

Intersecting sets in probability spaces and Shelah's classification

Artem Chernikov^{1*} and Henry Towsner^{2†}

¹ University of Maryland, College Park, MD, US and University of California, Los Angeles, CA, US
 artem@umd.edu

² University of Pennsylvania, Philadelphia, PA, US
 htowsner@math.upenn.edu

Abstract

For $n \in \mathbb{N}$ and $\varepsilon > 0$, given a sufficiently long sequence of events in a probability space all of measure at least ε , some n of them will have a common intersection. A more subtle pattern: for any $0 < p < q < 1$, we cannot find events A_i and B_i so that $\mu(A_i \cap B_j) \leq p$ and $\mu(A_j \cap B_i) \geq q$ for all $1 < i < j < n$, assuming n is sufficiently large. This is closely connected to model-theoretic stability of probability algebras. We survey some results from our recent work in [7] on more complicated patterns that arise when our events are indexed by multiple indices. In particular, how such results are connected to higher arity generalizations of de Finetti's theorem in probability, structural Ramsey theory, hypergraph regularity in combinatorics, and model theory.

1 Intersections in a sequence of sets of positive measure

The following is a basic fact on intersections of sets of positive measure in probability spaces (there exist more precise infinitary/density versions, e.g. Bergelson's lemma in dynamics [5]):

Fact 1. *For every $\varepsilon \in \mathbb{R}_{>0}$ and $n \in \mathbb{N}$ there exists $N \in \mathbb{N}$ satisfying the following. If (X, \mathcal{B}, μ) is a probability space (i.e. \mathcal{B} is a σ -algebra of subsets of X and μ is a countably additive probability measure on \mathcal{B}) and $A_i \in \mathcal{B}$ are measurable sets with $\mu(A_i) \geq \varepsilon$ for $1 \leq i \leq N$, then $\mu(\bigcap_{i \in I} A_i) > 0$ for some $I \subseteq [N] = \{1, \dots, N\}$ with $|I| = n$.*

We sketch an overcomplicated proof of this fact in the remainder of Section 1, as a warm up for what comes later. If the random variables $\mathbf{1}_{A_i} : X \rightarrow \{0, 1\}$ in Fact 1 were independent, then of course $\mu(\bigcap_{i \in [n]} A_i) = \prod_{i \in [n]} \mu(A_i) \geq \varepsilon^n > 0$. We will reduce to this case. Assume from now on that for some fixed $\varepsilon > 0$ and n , no $N \in \mathbb{N}$ satisfies the claim.

1.1 Homogenizing the sequence

Using Ramsey's theorem, we can homogenize our sequence arbitrarily well. E.g., we could assume that for any fixed $\delta > 0$ and k , $\mu(A_{i_1} \cap \dots \cap A_{i_k}) \approx^\delta \mu(A_{j_1} \cap \dots \cap A_{j_k})$ for any $i_1 < \dots < i_k$, $j_1 < \dots < j_k$, and similarly for arbitrary Boolean combinations of the A_i 's.

Using a compactness argument (e.g. taking Loeb measure on an ultraproduct of counterexamples), we can thus find some large probability space (X, \mathcal{B}, μ) and sets $A_i \in \mathcal{B}$, $\mu(A_i) \geq \varepsilon$ for $i \in \mathbb{N}$, still intersection of any n of them has measure 0, so that the sequence of random variables $(\mathbf{1}_{A_i} : i \in \mathbb{N})$ is *spreadable*.

*Chernikov was partially supported by the NSF CAREER grant DMS-1651321 and by the NSF Research Grant DMS-2246598.

†Towsner was partially supported by the NSF grant DMS-2054379.

1.2 de Finetti's theorem

Definition 1. A sequence of $[0, 1]$ -valued random variables $(\xi_i : i \in \mathbb{N})$ is spreadable if for every $n \in \mathbb{N}$ and $i_1 < \dots < i_n, j_1 < \dots < j_n$ we have $(\xi_{i_1}, \dots, \xi_{i_n}) =^{\text{dist}} (\xi_{j_1}, \dots, \xi_{j_n})$.

