1 Graph regularity

The *regularity lemma* asserts, in the broadest possible generality, that any discrete object can be decomposed into a small number of random-like pieces. Such a statement is of immense utility in combinatorics, computer science, number theory, and related fields: it means that for a wide variety of applications, we do not need to solve the problem on an arbitrary object, but only on an appropriate pseudorandom object.

In the statement of the regularity lemma, "small" usually means "a constant", where the (constant) number of pieces in the decomposition depends only on the desired quality of the pseudorandomness, and not on the size of the input. Nevertheless, for applications, we are often interested in exactly *how* small this constant is, and, unfortunately, it is usually absolutely enormous. It is therefore very natural to wonder whether adding extra assumptions allows one to improve this: if we assume that the input object is simple in some appropriate sense, can we obtain better guarantees on the size of the regularity decomposition than what we have in the worst case?

In this talk, my goal is to try to answer these questions in the setting of graph and hypergraph regularity. As we will discover, the answers to these questions turn out to be intimately related to the notion of *Vapnik–Chervonenkis (VC) dimension* and various generalizations of it.

Given a bipartite graph between vertex sets X, Y, we denote by $e(X,Y)$ the number of edges in it, and by $d(X,Y) \coloneqq e(X,Y)/(|X||Y|)$ the *edge density*. The pair (X,Y) is said to be *ε-regular* if for all $X' \subseteq X, Y' \subseteq Y$ with $|X'| \geqslant \varepsilon|X|, |Y'| \geqslant \varepsilon|Y|$, we have

$$|d(X',Y') - d(X,Y)| \leqslant \varepsilon.$$

Intuitively, this says that the edges between X and Y are "well-distributed". Moreover, classical results due (essentially) to Thomason and Chung–Graham–Wilson imply that this notion of pseudorandomness is, in some sense, the only one: it is equivalent, up to a polynomial change in ε, to any other reasonable notion one could come up with. Note that in the definition of ε-regularity, the parameter ε is our "quality" parameter, and the smaller it is, the better pseudorandomness guarantee we have. In particular, this definition is extremely meaningful once ε is much smaller than $d(X,Y)$, whereas if $d(X,Y)$ is much smaller than ε, then it is essentially meaningless.

The main reason why ε-regularity is a useful notion is that ε-regular graphs enjoy many of the properties shared by truly random graphs. One important example is the *counting lemma*, which guarantees the presence of desired substructures within regular pairs. There are many variants of the counting lemma; for our purposes, the following one will be particularly relevant.

Theorem 1.1 (Induced counting lemma)**.** *Let B be a fixed bipartite graph. If ε is sufficiently small, then any ε-regular pair (X,Y) contains a bi-induced copy of B, assuming that $d(X,Y)$ is bounded away from 0 and 1.*

Here, a *bi-induced* copy of B means a copy of B in which all edges are present, all non-edges across are not present, but we don't care what happens within X and Y. The condition that ε be sufficiently small is clearly necessary, for the reasons discussed above (ε-regularity is meaningless if ε is large), as is the condition that $d(X,Y)$ is bounded away from 0 and 1: an empty or complete pair is trivially ε-regular for all $\varepsilon > 0$, but will not contain a bi-induced copy of any interesting bipartite graph B.

The other main utility of the notion of ε-regularity is, as indicated above, that it supports a *regularity lemma*, stating that every graph can be decomposed into regular pieces. To make this formal, let us say that an equipartition $V(G) = V_1 \sqcup \cdots \sqcup V_m$ of the vertex set of a graph is *ε-regular* if all but at most εm^2 pairs (V_i, V_j) are ε-regular. With this notion, Szemerédi's regularity lemma states the following.

Theorem 1.2 (Szemerédi's regularity lemma)**.** *Every graph has an ε-regular equipartition into at most $M(\varepsilon)$ parts.*

Crucially, the number of parts is *small*: an absolute constant depending only on ε, the quality of the pseudorandomness we wish to achieve. Unfortunately, this "small" is really rather large: Szemerédi's proof gave a bound of the form

$$M(\varepsilon) \leqslant \operatorname{twr}(\varepsilon^{-5}) = \left.2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}\right\}\varepsilon^{-5}.$$

For several decades it remained a major open problem to improve this horrendous bound, until Gowers showed that this was impossible: he constructed a graph G for which every ε-regular equipartition contains at least $\operatorname{twr}(\varepsilon^{-c})$ many parts, for some fixed $c > 0$. So in the worst case, Szemerédi's regularity lemma really does require a tower-type number of parts in the regular decomposition. As a consequence, any application of the regularity lemma will necessarily also encounter such terrible bounds.

Nevertheless, in applications we don't necessarily care about the worst-case graph, but rather about the graph we are presented with. So it is very natural to wonder whether Gowers's example is particularly pathological, and thus whether naturally-occurring graph classes support much smaller ε-regular partitions. Unfortunately, the answer to this question is generally no: it is not hard to adapt Gowers's example to lie in many natural graph classes, e.g. to make it bipartite or chordal.

That being said, a few decades ago people started proving results stating that *geometrically defined* graphs actually do support much smaller ε-regular partitions, of size only $\operatorname{poly}(\varepsilon^{-1})$. These days, the way we think about all of these results is that graphs with *bounded VC dimension* have small regular partitions.

2 VC dimension

Over the past 50 years, VC dimension has become one of the central notions in combinatorics, statistical learning theory, model theory, and group theory, and has proved itself to be one of the most important notions of "low complexity" for discrete objects. I will begin by giving a wrong (but nearly right) definition of VC dimension for graphs, in order to explain why this notion relates to the regularity lemma.

2.1 Fake VC dimension

Definition 2.1. The *fake VC dimension* of a graph G is the smallest integer d such that there is a d-vertex bipartite graph B such that B is not a bi-induced subgraph of G.

Equivalently, if B is not a bi-induced subgraph of G, then $\dim_{\mathrm{fVC}}(G) \leqslant |B|$.

In all of our applications, we will think of G as some enormous graph (or, more precisely, as an element of a sequence of graphs whose order tends to infinity), whereas the VC dimension will always be some unspecified constant. Thus, informally, we will say that a graph has *bounded (fake) VC dimension* if there is some fixed bipartite B which is not a bi-induced subgraph of G.

We've already encountered bi-induced subgraphs, and this is not a coincidence: the induced counting lemma immediately implies the following.

Lemma 2.2. *If G has bounded fake VC dimension, then every ε-regular pair in G must have density close to 0 or 1.*

Indeed, if there were some ε-regular pair whose density was bounded away from 0 and 1, we could apply the induced counting lemma to find a bi-induced copy of any fixed bipartite B; by assumption, there is some such B which is not a bi-induced subgraph of G, yielding the lemma.

Motivated by this lemma, let us say that a pair (X, Y) is *ε-homogeneous* if $d(X, Y) \in [0, \varepsilon] \cup [1-\varepsilon, 1]$. It is not hard to show that this is a strictly stronger notion than ε-regularity, up to a polynomial change in ε. More precisely, if a pair is ε^3-homogeneous, then it is ε-regular; however, there exist ε-regular pairs which are not $\frac{1}{3}$-homogeneous (e.g. a random bipartite graph will have this property).

