Complete r-partite r-graphs are Sidorenko
A brief exposition

Yuval Wigderson

Abstract

The result in the title of this note is extremely useful, and also well-known. However, I have been unable to find a concise proof of it, using modern terminology and notation. The aim of this note is to fill this gap.

Since the goal of this note is to present the result quickly, the note is structured somewhat unusually: we begin with the statement and proof, and only afterwards give a historical survey and indicate why this result is interesting.

1 Notation and terminology

For an integer $r \geq 2$, an *r-uniform hypergraph* (or *r-graph* for short) consists of a finite vertex set V and a collection of *(hyper)edges* $E \subseteq \binom{V}{r}$, each comprising r vertices. Given two r-graphs $\mathcal{H}$ and $\mathcal{G}$, a *homomorphism* $\varphi : \mathcal{H} \to \mathcal{G}$ is a function $V(\mathcal{H}) \to V(\mathcal{G})$ which maps edges to edges. The set of all homomorphisms $\varphi : \mathcal{H} \to \mathcal{G}$ is denoted by $\hom(\mathcal{H}, \mathcal{G})$, and we define the *homomorphism density* to be

$$t(\mathcal{H}, \mathcal{G}) := \frac{|\hom(\mathcal{H}, \mathcal{G})|}{|V(\mathcal{G})|^{|V(\mathcal{H})|}}.$$

This is the fraction of all functions $V(\mathcal{H}) \to V(\mathcal{G})$ which are homomorphisms; equivalently, it is the probability that a random function $V(\mathcal{H}) \to V(\mathcal{G})$ is a homomorphism. If the vertices of $\mathcal{H}$ are $v_1, \ldots, v_h$, then we may equivalently write

$$t(\mathcal{H}, \mathcal{G}) = \mathop{\mathbb{E}}_{X_1,\ldots,X_h} \left[\prod_{e \in E(\mathcal{H})} \mathcal{G}(\{X_i : v_i \in e\}) \right].$$

Here, the expectation is over uniformly and independently random vertices $X_1, \cdots, X_h \in V(\mathcal{G})$, and we abuse notation and denote by $\mathcal{G}$ also the indicator function of edges of $\mathcal{G}$; in other words, $\mathcal{G}(\{u_1, \ldots, u_r\})$ equals 1 if $\{u_1, \ldots, u_r\} \in E(\mathcal{G})$, and 0 otherwise.

Let $\mathcal{K}_r$ denote the r-graph with r vertices and a single edge. An r-graph $\mathcal{H}$ is called *r-partite* if there exists a homomorphism $\mathcal{H} \to \mathcal{K}_r$. The *complete r-partite r-graph* with parts of size $s_1, \ldots, s_r$, denoted $\mathcal{K}_{s_1,\ldots,s_r}$, is the r-graph whose vertex set is partitioned into r

blocks of sizes $s_1, \ldots, s_r$, and an r-tuple is an edge if and only if it contains one vertex from each block. Finally, we say that an r-graph $\mathcal{H}$ is *Sidorenko* if

$$t(\mathcal{H}, \mathcal{G}) \geq t(\mathcal{K}_r, \mathcal{G})^{|E(\mathcal{H})|} \tag{$*$}$$

holds for every r-graph $\mathcal{G}$. It is easy to see that $\mathcal{H}$ being Sidorenko implies that $\mathcal{H}$ is r-partite, for if $\mathcal{H}$ is not r-partite then $(*)$ fails with $\mathcal{G} = \mathcal{K}_r$.

2 The result and its proof

The main result of this note is the following theorem, essentially due to Erdős [4] (though see Section 3 for more historical backround).

Theorem 1. *If $\mathcal{H}$ is a complete r-partite r-graph, then $\mathcal{H}$ is Sidorenko.*

Theorem 1 will follow from the following lemma. Given a positive integer s and an r-graph $\mathcal{H}$ with vertices $v_1, \ldots, v_h$, we denote by $\mathcal{H}^+(s)$ the r-graph gotten by making s clones of v_h, i.e. by replacing v_h by s vertices $v_h^{(1)}, \ldots, v_h^{(s)}$ such that the induced subhypergraph $\mathcal{H}[\{v_1, \ldots, v_{h-1}, v_h^{(j)}\}]$ is isomorphic to $\mathcal{H}$ for all $j \in [s]$.