For example, every i.i.d. (independent, identically distributed) sequence of random variables is spreadable. The converse holds “up to mixing”:

Fact 2 (de Finetti's theorem). *If an infinite sequence of random variables $(\xi_i : i \in \mathbb{N})$ on (X, \mathcal{B}, μ) is spreadable then there exists a probability space (X', \mathcal{B}', μ') , a Borel function $f : [0, 1]^2 \rightarrow [0, 1]$ and a collection of Uniform $[0, 1]$ i.i.d. random variables $\{\zeta_i : i \in \mathbb{N}\} \cup \{\zeta_\emptyset\}$ on X' so that $(\xi_i : i \in \mathbb{N}) =^{\text{dist}} (f(\zeta_i, \zeta_\emptyset) : i \in \mathbb{N})$.*

This gives us an i.i.d. counterexample to Fact 1, and we can conclude.

1.3 Exchangeable versus spreadable sequences

More precisely, de Finetti obtained this conclusion under a stronger assumption that the sequence $(\xi_i : i \in \mathbb{N})$ is *exchangeable*, that is: for any $n \in \mathbb{N}$, any permutation $\sigma \in \text{Sym}(n)$ and $i_1 < \dots < i_n$ we have $(\xi_{i_1}, \dots, \xi_{i_n}) =^{\text{dist}} (\xi_{i_{\sigma(1)}}, \dots, \xi_{i_{\sigma(n)}})$. And then Ryll-Nardzewski [20] proved that exchangeability is equivalent to spreadability. Curiously, Ryll-Nardzewski has a well-known theorem in model theory, but here he worked as a probabilist. It turns out that this result is closely connected to *stability* — a central notion in modern model theory.

2 Model theoretic stability of probability algebras

Modern model theory begins with *Morley's Categoricity Theorem*: for a countable theory T , if it has only one model of some uncountable cardinality (up to isomorphism), then it has only one model of *every* uncountable cardinality. Morley conjectured [18] a generalization: for a countable theory T , the number of its models of size κ is non-decreasing on uncountable κ .

In his solution of Morley's conjecture [21], Shelah isolated the importance of *stable theories* and developed a lot of machinery to analyze models of stable theories. Stability was rediscovered many times in various contexts, e.g. by Grothendieck in his work on Banach spaces, in dynamics as WAP systems (Weakly Almost Periodic), in machine learning as Littlestone dimension, etc.

In particular, probability algebras are stable, viewed as structures in continuous logic. This is implicit in Ryll-Nardzewski's theorem (“every indiscernible sequence is totally indiscernible”), later in Krivine and Maurey [17], explicit in Ben Yaacov [24]. A more general version was given by Hrushovski (proved using array de Finetti, discussed in Section 3.3), and Tao gave a short elementary proof [23]:

Fact 3. *For any $0 \leq p < q \leq 1$ there is N satisfying: if (X, \mathcal{B}, μ) is a probability space, and $A_1, \dots, A_n, B_1, \dots, B_n \in \mathcal{B}$ satisfy $\mu(A_i \cap B_j) \geq q$ and $\mu(A_j \cap B_i) \leq p$ for all $1 \leq i < j \leq n$, then $n \leq N$.*

This result has many applications: Hrushovski's work on approximate subgroups [10], Tao's algebraic regularity lemma [22], work in topological dynamics by Tsankov, Ibarlucia [11], etc.

3 Intersecting multi-parametric families of events

We obtain a *higher arity* generalization of Fact 1:

Theorem 1 (Chernikov, Towsner [7]). *For every finite bipartite graph $H = (V_0, W_0, E_0)$ and $\varepsilon \in (0, 1]$ there exists a finite bipartite graph $G = (V, W, E)$ and $\delta > 0$ (depending only on H and ε) satisfying the following. Assume that (X, \mathcal{B}, μ) is a probability space, and for every $(v, w) \in V \times W$ a measurable set $A_{v,w} \in \mathcal{B}$ so that: for any $(v, w) \in E, (v', w') \notin E$ we have $\mu(A_{v,w}) - \mu(A_{v',w'}) \geq \varepsilon$. Then there exists an induced subgraph $H' = (V', W', E')$ of G (i.e. $V' \subseteq V, W' \subseteq W$ and $E' = E \cap (V' \times W')$) isomorphic to H so that:*

$$\mu \left(\left(\bigcap_{(v,w) \in E'} A_{v,w} \right) \cap \left(\bigcap_{(v,w) \in (V' \times W') \setminus E'} X \setminus A_{v,w} \right) \right) \geq \delta.$$

