–309.
395. ———, *Irregularities of distribution, VII*, Acta Arithmetica **21** (1972), 45–50.
396. C. P. Schnorr, *A unified approach to the definition of random sequences*, Mathematical systems theory **5** (1971), 246–258.
397. ———, *Zufälligkeit und Wahrscheinlichkeit*, Lecture Notes in Mathematics, no. 218, Springer, 1971.
398. G. Shafer, *A mathematical theory of evidence*, Princeton University Press, 1976.
399. A. Shamir, **IP = PSPACE**, Journal of the Association for Computing Machinery **39** (1992), 869–877.
400. C. E. Shannon, *A mathematical theory of communication*, The Bell System Technical Journal **27** (1948), 379–423.
401. L. S. Shapley, *Stochastic games*, Proceedings of the National Academy of Sciences **39** (1953), 1095–1100.
402. S. Shelah, *A combinatorial problem; stability and order for models and theories in infinitary languages*, Pacific Journal of Mathematics **41** (1972), 247–261.
403. ———, *Classification theory and the number of non-isomorphic models*, revised ed., Studies in Logic and the Foundations of Mathematics, no. 92, North-Holland, 1990.
404. S. Shelah and J. Spencer, *Zero-one laws for sparse random graphs*, Journal of the American Mathematical Society **1** (1988), 97–115.
405. A. Shen, *IP = PSPACE: Simplified proof*, Journal of the Association for Computing Machinery **39** (1992), 878–880.
406. A. K. Shen, *On relations between different algorithmic definitions of randomness*, Sovient Mathematics Doklady **38** (1989), 316–319.
407. A. N. Shiryaev, *Probability*, second ed., Graduate Texts in Mathematics, no. 95, Springer, 1996.
408. I. Shmulevich and E. R. Dougherty, *Genomic signal processing*, Princeton Series in Applied Mathematics, Princeton University Press, 2007.

409. _____, *Probabilisitc boolean networks*, Society for Industrial and Applied Mathematics, 2010.
410. H. A. Simon, *On a class of skew distribution functions*, Biometrika **42** (1955), 425–440.
411. P. Simon, *A guide to NIP theories*, Lecture Notes in Logic, no. 44, Cambridge, 2015.
412. R. I. Soare, *Recursively enumerable sets and degrees*, Springer-Verlag, 1987.
413. J. Spencer, *Asymptotic lower bounds for Ramsey functions*, Discrete Mathematics **20** (1977), 69–76.
414. _____, *Threshold functions for extension statements*, Journal of Combinatorial Theory A **53** (1990), 286–305.
415. _____, *Threshold spectra via the Ehrenfeucht game*, Discrete Applied Mathematics **30** (1991), 235–252.
416. _____, *The strange logic of random graphs*, Algorithms and Combinatorics, no. 22, Springer, 2001.
417. J. Spencer and K. St. John, *The tenacity of zero-one laws*, The Electronic Journal of Combinatorics **8** (2001), R17.1–R17.14.
418. J. Spencer and M. E. Zhukovskii, *Bounded quantifier depth spectra for random graphs*, Discrete Mathematics **339** (2016), 1651–1664.
419. L. Staiger, *Kolmogorov complexity and Hausdorff dimension*, Information and Computation **103** (1993), 159–194.
420. _____, *A tight upper bound on Kolmogorov complexity and uniformly optimal prediction*, Theory of Computing Systems **31** (1998), 215–229.
421. C. I. Steinhorn, *Borel structures and measure and category logics*, Model Theoretic Logics, Perspectives in Logic, no. 8, Springer, 1985, pp. 579–596.
422. G. Stengle and J. E. Yukich, *Some new Vapnik-Chervonenkis classes*, The Annals of Statistics **17** (1989), 1441–1446.
423. V. E. Stepanov, *Phase transitions in random graphs*, Theory of Probability and its Applications **15** (1970), 187–203.
424. F. Stephan and Yu. Ventsov, *Learning algebraic structures from text*, Theoretical Computer Science **268** (2001), 221–273.
425. L. J. Stockmeyer, *The polynomial-time hierarchy*, Theoretical Computer Science **3** (1976), 1–22.
426. G. Stuck and R. Zimmer, *Stabilizers for ergodic actions of higher rank semisimple groups*, Annals of Mathematics **139** (1994), 723–747.
427. M. Studený, *Conditional independence relations have no finite complete characterization*, preprint, 1992.
428. D. Sussillo and O. Barak, *Opening the black box: Low-dimensional dynamics in high-dimensional recurrent neural networks*, Neural Computation **25** (2013), 626–649.
429. E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arithmetica **27** (1975), 199–245.
430. _____, *Regular partitions of graphs*, preprint, 9 pp., 1975.
431. T. Tao, *Expanding polynomials over finite fields of large characteristic, and a regularity lemma for definable sets*, Contributions to Discrete Mathematics **10** (2015), 22–98.
432. _____, *Szemerédi's proof of Szemerédi's theorem*, Acta Mathematica Hungarica **161** (2020), 443–487.
433. A. Tarski, *Algebraische Fassung des Maßproblems*, Fundamenta Mathematicae **31** (1938), 47–66.
434. _____, *Cardinal algebras*, Oxford University Press, 1949.