Using the previous lemma, we immediately discover that bounded (fake) VC dimension allows us to substantially strengthen the regularity lemma. Let us say that an equipartition $V(G) = V_1 \sqcup \cdots \sqcup V_m$ is ε-homogeneous if all but at most εm^2 pairs (V_i, V_j) are ε-homogeneous.

Theorem 2.3. *If G has bounded (fake) VC dimension, then G has an ε-homogeneous partition into at most $M(\varepsilon)$ parts.*

Proof sketch. Apply Szemerédi's regularity lemma to find an ε'-regular partition of G, for some ε' depending on ε. Now apply Lemma 2.2 to conclude that every ε'-regular pair is actually ε-homogeneous. □

At the moment, we have not learned anything new about the number of parts in the homogeneous partition, since we obtained it by just applying the usual regularity lemma. However, let me stress that the conclusion of this theorem is very strong, and genuinely much stronger than what holds in the worst case: without the assumption of bounded VC dimension, we should not expect G to have an ε-homogeneous partition into *any* bounded number of parts. For example, it is not hard to show that a random graph does not have such a partition.

2.2 Real VC dimension

We have already obtained a qualitative strengthening of the regularity lemma for graphs of bounded VC dimension, but I also promised you a quantitative strengthening. Clearly if that is our goal, we should not simply apply Szemerédi's regularity lemma as a black box, and should instead find a proof that is tailored to this special setting.

In order to do this, we will need to learn a bit more about VC dimension, including its real definition.

Definition 2.4. Let V be a finite set, and let $\mathcal{F} \subseteq 2^V$ be a family of subsets of V. A set $S \subseteq V$ is *shattered by* V if for every $R \subseteq S$, there is some $F \in \mathcal{F}$ such that $F \cap S = R$. Equivalently, S is shattered by $\mathcal{F}$ if $\{F \cap S : F \in \mathcal{F}\} = 2^S$.

Equivalently, let us make an $\mathcal{F} \times V$ matrix whose rows are the indicator vectors of the sets in $\mathcal{F}$. A shattered set S is a set of columns in which we see all the $2^{|S|}$ possible binary patterns.

The *VC dimension* of $\mathcal{F}$ is defined to be the maximum size of a set shattered by $\mathcal{F}$.

Before continuing with our discussion of abstract VC dimension, let us briefly see the connection to the fake VC dimension discussed above. For a graph G, we define its VC dimension to be the VC dimension of the set system $\mathcal{N}(G) = \{N(v) : v \in V(G)\} \subseteq 2^{V(G)}$ consisting of all neighborhoods of vertices. A shattered set here is precisely a set S of vertices, as well as for every $R \subseteq S$, a special vertex v_R such that v_R is adjacent to all vertices in R and non-adjacent to all vertices in $S \setminus R$. But this is precisely[1] a bi-induced copy of a certain bipartite graph, which we denote by V_d: its two parts are indexed by $[d]$ and by $2^{[d]}$, and a vertex on the right is joined to precisely the set of vertices on the left that it indexes. Hence, if $\dim_{\mathrm{VC}}(G) < d$, meaning that there is no shattered set of size d, then G avoids V_d as a bi-induced subgraph, and hence has bounded fake VC dimension (specifically $\dim_{\mathrm{fVC}}(G) \leqslant d + 2^d$). Conversely, it is not hard to show that $V_{d+\log d}$ contains *every* bipartite graph on d vertices as a bi-induced subgraph, hence if $\dim_{\mathrm{fVC}}(G) \leqslant d$, then G avoids some such bi-induced subgraph, and hence avoids $V_{d+\log d}$, implying $\dim_{\mathrm{VC}}(G) \leqslant d + \log d$. In particular, since we are only interested in whether the VC dimension is bounded or not, the real notion is equivalent to the fake one. Note too that the graph perspective is fairly natural, since any set system $\mathcal{F} \subseteq 2^V$ can be viewed as a bipartite graph with parts V and $\mathcal{F}$ with edges for inclusion, and then $\mathcal{F}$ can be recovered as the neighborhood system of this bipartite graph[2].

Arguably the most basic and fundamental fact about VC dimension is the so-called *Sauer–Shelah lemma* which, naturally, was first proven by Vapnik and Chervonenkis.

Lemma 2.5 (Vapnik–Chervonenikis–Perles–Sauer–Shelah)**.** *Let U be a set of size k and let $\mathcal{F} \subseteq 2^U$ be a family with $\dim_{\mathrm{VC}}(\mathcal{F}) \leqslant d$. Then*

$$|\mathcal{F}| \leqslant \binom{k}{0} + \binom{k}{1} + \cdots + \binom{k}{d}.$$

[1]Actually, not quite: I am cheating a tiny bit by assuming that $v_R \notin S$.

[2]More precisely, we should only look at neighborhoods of vertices on the right, viewed as a family of subsets of the left vertices.

The bound in this result is tight, but we will only ever use the fact that it guarantees $|\mathcal{F}| \leqslant 2m^d$. This result is already quite interesting, showing that set systems of bounded VC dimension must be fairly small. However, in our kinds of applications it seems basically useless: for example, if G is an n-vertex graph, then trivially the neighborhood family $\mathcal{N}(G)$ has at most n sets in it, whether or not G has bounded VC dimension. But this is actually far from being the case, as we can very often use the Sauer–Shelah lemma even with it seems trivial, by applying it to certain new families derived from $\mathcal{F}$; we will shortly see an example of this.

The example concerns Haussler's packing lemma, which is another foundational result about VC dimension and which helps justify its "dimension" name. The setup is as follows: let us say that a family $\mathcal{F} \subseteq 2^V$ is *ε-separated* if for all distinct $F, F' \in \mathcal{F}$, we have $|F \triangle F'| \geqslant \varepsilon|V|$. If we again view the elements of $\mathcal{F}$ via their indicator vectors (so that they are elements of $\{0,1\}^n$, where $n = |V|$), then this is nothing more than the definition of a code of relative distance ε: the symmetric difference of two sets is the same as the Hamming distance between their indicator vectors, and a code of relative distance ε is precisely a collection of vectors whose pairwise Hamming distance is at least εn.

In general, we know that the number of codewords in such a code can be exponential in n. However, Haussler's packing lemma states that if we impose the added condition that $\mathcal{F}$ has bounded VC dimension, this number drops substantially, to exponential in d. This both justifies the name dimension—since it essentially states that if a family has VC dimension d, it "looks like it lives in d dimensions"—and implies that if the VC dimension is an absolute constant, then the number of such codewords is itself a constant depending on ε. Here is the formal statement.

Theorem 2.6 (Haussler's packing lemma)**.** *There is an absolute constant C such that the following holds. If $\mathcal{F} \subseteq 2^V$ is a set system of VC dimension d which is ε-separated, then $|\mathcal{F}| \leqslant (C/\varepsilon)^d$.*

The proof of the packing lemma is a bit involved, but there is a short and elegant proof of the slightly weaker bound

$$|\mathcal{F}| \leqslant \left(\frac{C \log \frac{1}{\varepsilon}}{\varepsilon}\right)^d,$$

which we will now see. We stress that this weaker bound is sufficient for all our applications, since we will only ever use that we have a $\mathrm{poly}(\varepsilon^{-1})$ bound for any constant d.