More precisely, Theorem 1 follows from [7, Lemma 10.13] and compactness. With high probability, a sufficiently large G taken at random will work. More generally, we prove it there for partite hypergraphs of any arity instead of just graphs. The question is motivated by *Keisler randomizations* of first-order structures [16] and whether they preserve NIP (Ben Yaacov, related to work of Talagrand on VC-dimension for functions [4]) and its higher arity generalization *n-dependence* (where Ben Yaacov's analytic proof for $n = 1$ does not seem to generalize).

In what follows we outline a proof of Theorem 1. The overall strategy is similar to the proof above for sequences of events, but each of the steps becomes harder.

3.1 Structural Ramsey theory, and infinite limits of Ramsey classes

Let \mathcal{K} be a class of finite \mathcal{L}_0 -structures, where \mathcal{L}_0 is a relational language (for example, finite graphs). For $A, B \in \mathcal{K}$, let $\binom{B}{A}$ be the set of all $A' \subseteq B$ s.t. $A' \cong A$ (we work with substructures instead of embeddings for simplicity).

Definition 2. \mathcal{K} is Ramsey if for any $A, B \in \mathcal{K}$ and $k \in \omega$ there is some $C \in \mathcal{K}$ s.t. for any coloring $f : \binom{C}{A} \rightarrow k$, there is some $B' \in \binom{C}{B}$ s.t. $f \upharpoonright \binom{B'}{A}$ is constant.

The usual Ramsey theorem means: the class of finite linear orders is Ramsey. The subject of structural Ramsey theory started with the following fundamental result of Nesetril, Rödl [19] and Abramson, Harrington [1]:

Fact 4. For any $k \in \mathbb{N}_{\geq 1}$, the class of all finite ordered k -hypergraphs is Ramsey.

Fact 5. Given a Ramsey class \mathcal{K} of finite structures, there exists a unique (up to isomorphism) countable structure \tilde{K} (called the Fraïssé limit of \mathcal{K}) so that the class of its finite substructures is precisely \mathcal{K} and \tilde{K} is homogeneous, i.e. if K_0 and K_1 are finite substructures of \tilde{K} and $f : K_0 \rightarrow K_1$ is an isomorphism, then f extends to an automorphism of the whole structure \tilde{K} .

E.g., if \mathcal{K} is the class of all graphs, its limit \tilde{K} is the countable Rado's random graph; and if \mathcal{K} is the class of finite linear orders, then its limit is $(\mathbb{Q}, <)$.

Understanding which structures are Ramsey is an active subject, with connections to model theory and topological dynamics (Ramsey property of \mathcal{K} is equivalent to the extreme amenability of the group $\text{Aut}(\tilde{K})$ — via Kechris-Pestov-Todorčević correspondence [15]).

3.2 Finding an “exchangeable” counterexample

For any $k \in \mathbb{N}_{\geq 1}$, using that the class of all finite ordered partite k -hypergraphs is Ramsey (viewed as structures in the language $E, P_1, \dots, P_k, <$ with P_i a partition of vertices, $E \subseteq$

$P_1 \times \dots \times P_k$ and $P_i < P_j$ for $i < j$, e.g. [6, Appendix A]), we let \mathcal{OH}_k denote its *Fraïssé limit*. And we let \mathcal{H}_k be its reduct forgetting the ordering.