435. A. Taveneaux, *Randomness zoo*, preprint, 2012.
436. E. Thoma, *Die unzerlegbaren, positiv-definiten Klassenfunktionen der abzählbar unendlichen, symmetrischen Gruppe*, Mathematische Zeitschfrift **85** (1964), 40–61.
437. _____, *Über unitäre Darstellungen abzählbarer, diskreter Gruppen*, Mathematische Annalen **153** (1964), 111–138.
438. S. Thomas, *The classification problem for torsion-free Abelian groups of finite rank*, Journal of the American Mathematical Society **16** (2003), 233–258.
439. S. Thomas and R. Tucker-Drob, *Invariant random subgroups of strictly diagonal limits of finite symmetric groups*, Bulletin of the London Mathematical Society **46** (2014), 1007–1020.
440. _____, *Invariant random subgroups of inductive limits of finite alternating groups*, Journal of Algebra **503** (2018), 474–533.

441. A. Thue, *Probleme über Veränderungen von Zeichenreihen nach gegeben regeln*, Skrifter utgit av Videnskapsselskapet i Kristiania **1** (1914), no. 10, 1–34.
442. N. Tishby and N. Zaslavsky, *Deep learning and the information bottleneck principle*, IEEE Information Theory Workshop, 2015.
443. S. Toda, *PP is as hard as the polynomial-time hierarchy*, SIAM Journal of Computing **20** (1991), 878–880.
444. H. Towsner, *σ-algebras for quasirandom hypergraphs*, Random Structures and Algorithms **50** (2017), 114–139.
445. ———, *Algorithmic randomness in ergodic theory*, Algorithmic Randomness: Progress and Prospects (J. N. Y. Franklin and C. P. Porter, eds.), Lecture Notes in Logic, no. 50, Cambridge, 2020, pp. 40–57.
446. S. Vadhan, *Pseudorandomness*, Foundations and Trends in Theoretical Computer Science, vol. 7, Now, 2012.
447. L. G. Valiant, *The complexity of computing the permanent*, Theoretical Computer Science **8** (1979), 189–201.
448. ———, *A theory of the learnable*, Communications of the ACM **27** (1984), 1134–1142.
449. P. van der Hoorn, G. Lippner, and D. Krioukov, *Sparse maximum-entropy random graphs with a given power-law degree distribution*, Journal of Statistical Physics **173** (2018), 803–844.
450. V. N. Vapnik and A. Ya. Chervonenkis, *On the uniform convergence of relative frequencies of events to their probabilities*, Theory of probability and its applications **16** (1971), 264–280.
451. A. M. Vershik, *Totally nonfree actions and the infinite symmetric group*, Moscow Mathematical Journal **12** (2012), 193–212.
452. N. Vieille, *Stochastic games: Recent results*, Handbook of Game Theory, vol. 3, Elsevier, 2002, pp. 1833–1850.
453. Susan Vineberg, *Dutch Book Arguments*, The Stanford Encyclopedia of Philosophy (Edward N. Zalta, ed.), Metaphysics Research Lab, Stanford University, spring 2016 ed., 2016.
454. A. Visser, *Numerations, λ-calculus, & arithmetic*, To H.B. Curry: Essays on Combinatory Logic, Lambda Calculus, and Formalism, Academic Press, 1980, pp. 259–284.
455. J. von Neumann, *Zur allgemeinen Theorie des Masses*, Fundamenta Mathematicae **13** (1929), 73–116.
456. P. Walters, *An introduction to ergodic theory*, Graduate Texts in Mathematics, no. 79, Springer, 1982.
457. R. Weber, *Computability theory*, Student Mathematical Library, no. 62, American Mathematical Society, 2012.
458. K. Weihrauch, *Computable analysis*, Texts in Theoretical Computer Science, Springer, 2000.
459. J. Williamson, *Probability logic*, Handbook of the Logic of Argument and Inference, Studies in Logic and Practical Reasoning, vol. 1, Elsevier, 2002, pp. 397–424.
460. J. S. Wilson, *On simple pseudofinite groups*, Journal of the London Mathematical Society **51** (1995), 471–490.
461. D. Xiao, *New perspectives on the complexity of computational learning, and other problems in theoretical computer science*, Ph.D. thesis, Princeton University, 2009.
462. ———, *On basing $\mathbf{ZK} \neq \mathbf{BPP}$ on the hardness of PAC learning*, 24th Annual IEEE Conference on Computational Complexity, 2009, pp. 304–315.
463. A. C. Yao, *Theory and applications of trapdoor functions*, 23rd Annual Symposium on Foundations of Computer Science, IEEE, 1982, pp. 80–91.
464. C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals, *Understanding deep learning requires rethinking generalization*, ICLR, 2017.
465. X. Zheng and R. Rettinger, *On the extnesions of Solovay reducibility*, Computing and Combinatorics: 10th Annual International Conference, COCOON 2004, Lecture Notes in Computer Science, no. 3106, Springer, 2004, pp. 360–369.
466. M. E. Zhukovskii, *On infinite spectra of first order properties of random graphs*, Moscow Journal of Combinatorics and Number Theory **4** (2016), 73–102.
467. R. J. Zimmer, *Ergodic theory and semisimple groups*, Monographs in Mathematics, no. 81, Birkhäuser, 1984.