Proof of the weaker bound. For every set $U \subseteq V$, let $\mathcal{F}|_U = \{F \cap U : F \in \mathcal{F}\}$. If we think of the $\mathcal{F} \times V$ matrix again, this just boils down to restricting to the columns in U; if we think of $\mathcal{F}$ as the neighborhood set system of some bipartite graph, then this is just the induced subgraph obtained by only keeping the vertices of U on the left. As both of these perspectives make obvious, this operation cannot create new shattered sets, hence $\dim_{\mathrm{VC}}(\mathcal{F}|_U) \leqslant \dim_{\mathrm{VC}}(\mathcal{F}) \leqslant d$.

Define $k = (2 \log |\mathcal{F}|)/\varepsilon$. Now, let $u_1, \ldots, u_k$ be k independent, uniformly random elements of V, and let $U = \{u_1, \ldots, u_k\}$. Note that we allow repetitions, but we certainly have

$|U| \leqslant k$. Deterministically, we have by the Sauer–Shelah lemma that

$$|\mathcal{F}|_U| \leqslant 2|U|^d \leqslant 2k^d.$$

On the other hand, for every fixed $F \neq F' \in \mathcal{F}$, we claim that with high probability $F \cap U \neq F' \cap U$. Indeed, the event $F \cap U = F' \cap U$ is precisely the event that $u_i \notin F \triangle F'$ for all $i \in [k]$. By assumption, we know that $|F \triangle F'| \geqslant \varepsilon|V|$, hence

$$\Pr(F \cap U = F' \cap U) = \Pr(u_i \notin F \triangle F' \text{ for all } i) \leqslant (1-\varepsilon)^k \leqslant e^{-\varepsilon k} = |\mathcal{F}|^{-2},$$

where the final equality is by our choice of $k = (2\log|\mathcal{F}|)/\varepsilon$. We now apply a union bound to conclude that with positive probability, we have $F \cap U \neq F' \cap U$ for all distinct $F, F' \in \mathcal{F}$. But this precisely means that $|\mathcal{F}|_U| = |\mathcal{F}|$, from which we conclude that

$$|\mathcal{F}| = |\mathcal{F}|_U| \leqslant 2k^d = 2\left(\frac{2\log|\mathcal{F}|}{\varepsilon}\right)^d$$

which implies the claimed bound. □

Haussler's packing lemma has a number of nice applications. One of the most important is the following result, which I view as a regularity lemma for set systems of bounded VC dimension: it states that every such set system can be partitioned into $\mathrm{poly}(\varepsilon^{-1})$ parts, such that all the sets within a single part are nearly identical, in the sense that they have small symmetric difference.

Corollary 2.7. *Let $\mathcal{F} \subseteq 2^V$ be a set system with $\dim_{\mathrm{VC}}(\mathcal{F}) \leqslant d$, and let $\varepsilon > 0$. We can write $\mathcal{F} = \mathcal{F}_1 \sqcup \cdots \sqcup \mathcal{F}_m$, for some $m \leqslant (C/\varepsilon)^d$, such that for all $i \in [m]$ and all $F, F' \in \mathcal{F}_i$, we have $|F \triangle F'| \leqslant \varepsilon|V|$.*

Proof. Let $\{F_1, \ldots, F_m\}$ be a maximal $(\varepsilon/2)$-separated subfamily of $\mathcal{F}$. By the packing lemma we know that $m \leqslant (C/\varepsilon)^d$ for some constant C. Now, by maximality, we know that for every $F \in \mathcal{F}$, there must be some $i \in [m]$ such that $|F \triangle F_i| \leqslant \varepsilon/2$. Let $\mathcal{F}_i$ be the collection of all such sets; by the triangle inequality, we know that $|F \triangle F'| \leqslant \varepsilon$ for all $F, F' \in \mathcal{F}_i$. □

Using this corollary, we can finally go back to the world of graphs, and prove the regularity lemma for graphs of bounded VC dimension.

Theorem 2.8. *If G is a graph with $\dim_{\mathrm{VC}}(G) \leqslant d$, then G has an ε-homogeneous equipartition into at most $(C/\varepsilon)^{2d}$ parts, where C is an absolute constant.*

Versions of this theorem were first proved independently by Lovász–Szegedy and Alon–Fischer–Newman; the version stated above, as well as the proof given here, is due to Fox–Pach–Suk.

Proof sketch. Recall that $\mathcal{N}(G)$ denotes the neighborhood set system of G, and by assumption $\dim_{\mathrm{VC}}(\mathcal{N}(G)) \leqslant d$. Apply Corollary 2.7 to this set system (with parameter ε^2) to obtain a partition $\mathcal{N}(G) = \mathcal{F}_1 \sqcup \cdots \sqcup F_m$, for some $m \leqslant (C/\varepsilon^2)^d$. The sets in $\mathcal{N}(G)$ are in bijection with vertices of G, so this actually corresponds to a partition $V(G) = V_1 \sqcup \cdots \sqcup V_m$. The key property this partition has, coming directly from Corollary 2.7, is that for all $v, v' \in V_i$, we have $|N(v) \triangle N(v')| \leqslant \varepsilon^2 |V|$. That is, all pairs of vertices in the same part have almost exactly the same neighborhood in G.

All that remains is to prove that this partition is $O(\varepsilon)$-homogeneous, which is not hard but which I will skip. It boils down to a clever double-counting argument: if many pairs of parts had edge density far from 0 and 1, then we could average to find two vertices in the same part with rather different neighborhoods. □

Recall that ε-homogeneity is a strictly stronger notion than ε-regularity (up to a polynomial change in ε). So we have found that the assumption of bounded VC dimension allows us to strengthen Szemerédi's regularity lemma in two ways: quantitatively, the number of parts drops from tower-type in the worst case to polynomial, and qualitatively, the condition we get is the stronger property of ε-homogeneity.

It is now natural to wonder if there are intermediate statements: can we impose some weaker assumption than bounded VC dimension to obtain weaker conclusions, but still ones that give us better control than in the worst case? Rather amazingly, the answer is no in the following very strong sense, which was perhaps first noticed by Alon, Fox, and Zhao.

Theorem 2.9. *The following are equivalent for a hereditary family of graphs G.*

- *G has bounded VC dimension.*
- *G has an ε-homogeneous partition with* $\mathrm{poly}(\varepsilon^{-1})$ *many parts.*
- *G has an ε-homogeneous partition with any bounded number of parts.*
- *G has an ε-regular partition with* $\mathrm{poly}(\varepsilon^{-1})$ *many parts.*
- *G has an ε-regular partition with any number of parts that is less than* $\mathrm{twr}(\varepsilon^{-c})$.

Unfortunately, while we should really think of this as a statement about a single graph G, it doesn't really make sense for a single graph (and most ways of trying to make it make sense makes it false). So for this statement we work with hereditary families of graphs, where we recall that a family of graphs is *hereditary* if any induced subgraph of a graph in the family is also in the family. If you don't like this, there are other ways of formalizing a statement like Theorem 2.9, but I will stick with this because it's simplest.

We have already seen that the first statement implies the other four, and of course the interesting thing is the converse implications. In particular, it is quite surprising (a) that there is no intermediate growth rate for regular partitions between polynomial and tower, and (b) that the only classes of graphs which admit homogeneous partitions of any bounded size are those in which this bounded size is actually automatically a polynomial. However, despite these statements being quite surprising, they are rather easy to prove with what we already know.