Assuming that Theorem 1 fails, by Ramsey property and compactness (model theoretic jargon: extracting a generalized indiscernible) we can find some large probability space (X, \mathcal{B}, μ) , $0 < r < s < 1$ and sets $A_{v,w} \in \mathcal{B}$ for all v, w vertices of $\mathcal{OH}_2 = (E; V, W)$ so that:

- $(v, w) \in E \implies \mu(A_{v,w}) \geq s$ and $(v, w) \notin E \implies \mu(A_{v,w}) \leq r$;
- for any two isomorphic (as ordered bipartite graphs) substructures H_1, H_2 of \mathcal{OH}_2 ,

$$(\mathbf{1}_{A_{v,w}} : v, w \in H_1) =^{\text{dist}} (\mathbf{1}_{A_{v,w}} : v, w \in H_2).$$

3.3 (Relatively) Exchangeable random structures

This indiscernibility guarantees certain “exchangeability” in the probabilistic sense. Exchangeable sequences (de Finetti, Section 1.2) and arrays (Aldous-Hoover-Kallenberg, see [14]) of random variables can be presented “up to mixing” using i.i.d. random variables, and we need a certain generalization to relational structures which were studied recently by a number of authors [8, 2, 12].

Definition 3. (1) Let $\mathcal{L}' = \{R'_1, \dots, R'_{k'}\}$, R'_i a relation symbol of arity r'_i . By a random \mathcal{L}' -structure we mean a (countable) collection of random variables $(\xi_{\bar{n}}^i : i \in [k'], \bar{n} \in \mathbb{N}^{r'_i})$ on some probability space $(\Omega, \mathcal{F}, \mu)$ with $\xi_{\bar{n}}^i : \Omega \rightarrow \{0, 1\}$.

(2) Let now $\mathcal{L} = \{R_1, \dots, R_k\}$ be another relational language, with R_i a relation symbol of arity r_i , and let $\mathcal{M} = (\mathbb{N}, \dots)$ be a countable \mathcal{L} -structure with domain \mathbb{N} . We say that a random \mathcal{L}' -structure $(\xi_{\bar{n}}^i : i \in [k'], \bar{n} \in \mathbb{N}^{r'_i})$ is \mathcal{M} -exchangeable if for any two finite subsets $A = \{a_1, \dots, a_\ell\}, A' = \{a'_1, \dots, a'_{\ell'}\} \subseteq \mathbb{N}$

$$\begin{aligned} \text{qftp}_{\mathcal{L}}(a_1, \dots, a_\ell) = \text{qftp}_{\mathcal{L}}(a'_1, \dots, a'_{\ell'}) \implies \\ (\xi_{\bar{n}}^i : i \in [k'], \bar{n} \in A^{r'_i}) =^{\text{dist}} (\xi_{\bar{n}}^i : i \in [k'], \bar{n} \in (A')^{r'_i}). \end{aligned}$$

3.4 A higher amalgamation condition on the indexing structure

Let \mathcal{K} be a collection of finite structures in a relational language \mathcal{L} closed under isomorphism.

Definition 4. For $n \in \mathbb{N}_{\geq 1}$, we say that \mathcal{K} satisfies the n -disjoint amalgamation property (n -DAP) if for every collection of \mathcal{L} -structures $(\mathcal{M}_i = (M_i, \dots) : i \in [n])$ with $\mathcal{M}_i \in \mathcal{K}$, $M_i = [n] \setminus \{i\}$ and $\mathcal{M}_i|_{[n] \setminus \{i,j\}} = \mathcal{M}_j|_{[n] \setminus \{i,j\}}$ for all $i \neq j \in [n]$, there exists an \mathcal{L} -structure $\mathcal{M} = (M, \dots) \in \mathcal{K}$ such that $M = [n]$ and $\mathcal{M}|_{[n] \setminus \{i\}} = \mathcal{M}_i$ for every $1 \leq i \leq n$.

We say that an \mathcal{L} -structure \mathcal{M} satisfies n -DAP if the collection of its finite substructures does. E.g., the generic k -hypergraph \mathcal{H}_k satisfies n -DAP for all n [7, Proposition 9.6], but $(\mathbb{Q}, <)$ fails 3-DAP.