Lemma 2. *Let $\mathcal{H}$ be an r-graph with vertices $v_1, \ldots, v_h$, let s be a positive integer, and suppose that every edge of $\mathcal{H}$ contains v_h. If $\mathcal{H}$ is Sidorenko, then so is $\mathcal{H}^+(s)$.*

Proof. For any r-graph $\mathcal{G}$, we may write

$$t(\mathcal{H}, \mathcal{G}) = \mathop{\mathbb{E}}_{X_1, \ldots, X_h} \left[\prod_{e \in E(\mathcal{H})} \mathcal{G}(\{X_i : v_i \in e\}) \right] = \mathop{\mathbb{E}}_{X_1, \ldots, X_{h-1}} \left[\mathop{\mathbb{E}}_{X_h} \left[\prod_{e \in E(\mathcal{H})} \mathcal{G}(\{X_i : v_i \in e\}) \right] \right].$$

We now apply Jensen's inequality to the outer expectation, using convexity of the function $x \mapsto x^s$, which yields

$$\begin{aligned} t(\mathcal{H}, \mathcal{G}) &\leq \mathop{\mathbb{E}}_{X_1, \ldots, X_{h-1}} \left[\left(\mathop{\mathbb{E}}_{X_h} \left[\prod_{e \in E(\mathcal{H})} \mathcal{G}(\{X_i : v_i \in e\}) \right] \right)^s \right]^{1/s} \\ &= \mathop{\mathbb{E}}_{X_1, \ldots, X_{h-1}} \left[\mathop{\mathbb{E}}_{X_h^{(1)}, \ldots, X_h^{(s)}} \prod_{j=1}^{s} \prod_{e \in E(\mathcal{H})} \mathcal{G}(\{X_i : v_i \in e \setminus \{v_h\}\} \cup \{X_i^{(j)}\}) \right]^{1/s} \\ &= t(\mathcal{H}^+(s), \mathcal{G})^{1/s}. \end{aligned}$$

The notation is somewhat convoluted, but what's really going on is simple: by expanding the sth power on the inner expectation, we get s iid copies $X_h^{(1)}, \ldots, X_h^{(s)}$ of the random variable X_h. By doing this, the factor corresponding to any edge $e \in E(\mathcal{H})$ now turns into

s factors, one for each of the variables $X_h^{(1)}, \ldots, X_h^{(s)}$. Thus, the complicated product we see is really a product over the edges of $\mathcal{H}^+(s)$, which yields the final equality.

Finally, since every edge of $\mathcal{H}$ contains v_h, we see that $|E(\mathcal{H}^+(s))| = s|E(\mathcal{H})|$, so the fact that $\mathcal{H}$ is Sidorenko yields

$$t(\mathcal{H}^+(s), \mathcal{G}) \geq t(\mathcal{H}, G)^s \geq \left(t(\mathcal{K}_r, \mathcal{G})^{|E(\mathcal{H})|}\right)^s = t(\mathcal{K}_r, \mathcal{G})^{|E(\mathcal{H}^+(s))|},$$

showing that $\mathcal{H}^+(s)$ is Sidorenko. □

We can now prove Theorem 1.

Proof of Theorem 1. We have that $\mathcal{K}_r$ is Sidorenko, since $(*)$ is simply an equality for $\mathcal{H} = \mathcal{K}_r$. By applying Lemma 2, this shows that $\mathcal{K}_{s_1,1,1,\ldots,1}$ is Sidorenko for any positive integer s_1. By applying Lemma 2 again, we find that $\mathcal{K}_{s_1,s_2,1,1,\ldots,1}$ is Sidorenko for all positive integers s_1, s_2. Continuing in this fashion, we conclude that $\mathcal{K}_{s_1,s_2,\ldots,s_r}$ is Sidorenko for any positive integers $s_1, s_2, \ldots, s_r$. □

Remark. Inspecting the proof of Lemma 2 (and thus of Theorem 1), one finds that we never really used the fact that $X_1, \ldots, X_h$ are uniformly random on the finite set $V(\mathcal{G})$, nor that the function $\mathcal{G}$ is a symmetric $\{0,1\}$-valued function (i.e. the indicator function of the edges of an r-graph). Because of this, the exact same proof shows that $(*)$ holds for $\mathcal{H}$ a complete r-partite r-graph even when $\mathcal{G}$ is an arbitrary non-negative measurable function on a product of r probability spaces. This perspective naturally leads to the study of *hypergraphons*, which we will not discuss further; see e.g. the book [10] for more.

3 History

The $r = 2$ version of Theorem 1 is essentially due to Kővári, Sós, and Turán [9] from 1954, who were arguably the first to observe that simple convexity arguments (i.e. applications of Jensen's inequality) could be applied to the indicator function of edges of graphs to yield interesting results in graph theory. The case for general r is essentially due to Erdős [4], who proved it by induction on r.