Proof sketch. We already showed that the first statement implies the remaining four; we will now show that the negation of the first statement implies the negations of the others. By definition, if a hereditary family *does not* have bounded VC dimension, then it contains every bipartite graph in a bi-induced fashion. But a random bipartite graph does not have ε-homogeneous partitions of any bounded size, and the bipartite version of Gowers's example does not have ε-regular partitions of any sub-tower size, from which we deduce the negations of the other statements. □

This proof makes it seem like I am cheating, and getting a characterization simply by using rather strongly the heredity. This is mostly false (as mentioned, one can come up with similar statements avoiding hereditary families), but also a bit true. Nevertheless, we will shortly see other results, in the setting of hypergraphs, which indicate that even when working with hereditary families we can get much more complex behavior.

3 Hypergraph regularity 101

We've seen that bounded VC dimension is a natural condition on graphs that allows to both qualitatively and quantitatively strengthen Szemerédi's regularity lemma, and moreover, that this condition actually characterizes when such strengthenings are possible. Our goal in the remainder of the talk is to understand what a corresponding statement for hypergraphs should look like. Unfortunately, it turns out that this is much harder than the graph case, for a number of reasons. First, when working with k-uniform hypergraphs, there are at least two (and in fact roughly $k-1$) different notions of regularity, as well as at least k different notions of bounded VC dimension. Moreover, these different notions can be extremely hard to work with. For the moment, we will only focus on 3-uniform hypergraphs, which already capture much of the difficulty of the general case.

We can begin by "following our nose" and trying to generalize what we know for graphs. With that in mind, let's define the density of a tripartite 3-graph (X, Y, Z) to be $d(X, Y, Z) := e(X, Y, Z)/(|X||Y||Z|)$, where $e(X, Y, Z)$ is the number of edges. Then we can say that the triple (X, Y, Z) is *weakly ε-regular* if, for all $X' \subseteq X, Y' \subseteq Y, Z' \subseteq Z$ with $|X'| \geqslant \varepsilon|X|, |Y'| \geqslant \varepsilon|Y|, |Z'| \geqslant \varepsilon|Z|$, we have

$$|d(X', Y', Z') - d(X, Y, Z)| \leqslant \varepsilon.$$

While this is a very natural generalization of the graph-theoretic notion, as its name indicates, it is not really the best definition. In particular, it turns out that weak regularity is too weak to support a counting lemma (we will see a proof of this later on), which means that it is too weak for most applications. However, it is still useful for some things, and is also an interesting notion to study in its own right. Additionally, one thing that is the same as the graph setting is that it does support a regularity lemma.

Theorem 3.1 (Chung)**.** *Every 3-graph has a weakly ε-regular equipartition into at most $M(\varepsilon)$ parts.*

As you would expect, we say that an equipartition $V_1 \sqcup \cdots \sqcup V_m$ is weakly ε-regular if all but at most εm^3 triples of parts are weakly ε-regular. Chung's proof, a natural extension of Szemerédi's, provides tower-type upper bounds on $M(\varepsilon)$, and it is straightforward to embed Gowers's example into a hypergraph and show that such bounds are necessary in the worst case.

Our next goal, of course, is to come up with a notion of VC dimension which would ideally allow us to both qualitatively and quantitatively strengthen Chung's regularity lemma. Continuing our inspiration from the graph world, a natural thing to try is to come up with a generalization of the neighborhood set system for a 3-graph. There are a couple of natural ways one could try to do this. The following turns out to be a good one; one reason it is natural is that it arises naturally from geometric applications (recall that geometrically defined graphs have bounded VC dimension, and hypergraphs that are geometrically defined in the same way turn out to have bounded VC dimension under the following notion).

Definition 3.2. Given a 3-graph H and a vertex $v \in V(H)$, its *link* $L(v)$ is the set of pairs $\{x, y\}$ such that $vxy \in E(H)$. We will later think of $L(v)$ as a graph on vertex set $V(H) \setminus \{v\}$, but for the moment we just think of it as an abstract collection of pairs.

The *neighborhood set system* $\mathcal{N}(H)$ is

$$\mathcal{N}(H) \coloneqq \{L(v) : v \in V(H)\} \subseteq 2^{\binom{V(H)}{2}}.$$

The *strong VC dimension* of H is defined as $\dim_{\mathrm{sVC}}(H) \coloneqq \dim_{\mathrm{VC}}(\mathcal{N}(H))$.

Thus, we have $\dim_{\mathrm{sVC}}(H) \geqslant d$ if and only if there is a *shattered set of d pairs*: that is, d pairs $\{x_1, y_1\}, \ldots, \{x_d, y_d\}$ of vertices in H such that for all $R \subseteq [d]$, we have a vertex v_R such that $v_R x_i y_i$ is an edge of H if and only if $i \in R$.

And, just as we would have hoped, this definition leads to a qualitatively and quantitatively stronger regularity lemma, which was proved independently by Chernikov–Starchenko and Fox–Pach–Suk.

Theorem 3.3. *If H is a 3-graph with $\dim_{\mathrm{sVC}}(H) \leqslant d$, then there is an ε-homogeneous equipartition of H into at most $(C/\varepsilon)^{2d}$ parts.*

As in the graph case, we say a triple (X, Y, Z) is *ε-homogeneous* if $d(X, Y, Z) \in [0, \varepsilon] \cup [1-\varepsilon, 1]$, and a partition is ε-homogeneous if all but an ε-fraction of triples of parts are. The proof of Theorem 3.3 is essentially identical to that of Theorem 2.8: we apply Corollary 2.7 to the set system $\mathcal{N}(H)$, which gives us a partition into clusters. Since the elements of $\mathcal{N}(H)$ are in bijection with vertices of H, this gives a vertex partition of H with the property that two vertices in the same part have nearly identical links. All that remains is to again apply a double-counting argument to conclude that this partition is ε-homogeneous.

Based on our experience with the graph world, we might now think that bounded strong VC dimension is a characterization: a (hereditary family of) 3-graph H has bounded strong VC dimension if and only if it admits homogeneous partitions of any bounded size (which is then polynomial), if and only if it admits weakly regular partitions of any sub-tower size (which is again then polynomial). However, it turns out that in the setting of hypergraphs, things are more complicated: such a characterization is partially true and partially false. Let us begin with the good news.

Theorem 3.4 (Terry)**.** *The following are equivalent for a hereditary family of* 3*-graphs* H*:*

- H *has bounded strong VC dimension.*
- H *has an* ε*-homogeneous partition with* $\operatorname{poly}(\varepsilon^{-1})$ *many parts.*
- H *has a weakly* ε*-regular partition with* $\operatorname{poly}(\varepsilon^{-1})$ *many parts.*

Again, Theorem 3.3 shows that the first item implies the second, which in turn implies the third since it's not hard to show that ε-homogeneity is strictly stronger than weak ε-regularity, up to a polynomial change in ε. And unlike in the graph case, the converses here are by no means trivial.

To understand why the converses are non-trivial, let's turn to the bad news.