3.5 Presentation for random relational structures

Fact 6 (Crane, Towsner [8]; generalizing Aldous-Hoover-Kallenberg [3, 9, 13]). Let $\mathcal{L}' = \{R'_i : i \in [k']\}$, $\mathcal{L} = \{R_i : i \in [k]\}$ be finite relational languages with all R'_i of arity at most r' , and $\mathcal{M} = (\mathbb{N}, \dots)$ a countable homogeneous \mathcal{L} -structure that has n -DAP for all $n \geq 1$. Suppose that

$(\xi_{\bar{n}}^i : i \in [k'], \bar{n} \in \mathbb{N}^{r'_i})$ is a random \mathcal{L}' -structure that is \mathcal{M} -exchangeable, such that the relations R'_i are symmetric with probability 1.

Then there exists a probability space $(\Omega', \mathcal{F}', \mu')$, $\{0, 1\}$ -valued Borel functions $f_1, \dots, f_{r'}$ and a collection of Uniform $[0, 1]$ i.i.d. random variables $(\zeta_s : s \subseteq \mathbb{N}, |s| \leq r')$ on Ω' so that

$$(\xi_{\bar{n}}^i : i \in [k'], \bar{n} \in \mathbb{N}^{r'_i}) = ^{\text{dist}} \left(f_i \left(\mathcal{M}|_{\text{rng } \bar{n}}, (\zeta_s)_{s \subseteq \text{rng } \bar{n}} \right) : i \in [k'], \bar{n} \in \mathbb{N}^{r'_i} \right),$$

where $\text{rng } \bar{n}$ is the set of its distinct elements, and \subseteq denotes “subsequence”.

3.6 Getting rid of the ordering

Our counterexample from Section 3.2 is only guaranteed to be \mathcal{OH}_n -exchangeable (and the ordering is unavoidable in the Ramsey theorem for hypergraphs) — but the presentation theorem in Fact 6 requires n -DAP (and linear orders fail 3-DAP). However, using Fact 3 inductively, we can show that \mathcal{OH}_n -exchangeability already implies \mathcal{H}_n -exchangeability (i.e., with respect to the reduct forgetting the ordering), using that the theory of probability algebras is *stable!* (See [7, Lemma 10.15] for the details.)

Applying the exchangeable presentation to the counterexample, we finally reduce (modulo some mixing) to working with *independent* random variables in the proof of Theorem 1, and can conclude the proof.

4 Open questions and future directions

Question 1. Our proof of Theorem 1 is non-constructive and relies on a compactness argument. It would be interesting to obtain explicit bounds on $|G|$ and δ in terms of H and ε . Do there exist infinitary/density versions of this result (similarly to Fact 1)?

Question 2. Apart from k -partite k -hypergraphs, which other classes of structures satisfy an analog of Theorem 1? E.g., there is a growing list of Ramsey classes of finite structures, for which also an appropriate analog of Fact 6 holds. The following example illustrates that these two properties alone are not sufficient:

Example (Tim Austin) Theorem 1 does not hold for graphs (as opposed to *bipartite* graphs). Indeed, let H be the triangle K_3 , and $\varepsilon = 1/2$. Consider any graph $G = (V, E)$. On some probability space (Ω, Σ, μ) , let $(\pi_v : v \in V)$ be a process of independent uniform $\{0, 1\}$ -valued random variables, and consider the events $A_{v,w}$ defined by: $A_{v,w} := (\pi_v \neq \pi_w)$ if $(v, w) \in E$, and $A_{v,w} := \emptyset$ if $(v, w) \notin E$. Then $\mu(A_{v,w})$ is equal to $1/2$ if $(v, w) \in E$, but equal to 0 if $(v, w) \notin E$. However, for any induced triangle in G , say with vertices u, v, w , we have $\mu(A_{u,v} \cap A_{v,w} \cap A_{w,u}) = \mu(\pi_u \neq \pi_v \neq \pi_w \neq \pi_u) = \mu(\emptyset) = 0$.

As mentioned above, Theorem 1 is the main ingredient in our proof that Keisler randomization of first-order theories preserves n -dependence, for all $n \in \mathbb{N}_{\geq 1}$ ([7, Corollary 11.3]).

Question 3. Apart from n -dependence, what are the other higher arity tameness notions from model theory preserved under Keisler randomization? E.g., is FOP_n preserved?