Although Erdős's simple convexity argument yields Theorem 1, the result stated in his paper is weaker in several fashions than Theorem 1. The first, not very significant, fashion is that his result is only stated for *balanced* complete r-uniform r-partite hypergraphs, i.e. for the case when $s_1 = \cdots = s_r$. The second fashion is that Erdős does not count homomorphisms $\mathcal{K}_{s,\ldots,s} \to \mathcal{G}$, but rather proves that if $\mathcal{G}$ is a *dense* hypergraph on sufficiently many vertices, then it contains a copy of $K_{s,\ldots,s}$ as a subhypergraph; this is a weaker result, since a lower bound on the homomorphism density can be used to prove the existence of copies (see Section 4 for more on this topic). Nikiforov [11] addressed both these aspects of Erdős's paper by proving a counting result for copies of arbitrary complete r-partite r-graphs, but his bound is weaker than that in Theorem 1; roughly, there is an extra factor of r in the exponent of his version of $(*)$.

The other main fashion in which Erdős's paper is different from the presentation above is in the language. His paper is written entirely in terms of hypergraphs and subhypergraphs, rather than in the analytic language of homomorphism densities and expectations. This analytic language was pioneered by Sidorenko (see e.g. [16, 17, 18]), and has since become extremely prevalent in extremal combinatorics. It really appears that an analytic framework is the "correct" framework for stating and proving results like Theorem 1, and it has had a huge influence over the development of extremal combinatorics in the past 30 years.

As stated in the abstract, Theorem 1 is well-known among experts in the field, though I haven't been able to find this sort of analytic statement or proof in the literature. The result follows from the techniques in [2, Section 5], but they prove a much more general result using many more techniques and ideas; nonetheless, to my knowledge, that is the only place in the literature from which one can deduce the full statement of Theorem 1 directly. If you are reading this and know of a statement of Theorem 1 anywhere else, please let me know!

4 Why do we care?

Why do we care about proving that some hypergraphs are Sidorenko? There are a number of reasons. One important reason is that the inequality $(*)$ is always tight if it is true. Indeed, if $p \in [0,1]$ is fixed and $\mathcal{G}$ is taken to be a random n-vertex hypergraph, where every r-tuple is taken to be an edge of $\mathcal{G}$ independently with probability p, then it is a simple exercise to show that with high probability,

$$t(\mathcal{K}_r, \mathcal{G}) = p + o(1) \qquad \text{and} \qquad t(\mathcal{H}, \mathcal{G}) = p^{|E(\mathcal{H})|} + o(1),$$

where the $o(1)$ terms tend to 0 as $n \to \infty$. Therefore, $(*)$ is the strongest inequality one can hope to hold for all r-graphs $\mathcal{G}$.

Another reason goes to the original motivation of Kővári–Sós–Turán [9] and Erdős [4] mentioned above, which is the classical extremal problem for graphs and hypergraphs. Given an r-graph $\mathcal{H}$ and an integer n, the *extremal number* $\text{ex}(n, \mathcal{H})$ is defined as the maximum number of edges of an r-graph $\mathcal{G}$ on n vertices which does not contain $\mathcal{H}$ as a subhypergraph. Using Theorem 1, we can obtain an upper bound on $\text{ex}(n, \mathcal{H})$ when $\mathcal{H}$ is r-partite. Note that we have the trivial bound $\text{ex}(n, \mathcal{H}) \le \binom{n}{r} = \Theta(n^r)$.

Theorem 3 (Kővári–Sós–Turán [9], Erdős [4])**.** *If $\mathcal{H}$ is an r-partite r-graph with parts of sizes $s_1 \ge s_2 \ge \cdots \ge s_r$, then $\text{ex}(n, \mathcal{H}) = O_{s_1,\ldots,s_r}(n^{r-1/(s_2 \cdots s_r)})$.*

Observe that since $\mathcal{H}$ is a subgraph of $\mathcal{K}_{s_1,\ldots,s_r}$, it suffices to prove this theorem for $\mathcal{H} = \mathcal{K}_{s_1,\ldots,s_r}$. The idea of the proof is as follows. If $\mathcal{G}$ is an n-vertex r graph with εn^r edges, then $t(\mathcal{K}_r, \mathcal{G}) \ge \varepsilon$. So by Theorem 1, there are at least $\varepsilon^{s_1 \cdots s_r} n^{s_1 + \cdots + s_r}$ homomorphisms $\mathcal{K}_{s_1,\ldots,s_r} \to \mathcal{G}$. If $\mathcal{G}$ does not contain $\mathcal{K}_{s_1,\ldots,s_r}$ as a subhypergraph, then all these homomorphisms must be non-injective, and one can upper-bound the number of non-injective homomorphisms. This upper bound yields an upper bound on ε, which implies the desired result.