Theorem 3.5 (Terry)**.** *There exists a* 3*-graph* H *which admits an* ε*-homogeneous partition into* $2^{O(\varepsilon^{-1})}$ *parts, but into no smaller number of parts.*

Thus, bounded strong VC dimension *does not* characterize when we have homogeneous partitions of any bounded size. As a consequence, the proof of the converses in Theorem 3.4 cannot be as trivial as those in Theorem 2.9, and must actually use the fact that we have only polynomially many parts.

I will not prove Theorem 3.4, but I do want to sketch the proof of Theorem 3.5, as it is instructive both in demonstrating the new kinds of phenomena that arise in hypergraphs, and in motivating our next topic.

Proof sketch of Theorem 3.5. Let $k = 1/(4\varepsilon)$. Construct a tripartite 3-graph on parts A, B, C, each of size n, as follows. First, equitably partition A and B into parts $A_1 \sqcup \cdots \sqcup A_k, B_1 \sqcup \ldots B_k$. Next, for each $i \in [k]$, pick a uniformly random subset $C_i \subseteq C$. The edges of H are then defined to be $\bigcup_{i=1}^k A_i \times B_i \times C_i$.

First, notice that H has an ε-homogeneous partition of size 2^k. Indeed, we simply partition A, B into $A_1, \ldots, A_k, B_1, \ldots, B_k$, and partition C into the common refinement of the random subsets $C_1, \ldots, C_k$. That is, we look at the Venn diagram these subsets define, and partition C into the cells of these Venn diagrams. Then the resulting partition is certainly ε-homogeneous, since in fact every triple of parts in it is either complete or empty.

On the other hand, it is not too hard to show that any partition of H into fewer than $2^{k/100}$ parts will not be ε-homogeneous. The reason is that in any such partition, we will have many parts that intersect $\Omega(k)$ of the sets C_i and their complements non-trivially. Thus, for any triple of such parts, we will see essentially the same structure as the global hypergraph H, whose edge density is $1/(2k) = 2\varepsilon$. In particular, any of these triples of parts will not be ε-homogeneous, from which the result follows. $\square$

4 Easy links

The example in Theorem 3.5 shows that we need to consider a more refined notion of "low complexity" than strong VC dimension if we want to characterize the existence of homogeneous partitions of 3-graphs. To figure out what this more refined notion should be, it is instructive to think a bit more about the example from Theorem 3.5. Since it does not have an ε-homogeneous partition of polynomial size, we know from Theorem 3.3 that it must have unbounded strong VC dimension. But in fact, we can also see this directly: if we select one vertex a_i from each A_i and one vertex b_i from each B_i, then the set of pairs $\{a_1, b_1\}, \ldots, \{a_k, b_k\}$ is shattered by vertices in C, and hence $\dim_{\mathrm{sVC}}(H) \geqslant k$. This suggests one possible issue with the notion of strong VC dimension: the set of pairs we are shattering is completely unstructured, and we would get more information if we shattered a set of pairs with some "meaning".

Another thing we can note about the example in Theorem 3.5 is that despite its complicated global structure, its links are actually very simple: for every $v \in V(H)$, the graph $L(v)$ is a complete bipartite graph or a disjoint union of complete bipartite graphs. This, in turn, suggests an alternative perspective we can take on the question of "what does it mean for a 3-graph to have a simple structure?" Namely, we can say that a 3-graph is "simple" if all of its links are "simple" graphs. And in fact, this kind of idea is quite common in other areas of hypergraph theory; for example, one of the most fundamental facts in the study of high-dimensional expanders is that if all links in a hypergraph are good expanders, then the hypergraph itself must be a good expander. More generally, the set of links provides a local view of the structure of the hypergraph, and it is often possible to leverage this local information into global information.

This motivates the following definition.

Definition 4.1. The *VC_1 dimension* of a 3-graph H is defined as

$$\dim_{\mathrm{VC}_1}(H) \coloneqq \max_{v \in V(H)} \dim_{\mathrm{VC}}(L(v)).$$

That is, H has bounded VC_1 dimension if and only if every link has bounded VC dimension, when viewed as a graph.

A moment's thought reveals that this is equivalent to the following definition. H has VC_1 dimension at least d if and only if there is a *shattered star* of size d, namely a collection of vertices $x, y_1, \ldots, y_d$ such that for all $R \subseteq [d]$, there is a vertex v_R for which $v_R x y_i \in E(H)$ if and only if $i \in R$.

This definition unites the two perspectives introduced above: it gives a local, link-based definition of the complexity of H, which is in turn equivalent to a shattering condition about pairs that are not completely arbitrary. This also implies that $\dim_{\mathrm{sVC}}(H) \geqslant \dim_{\mathrm{VC}_1}(H)$, i.e. that having bounded strong VC dimension is a stronger condition than having bounded VC_1 dimension.

The notion of VC_1 dimension was introduced independently by Chernikov–Towsner and Terry–Wolf, who independently proved the following theorem.

Theorem 4.2. *If H has bounded VC$_1$ dimension, then H has an ε-homogeneous partition into at most $M_1(\varepsilon)$ parts.*

In other words, the assumption of bounded VC$_1$ dimension allows one to again qualitatively strengthen Chung's regularity lemma, obtaining a homogeneous partition rather than a weakly regular one. Unfortunately, from the quantitative perspective, things have gotten *worse*: Chernikov and Towsner's proof used infinitary techniques, and thus gave no explicit bound on $M_1(\varepsilon)$, whereas Terry and Wolf used the full hypergraph regularity lemma, which yields a bound of

$$M_1(\varepsilon) \leqslant \mathrm{wow}(\varepsilon^{-C}) \coloneqq \underbrace{2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}\Big\} 2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}\Big\}\cdots}_{\varepsilon^{-C}}.$$

Note that this is (much) worse than the tower-type bounds given by Chung's regularity lemma. However, the situation was substantially improved a few years ago, when Terry found an alternative proof of Theorem 4.2, which yielded a bound on the number of parts that is double-exponential, i.e. $M_1(\varepsilon) \leqslant 2^{2^{\varepsilon^{-C}}}$. In particular, this shows that bounded VC$_1$ dimension also yields a *quantitative* improvement over Chung's regularity lemma. Additionally, Terry's bound is close to tight, since Theorem 3.5 shows that $M_1(\varepsilon)$ cannot in general be taken to be smaller than single exponential. Terry moreover conjectured that her double-exponential upper bound is closer to the truth. As it turns out, this conjecture is false, and the lower bound in Theorem 3.5 is essentially sharp.

Theorem 4.3 (Gishboliner–Shapira–W.)**.** *If H has bounded VC$_1$ dimension, then H has an ε-homogeneous partition into at most $2^{\varepsilon^{-C}}$ parts.*

We actually proved this theorem by proving something somewhat stronger.

Theorem 4.4 (Gishboliner–Shapira–W.)**.** *Suppose that in a 3-graph H, every link $L(v)$ has an ε-homogeneous partition into at most m parts. Then H itself has an ε-homogeneous partition into at most $2^{(m/\varepsilon)^C}$ parts.*

In other words, we found a way to stitch together the local information from the links to a global partition. This result immediately implies Theorem 4.3, since, if H has bounded VC$_1$ dimension, then every link has bounded VC dimension, yielding an ε-homogeneous partition into $m \leqslant \varepsilon^{-C}$ parts by Theorem 2.8.