References

- [1] Fred G Abramson and Leo A Harrington. Models without indiscernibles. *The Journal of Symbolic Logic*, 43(3):572–600, 1978.
- [2] Nathanael Ackerman. Representations of $\text{Aut}(M)$ -invariant measures. *Preprint, arXiv:1509.06170*, 2015.

- [3] David J Aldous. Representations for partially exchangeable arrays of random variables. *Journal of Multivariate Analysis*, 11(4):581–598, 1981.
- [4] Itaï Ben Yaacov. Continuous and random Vapnik-Chervonenkis classes. *Israel Journal of Mathematics*, 1(173):309–333, 2009.
- [5] Vitaly Bergelson. Sets of recurrence of \mathbb{Z}^m -actions and properties of sets of differences in \mathbb{Z}^m . *Journal of the London Mathematical Society*, 2(2):295–304, 1985.
- [6] Artem Chernikov, Daniel Palacin, and Kota Takeuchi. On n-dependence. *Notre Dame Journal of Formal Logic*, 60(2):195–214, 2019.
- [7] Artem Chernikov and Henry Towsner. Hypergraph regularity and higher arity VC-dimension. *Preprint*, <https://arxiv.org/abs/2010.00726>, 2020.
- [8] Harry Crane and Henry Towsner. Relatively exchangeable structures. *The Journal of Symbolic Logic*, 83(2):416–442, 2018.
- [9] Douglas N Hoover. Relations on probability spaces and arrays of random variables. *Preprint*, Institute for Advanced Study, 1979.
- [10] Ehud Hrushovski. Stable group theory and approximate subgroups. *Journal of the American Mathematical Society*, 25(1):189–243, 2012.
- [11] Tomás Ibarlucía and Todor Tsankov. A model-theoretic approach to rigidity of strongly ergodic, distal actions. *Annales Scientifiques de l’École Normale Supérieure*, 54(3):751–777, 2021.
- [12] Colin Jähel and Todor Tsankov. Invariant measures on products and on the space of linear orders. *Journal de l’École polytechnique—Mathématiques*, 9:155–176, 2022.
- [13] Olav Kallenberg. Spreading and predictable sampling in exchangeable sequences and processes. *The Annals of Probability*, pages 508–534, 1988.
- [14] Olav Kallenberg. *Probabilistic symmetries and invariance principles*, volume 9. Springer, 2005.
- [15] AS Kechris, VG Pestov, and S Todorcevic. Fraïssé limits, Ramsey theory, and topological dynamics of automorphism groups. *Geometric & Functional Analysis GAFA*, 15:106–189, 2005.
- [16] H Jerome Keisler. Randomizing a model. *Advances in Mathematics*, 143(1):124–158, 1999.
- [17] Jean-Louis Krivine and Bernard Maurey. Espaces de Banach stables. *Israel Journal of Mathematics*, 39:273–295, 1981.
- [18] Michael Morley. Categoricity in power. *Transactions of the American Mathematical Society*, 114(2):514–538, 1965.
- [19] Jaroslav Nešetřil and Vojtěch Rödl. Partitions of finite relational and set systems. *Journal of Combinatorial Theory, Series A*, 22(3):289–312, 1977.
- [20] Czesław Ryll-Nardzewski. On stationary sequences of random variables and the de Finetti's equivalence. In *Colloquium Mathematicum*, volume 4, pages 149–156. Polska Akademia Nauk. Instytut Matematyczny PAN, 1957.
- [21] Saharon Shelah. *Classification theory: and the number of non-isomorphic models*. Elsevier, 1990.
- [22] Terence Tao. Expanding polynomials over finite fields of large characteristic, and a regularity lemma for definable sets. *Contributions to Discrete Mathematics*, 10(1), 2015.
- [23] Terrence Tao. “A spectral theory proof of the algebraic regularity lemma”, blogpost. <https://terrytao.wordpress.com/2013/10/29/a-spectral-theory-proof-of-the-algebraic-regularity-lemma/>, 2013.
- [24] Itaï Ben Yaacov. On theories of random variables. *Israel Journal of Mathematics*, 194:957–1012, 2013.