Despite its simplicity, Theorem 3 is a foundational result in extremal hypergraph theory. It implies the important fact that Turán densities are blowup-invariant, which in turn can be used to give a quick proof of the Erdős–Stone–Simonovits theorem [5, 6], the cornerstone result in extremal graph theory. For more on these topics, see e.g. the survey [7].

One further reason to care about $(*)$ is that in general, it is extremely difficult to get lower bounds on $t(\mathcal{H}, \mathcal{G})$ when $\mathcal{H}$ is a general r-graph, despite the fact that the quantities $t(\mathcal{H}, \mathcal{G})$ are of central importance in extremal (hyper)graph theory. Even in the simplest case $r = 2$, and when $\mathcal{H}$ is the smallest non-bipartite graph K_3, it was a major breakthrough of Razborov [13] to determine the minimum value of $t(K_3, G)$ among all graphs G of any fixed edge density. This result was extended by Nikiforov [12] to K_4 and by Reiher [14] to all cliques. Because of the difficulty of this problem, and the complexity of the answer proved by Razborov, Nikiforov, and Reiher, it is extremely appealing to see that for some choices of $\mathcal{H}$, the simple (and tight!) lower bound $(*)$ holds. In the case $r = 2$, it is conjectured that this works in the maximum possible generality.

Conjecture 4 (Sidorenko's conjecture [15])**.** *Every bipartite graph is Sidorenko.*

Despite a lot of progress towards this conjecture (e.g. [1, 3, 8, 19]), it remains open in general. We remark that the natural hypergraph analogue of Sidorenko's conjecture is false, as shown by Sidorenko himself [15].

References

[1] D. Conlon, J. Fox, and B. Sudakov, An approximate version of Sidorenko's conjecture, *Geom. Funct. Anal.* **20** (2010), 1354–1366.

[2] D. Conlon and J. Lee, Finite reflection groups and graph norms, *Adv. Math.* **315** (2017), 130–165.

[3] D. Conlon and J. Lee, Sidorenko's conjecture for blow-ups, *Discrete Anal.* (2021), Paper No. 2, 13.

[4] P. Erdős, On extremal problems of graphs and generalized graphs, *Israel J. Math.* **2** (1964), 183–190.

[5] P. Erdős and M. Simonovits, A limit theorem in graph theory, *Studia Sci. Math. Hungar.* **1** (1966), 51–57.

[6] P. Erdös and A. H. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.* **52** (1946), 1087–1091.

[7] P. Keevash, Hypergraph Turán problems, in *Surveys in combinatorics 2011*, *London Math. Soc. Lecture Note Ser.*, vol. 392, Cambridge Univ. Press, Cambridge, 2011, pp. 83–139.

[8] J. H. Kim, C. Lee, and J. Lee, Two approaches to Sidorenko's conjecture, *Trans. Amer. Math. Soc.* **368** (2016), 5057–5074.

[9] T. Kövari, V. T. Sós, and P. Turán, On a problem of K. Zarankiewicz, *Colloq. Math.* **3** (1954), 50–57.

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

[11] V. Nikiforov, Complete r-partite subgraphs of dense r-graphs, *Discrete Math.* **309** (2009), 4326–4331.

[12] V. Nikiforov, The number of cliques in graphs of given order and size, *Trans. Amer. Math. Soc.* **363** (2011), 1599–1618.

[13] A. A. Razborov, On the minimal density of triangles in graphs, *Combin. Probab. Comput.* **17** (2008), 603–618.

[14] C. Reiher, The clique density theorem, *Ann. of Math. (2)* **184** (2016), 683–707.

[15] A. Sidorenko, A correlation inequality for bipartite graphs, *Graphs Combin.* **9** (1993), 201–204.

[16] A. Sidorenko, An analytic approach to extremal problems for graphs and hypergraphs, in *Extremal problems for finite sets (Visegrád, 1991)*, *Bolyai Soc. Math. Stud.*, vol. 3, János Bolyai Math. Soc., Budapest, 1994, pp. 423–455.

[17] A. F. Sidorenko, Extremal problems in graph theory and functional-analytic inequalities, in *Proceedings of the All-Union seminar on discrete mathematics and its applications (Russian) (Moscow, 1984)*, Moskov. Gos. Univ., Mekh.-Mat. Fak., Moscow, 1986, pp. 99–105.

[18] A. F. Sidorenko, Inequalities for functionals generated by bipartite graphs, *Diskret. Mat.* **3** (1991), 50–65.

[19] B. Szegedy, An information theoretic approach to Sidorenko's conjecture, 2014. Preprint available at arXiv:1406.6738.