I won't say very much about the proof of this theorem, but let me at least mention one of the main insights. In many proofs of regularity lemmas, the thing that leads to terrible bounds is almost always the need to do "common refinements". Namely, at some point in the proof, you have found various good subsets of your vertex set, but since you need to maintain a partition of the vertex set, you take the common refinement of all good sets (this is the same as the cells of the Venn diagram we encountered before). This step is often necessary, but it has a huge cost, because in the worst case it will cost an exponential. For example, the need to do this common refinement iteratively is what leads to the tower-type bounds in Szemerédi's and Chung's regularity lemmas. Similarly, Terry obtained double-exponential bounds for $M_1(\varepsilon)$ by essentially first classifying vertices into exponentially many types, and later taking a common refinement of certain sets arising from this classification. In our proof, we work with a very non-standard type of partition, which is unnatural to think about but whose manipulations only cost polynomial losses. Then, only at the final step of the proof do we take a common refinement to return to a true partition; the fact that we only do this once is why our bound is single exponential. Moreover, the single-exponential lower bound in Theorem 3.5 essentially implies that doing one such step is unavoidable, so our proof is in some sense as efficient as possible. We will later return to these notions of non-standard partitions and minimizing the use of common refinements.

We have mostly been focusing on homogeneous partitions recently, but it is natural to ask whether there is a version of Theorem 4.4 for regular partitions. That is, if we know that the link of every vertex of H has a small ε-regular partition, does that imply that H itself has a small weakly ε-regular partition? Somewhat surprisingly, we also proved that the answer is no in a very strong sense: there is a 3-graph H all of whose links have an ε-regular partition with $\mathrm{poly}(\varepsilon^{-1})$ number of parts, but every weakly ε-regular partition of H requires at least a tower-type number of parts. Since Chung's regularity lemma states that *every* 3-graph has a weakly regular partition with a tower-type number of parts, this result indicates that regular partitions of the links are essentially useless! I find this statement quite surprising, and it demonstrates that there is a big difference between the notions of homogeneity and (weak) regularity.

Before moving on to our final topic, let me just mention that bounded VC_1 dimension does indeed give us the characterization that we were looking for before. This result was proved by Terry (building on earlier results of Chernikov–Towsner and Terry–Wolf), but the version I state below includes the optimal single-exponential bound arising from Theorem 4.3.

Theorem 4.5. *The following are equivalent for a hereditary class of 3-graphs $\mathcal{H}$.*

- *$\mathcal{H}$ has bounded VC_1 dimension.*
- *$\mathcal{H}$ has an ε-homogeneous partition with at most $2^{\mathrm{poly}(\varepsilon^{-1})}$ parts.*
- *$\mathcal{H}$ has an ε-homogeneous partition with any bounded number of parts.*
- *$\mathcal{H}$ has a weakly ε-regular partition with at most $2^{\mathrm{poly}(\varepsilon^{-1})}$ parts.*
- *$\mathcal{H}$ has a weakly ε-regular partition with any sub-tower number of parts.*

5 Hypergraph regularity 201

The very first low-complexity notion we saw in this talk was what I called fake VC dimension: a graph has bounded fake VC dimension if it forbids some fixed bipartite B as a bi-induced subgraph. From this perspective, there is an obvious extension to hypergraphs, which is perhaps the most combinatorially natural of all definitions we have considered. Namely, let us say a 3-graph H has *bounded VC_2 dimension* if there is some fixed tripartite 3-graph T

which is not a tri-induced subgraph of H. Here, a tripartite 3-graph is *tri-induced* if all edges between the 3 parts appear, all non-edges between the 3 parts do not appear, and we don't care about any other edges (i.e. entirely within one part, or between two parts).

This is the "fake" definition, which is arguably the most convenient to think about. The real definition, originally due to Shelah, and motivated by certain concepts in model theory, is the following.

Definition 5.1. A *shattered* $K_{d,d}$ in a 3-graph is a collection of vertices $x_1, \dots, x_d, y_1, \dots, y_d$ such that for every $R \subseteq [d] \times [d]$, there is a vertex v_R such that $v_R x_i y_j \in E(H)$ if and only if $(i, j) \in R$. The *VC_2 dimension* of H is the maximum d such that there is a shattered $K_{d,d}$ in H.

Equivalently, we can define the tripartite 3-graph $V_{d\times d}$ whose vertex set consists of three parts labeled by $[d], [d], 2^{[d]^2}$ respectively, and whose edges are all triples whose first two coordinates are in the set indexed by the third. Then the VC_2 dimension of H is the maximum d such that $V_{d\times d}$ is a tri-induced subgraph of H.

It is again not hard to check that every tripartite 3-graph is a tri-induced subgraph of $V_{d\times d}$ for a sufficiently large d, hence the "real" and "fake" notions are equivalent if we only care about whether the VC_2 dimension is bounded.

There are a wide array of rich examples of hypergraphs of bounded VC_2 dimension, and this notion does not, at first glance, seem like a particularly restrictive one. For example, the following simple construction can be used to construct many hypergraphs of bounded VC_2 dimension, and to disprove many properties we might have hoped are enjoyed by such hypergraphs.

Proposition 5.2. *Let G be any tripartite[3] graph. Let $H = \Delta(G)$ be the hypergraph whose edges are all triangles in G. Then H has bounded VC_2 dimension (in fact, $\dim_{\mathrm{VC}_2}(H) \leqslant 1$).*

Proof. We claim that there is no tri-induced copy of $K_{2,2,2} \setminus e$. Indeed, fixing any six vertices, if we find that seven of the eight triangles they span are triangles in G, then we can conclude that all edges they span are in G. But in that case also the eighth triangle is present in G, proving the claim. By the exact same logic, it is easy to verify that there is no shattered $K_{2,2}$ in H, implying the more precise bound $\dim_{\mathrm{VC}_2}(H) \leqslant 1$. □

Here are some immediate corollaries of this simple construction.

- 3-graphs of bounded VC_2 dimension may not admit homogeneous partitions of any bounded size.

  Proof: If G is a random tripartite graph, then $H = \Delta(G)$ has bounded VC_2 dimension, but in any partition into a bounded number of parts we will see the same edge density among all the parts as globally. Since H has edge density bounded away from 0 and 1, it has no bounded homogeneous partition.

[3] This is actually completely immaterial, and everything I'll say holds for general graphs. But it is helpful to focus on the tripartite setting.

- 3-graphs of bounded VC_2 dimension may require tower-type bounds in Chung's regularity lemma.

  Proof: Let the three bipartite graphs comprising G be Gowers's example. It is easy to check that any weakly ε-regular partition of $H = \Delta(G)$ corresponds to an ε-regular partition of G, which must have tower-type many parts.

- There is no analogue of Haussler's packing lemma for VC_2 dimension.

  Proof: Again, let G be random tripartite on parts A, B, C, and let $H = \Delta(G)$, so H has bounded VC_2 dimension. The kind of analogue of Haussler's packing lemma we would want is that we can partition the vertices of C into $O_\varepsilon(1)$ parts such that any pair of vertices in the same part have nearly identical links. But this is impossible: with high probability, for all distinct $c, c' \in C$ we have $|L(c) \triangle L(c')| \geqslant |A||B|/10$. So any partition as in Corollary 2.7 would need to put each vertex of C into a distinct part, and in particular the number of parts cannot be a constant.

- Weak regularity is too weak to support an induced counting lemma.

  Proof: Again, let G be random tripartite and $H = \Delta(G)$. Then it is easy to check that H is weakly ε-regular for any fixed ε, but we already saw that H does not contain a tri-induced copy of $K_{2,2,2} \setminus e$. Hence there is no analogue of the induced counting lemma for weak regularity.

The third point suggests that if we want to prove any sort of strengthened regularity lemma for VC_2 dimension, we would need a completely different argument: we cannot go through Corollary 2.7. However, the fourth point should make us rethink what it even is that we want to prove: is there a notion of hypergraph regularity that supports an induced counting lemma?

The answer is yes, and the example showing why weak regularity is not enough also suggests what we should be trying. Namely, the "reason" why weak regularity is not sufficient is that it does not see that the hypergraph $H = \Delta(G)$ "is defined by" an underlying graph G. It is the presence of this graph G that causes the counting lemma to fail. But of course, any regularity lemma for hypergraphs will not a priori "know about" the relevant lower-uniformity information, so whatever kind of structure it outputs should also output such information.

With that in mind, let me try to state the hypergraph regularity lemma. Precise versions were developed in the early 2000s by Gowers and by Rödl et al., and every precise version is far too complicated for me to state exactly. But I will try to give you a sense of what the statement looks like.

Definition 5.3. A *hypergraph regularity partition* of H consists of a partition $V(H) = V_1 \sqcup \cdots \sqcup V_m$, as well as, for each $i \neq j$, a partition of $V_i \times V_j$ into bipartite graphs $G_{ij}^1, \dots, G_{ij}^t$. We further require that for almost all triples, the edges of H "live ε-regularly" within $\Delta(G_{ij}^a \cup G_{jk}^b \cup G_{ik}^c)$, and morevoer each graph G_{ij}^a is itself extremely regular.

I will not specify what the key definition of "lives ε-regularly within" means, but hopefully you can get a picture of what kind of statement it should be: the edges of H should form a random-looking and well-distributed subset of the set of triangles $\Delta(G_{ij}^a \cup G_{jk}^b \cup G_{ik}^c)$.

For me, the least intuitive condition is the last one, but it makes sense why we would need to assume some quasirandomness about the graphs as well if our goal is to prove a counting lemma. However, one very subtle and important point is the fact that we need the graphs in the regularity partition to be extremely regular. The reason has to do with what we saw when we first defined ε-regularity: it is essentially meaningless unless ε is much smaller than the density of the pair we are considering. Since an average graph G_{ij}^a has edge density $1/t$, the amount of regularity we will require of our graphs is $f(1/t)$-regularity, for some function f that we can choose. In most applications, including the counting lemma, it suffices to take f a polynomial, but in principle it can be anything. And, as you might expect, this notion of hypergraph regularity is indeed strong enough to support an induced counting lemma.

With these definitions out of the way, we can finally state the hypergraph regularity lemma.

Theorem 5.4 (Hypergraph regularity lemma)**.** *Every 3-graph has an (ε, f) hypergraph regularity partition with at most $M_f(\varepsilon)$ parts.*

Once you know the basic facts of hypergraph regularity theory, it is actually quite straightforward to obtain a qualitative strengthening for 3-graphs of bounded VC_2 dimension. This was done independently by Chernikov–Towsner and Terry–Wolf.

Theorem 5.5. *If H has bounded VC_2 dimension, then it has a hypergraph regularity partition in which, for almost all triples, $E(H)$ lies ε-homogeneously within $\Delta(G_{ij}^a \cup G_{jk}^b \cup G_{ik}^c)$.*

The proof of this theorem is basically identical to that of Theorem 2.3: you apply the full regularity lemma, and then notice that, by the induced counting lemma, every regular triple must actually have edge density very close to 0 or 1, for otherwise we could embed any fixed tripartite 3-graph as a tri-induced subgraph.

While the proof is simple, this result actually has a really significant meaning. It says that, up to some small noise, every 3-graph of bounded VC_2 dimension looks like the union of a constant number of 3-graphs of the form $\Delta(G)$, for some collection of graphs G. In other words, the versatile example we saw before is essentially the *only* example! Said differently, if a 3-graph has bounded VC_2 dimension, it is "not really 3-uniform": all of the information in it is determined by a lower-uniformity structure, i.e. a graph.

We now see how to qualitatively strengthen the hypergraph regularity lemma, but what about the quantitative aspects? The proof of the hypergraph regularity lemma proceeds in phases, and in every phase one has to apply Szemerédi's regularity lemma. As a consequence, the bounds it gives are of wowzer type, i.e. $M_f(\varepsilon) \leqslant \mathrm{wow}(\mathrm{poly}(\varepsilon^{-1}))$. In fact, this is only true if f is some sort of "sane" function, e.g. a polynomial, or quadruple-exponential, or really anything that is of sub-tower type; in all of these cases, the contribution of f is completely swallowed by the rest of the losses. However, if f is absolutely insane (e.g. if f is itself of wowzer type), then the bounds on $M_f(\varepsilon)$ will be even worse, namely one level higher on the Ackermann hierarchy than f itself is. Moreover, Moshkovitz and Shapira proved that such bounds are necessary in general: even if f is a polynomial, there are 3-graphs which do not admit hypergraph regularity partitions with fewer than wowzer-type many parts.

As you might expect, we should be able to improve this dependency under the assumption of bounded VC_2 dimension. And indeed, that is what the next theorem states.

Theorem 5.6 (Gishboliner–Shapira–W.)**.** *If H has bounded VC_2 dimension and f is sane, then H has an (ε, f) hypergraph regularity partition with only* $\mathrm{twr}(\mathrm{twr}(\varepsilon^{-C}))$ *many parts.*

The fact that we get double-tower behavior is probably an artifact of the proof, and the reality is probably single-tower; however, in any case, this is a much better bound than the wowzer-type bounds that occur in the worst case. Moreover, this is nearly best possible: we already saw that hypergraphs of bounded VC_2 dimension might require tower-type bounds even in Chung's regularity lemma, so they certainly also require tower-type bounds in the full regularity lemma.

Again, these results imply that 3-graphs of bounded VC_2 dimension "look like" graphs. Even when it comes to the quantitative aspects of the regularity lemma, the bounds are of tower type—as they are in graphs—rather than of wowzer type, which is the truth for general 3-graphs.

Let me also mention a result of Terry's, which offers an interesting complement to ours. She showed that if f is allowed to be extremely quickly-growing—namely of wowzer type—then wowzer-type bounds are again required in the hypergraph regularity lemma even when the hypergraph has bounded VC_2 dimension. While this is by no means easy to prove, it is also not quite surprising: if f is so insane, then the *upper* bounds become worse than wowzer, so it is reasonable that the lower bounds should also deteriorate. In fact, our proof of Theorem 5.6 shows that the assumption of bounded VC_2 dimension allows one to improve the upper bound by one level of the Ackermann hierarchy, regardless of the specific function f.

6 Higher uniformities

I now want to briefly discuss what happens in higher uniformities. As you might imagine, a hypergraph regularity partition of a k-graph consists of a partition of the vertices, the pairs of vertices, the triples of vertices, ..., up to the $(k-1)$-tuples of vertices, such that each level sits quasirandomly within the level below it. The proof of the regularity lemma for k-graphs iterates the $(k-1)$-uniform regularity lemma, and hence the bounds provided are on the kth level of the Ackermann hierarchy; moreover, Moshkovitz and Shapira proved that such bounds are necessary in general.

We saw three notions of VC dimension for 3-graphs, and in general, in uniformity k, there are k such notions. First, we say that a k-graph H has bounded VC_{k-1} dimension if it forbids some fixed k-partite k-graph as a k-induced subgraph. Next, we say that it has bounded VC_ℓ dimension, for some $\ell < k-1$, if for all $(k-\ell-1)$-tuples of vertices, their joint link (which is an $(\ell+1)$-graph) has bounded VC_ℓ dimension. Chernikov and Towsner proved the following qualitative strengthening of the hypergraph regularity lemma.

Theorem 6.1. *If H is a k-graph with bounded VC_ℓ dimension, then it has a hypergraph regularity partition which is homogeneous at uniformity larger than ℓ.*

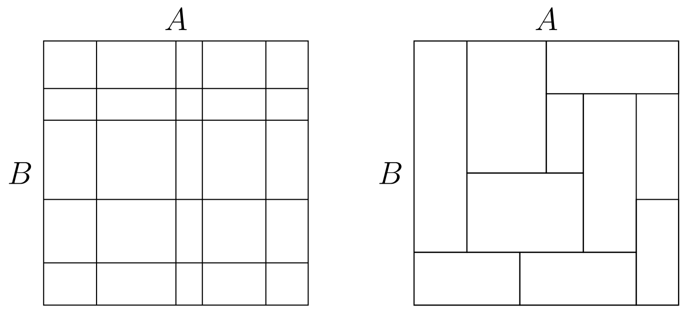
In other words, all information contained in the hypergraph regularity partition is trivial above uniformity ℓ. Again, the right way to think about this result is that k-graphs of bounded VC_ℓ dimension "look like ℓ-graphs": they can be approximated, up to small noise, by a constant-sized description that only uses information about ℓ-tuples of vertices. Given this, you might expect that also the quantitative aspects should behave as they do for ℓ-graphs, and this is indeed the case.

Theorem 6.2 (Gishboliner–Shapira–W.)**.** *If a k-graph has bounded VC_ℓ dimension, then it has a hypergraph regularity partition which is homogeneous at uniformity larger than ℓ, and whose number of parts is on the ℓth level of the Ackermann hierarchy. Moreover, this bound is essentially best possible.*

7 Proof sketch

At the end of the talk, I want to try to sketch what goes into the proof of our main theorems. Rather than proving anything for hypergraphs, I will return to graphs, and prove a weaker version of Theorem 2.8, the regularity lemma for graphs of bounded VC dimension. Namely, I will show that any such graph has an ε-homogeneous partition with only double-exponentially many parts. This is obviously much worse than the polynomial bound we already saw, but the advantage of the proof I will present is that it does not use Haussler's packing lemma, or indeed almost any special property of VC dimension. As a consequence, this proof will extend more or less readily to the higher-uniformity case, despite the fact that the packing lemma simply fails to generalize. For simplicity, let us work with a bipartite graph G (with parts A, B), although this restriction is by no means essential. We assume that G has bounded VC dimension, and our goal is to partition each of A and B into parts so that almost all pairs are ε-homogeneous.

As an intermediate goal, we will try to find the following weaker structure, which is precisely the kind of non-standard partition I mentioned earlier. Specifically, suppose we are able to find sets $A_1, \dots, A_m \subseteq A, B_1, \dots, B_m \subseteq B$ with the following properties: we have $A \times B = \bigsqcup_{i=1}^m A_i \times B_i$, and for almost every i, the bipartite graph $G[A_i, B_i]$ is ε-homogeneous. This is close to what we want, except that the A_i's and B_j's don't have to be disjoint, hence they do not necessarily form a partition of A and B; all we know is that their products partition $A \times B$, i.e. that every pair of vertices $a \in A, b \in B$ lies in exactly one choice of $A_i \times B_i$. Pictorially, the difference between a genuine partition and this kind of non-standard partition is the following.



Standard partition Non-standard partition

However, if we find such a non-standard partition, we can always turn it into a standard one, by taking the common refinement; this boils down to extending every vertical and horizontal line segment in the picture above in both directions. If all we knew is that the pairs (A_i, B_i) are ε-regular, then taking this common refinement would lose us all the information. However, if instead these pieces are ε-homogeneous, then it is not hard to show that the common refinement (or, indeed, *any* refinement) is $\sqrt{\varepsilon}$-homogeneous. This is nothing more than the statement that if you have a bunch of objects, the overwhelming majority of which are green, then no matter how you split them into buckets, most buckets will have most of their contents green. The formal proof is a one-line application of Markov's inequality.

The upshot is that, so long as we are willing to tolerate an exponential loss by taking the common refinement, we are happy if we can find a non-standard partition in which almost all the pairs (A_i, B_i) are ε-homogeneous. And once you set yourself this goal, you will readily discover that there are off-the-shelf tools for accomplishing this. Concretely, we can use the cylinder regularity lemma of Duke, Lefmann, and Rödl.

Theorem 7.1 (Cylinder regularity lemma)**.** *For every bipartite graph G on parts A, B, there is a non-standard partition $A \times B = \bigsqcup_{i=1}^{m} A_i \times B_i$ such that for all but εm choices of $i \in [m]$, the pair (A_i, B_i) is ε-regular. Moreover, we can take $m \leqslant 2^{\varepsilon^{-6}}$.*

Crucially, while this lemma returns a weaker structure than what is given by Szemerédi's regularity lemma, the bounds are much better: single exponential instead of tower-type. And once we have this tool, we are in business: we can again apply Lemma 2.2 to the non-standard partition to conclude that every ε-regular pair is actually ε-homogeneous, for otherwise we could apply the induced counting lemma inside this pair and contradict the assumption of bounded VC dimension.

Putting this all together, we obtain a proof that every bipartite graph of bounded VC dimension has an ε-homogeneous partition with at most a double-exponential number of parts: we pay one exponential to apply the cylinder regularity lemma, conclude that almost all pairs are homogeneous, pay another exponential to take the common refinement, and then apply Markov's inequality to show that the resulting partition remains homogeneous.

Essentially the same proof works for proving that k-graphs of bounded VC_k dimension have hypergraph regularity partitions whose number of parts is only on the $(k-1)$th level of the Ackermann hierarchy. The "only" thing one needs to prove is a hypergraph version of the cylinder regularity lemma, which does indeed save one level of the Ackermann hierarchy. This is very technically annoying but not very conceptually difficult, once one knows the proof of the usual hypergraph regularity lemma.

The much harder thing is to extend this to k-graphs of bounded VC_ℓ dimension. At a high level, we want to only apply the $(\ell + 1)$-uniform cylinder regularity lemma, but this doesn't really make sense, since we are presented with a k-graph. So we have to do a lot of acrobatics, including working with even more non-standard partitions—essentially working at multiple uniformities simultaneously—in order to prove the most general result.