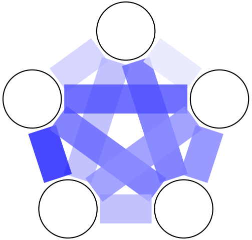
Efficient arithmetic regularity and removal lemmas for induced bipartite patterns

Yufei Zhao (MIT)

Joint work with Noga Alon (Princeton) and Jacob Fox (Stanford)

April 22, 2018

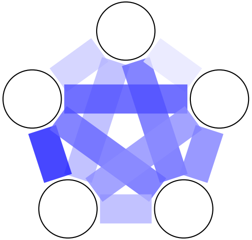
Szemerédi's graph regularity lemma



Graph regularity lemma

For every $\epsilon > 0$ there exists $M = M(\epsilon)$ so that every graph has a vertex partition into $\leq M$ parts so that all but $< \epsilon$ fraction of pairs are ϵ-regular

Szemerédi's graph regularity lemma



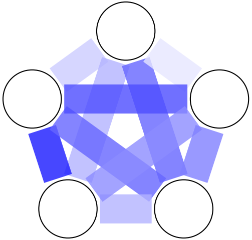
Graph regularity lemma

For every $\epsilon > 0$ there exists $M = M(\epsilon)$ so that every graph has a vertex partition into $\leq M$ parts so that all but $< \epsilon$ fraction of pairs are ϵ-regular

Graph removal lemma

For every $\forall\epsilon > 0$ and graph H there is some $\delta = \delta(H, \epsilon) > 0$ so that every n-vertex graph with H-density $< \delta$ can be made H-free by removing $< \epsilon n^2$ edges

Szemerédi's graph regularity lemma



Graph regularity lemma

For every $\epsilon > 0$ there exists $M = M(\epsilon)$ so that every graph has a vertex partition into $\leq M$ parts so that all but $< \epsilon$ fraction of pairs are ϵ-regular

Graph removal lemma

For every $\forall \epsilon > 0$ and graph H there is some $\delta = \delta(H, \epsilon) > 0$ so that every n-vertex graph with H-density $< \delta$ can be made H-free by removing $< \epsilon n^2$ edges

- $M(\epsilon) = 2^{2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}}$ tower of height $\epsilon^{-O(1)}$ (cannot be improved [Gowers])
- Removal lemma holds with $\delta = M^{-O(1)} = 1/2^{2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}}$
(possibly could be improved, but not beyond $\epsilon^{C \log(1/\epsilon)}$ when $H = K_3$)

When can you guarantee poly($1/\epsilon$) bounds?

When can you guarantee poly($1/\epsilon$) bounds?

For a graph with bounded VC dimension:

- Vertices can be partitioned into $\epsilon^{-O(1)}$ parts
- All but ϵ-fraction of pairs of vertex parts have densities $\leq \epsilon$ or $\geq 1 - \epsilon$

[Alon–Fischer–Newman, Lovász–Szegedy]

What is VC dimension?

Let $\mathcal{S}$ be a collection of subsets of Ω

$$\dim_{\mathrm{VC}}\mathcal{S} := \text{size of the largest shattered subset of } \Omega$$

$U \subset \Omega$ is shattered if for every $U' \subseteq U$ there exists $T \in \mathcal{S}$ such that $T \cap U = U'$

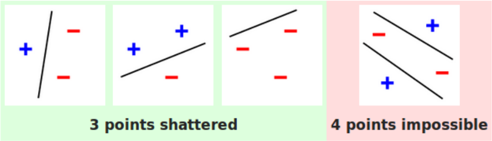
What is VC dimension?

Let $\mathcal{S}$ be a collection of subsets of Ω

$$\dim_{\mathrm{VC}}\mathcal{S} := \text{size of the largest shattered subset of } \Omega$$

$U \subset \Omega$ is shattered if for every $U' \subseteq U$ there exists $T \in \mathcal{S}$ such that $T \cap U = U'$

E.g., the VC-dimension of the collection of half-planes in $\mathbb{R}^2$ is 3



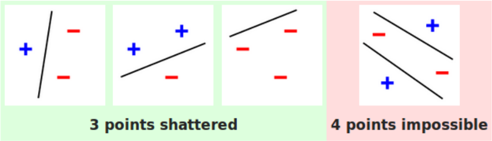
What is VC dimension?

Let $\mathcal{S}$ be a collection of subsets of Ω

$$\dim_{\mathrm{VC}} \mathcal{S} := \text{size of the largest shattered subset of } \Omega$$

$U \subset \Omega$ is shattered if for every $U' \subseteq U$ there exists $T \in \mathcal{S}$ such that $T \cap U = U'$

E.g., the VC-dimension of the collection of half-planes in $\mathbb{R}^2$ is 3



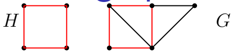
VC dimension of a graph G is defined to be the VC dimension of the collection of vertex neighborhoods ($\Omega = V(G)$):

$$\dim_{\mathrm{VC}} G := \dim_{\mathrm{VC}}\{N(v) : v \in V(G)\}$$

Bounded VC dimension $\iff$ forbidding a bi-induced subgraph

Bounded VC dimension $\iff$ forbidding a bi-induced subgraph

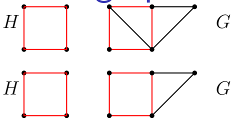
H as a subgraph of G (all edges of H are present in G)



Bounded VC dimension $\iff$ forbidding a bi-induced subgraph

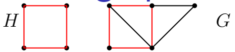
H as a subgraph of G (all edges of H are present in G)

H as an induced subgraph of G (all edges of H are present in G and non-edges are not present)

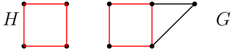


Bounded VC dimension $\iff$ forbidding a bi-induced subgraph

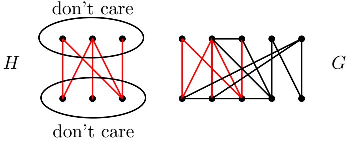
H as a subgraph of G (all edges of H are present in G)



H as an induced subgraph of G (all edges of H are present in G and non-edges are not present)



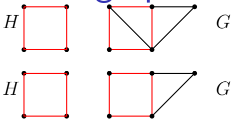
Bipartite H as a bi-induced subgraph (similar to induced but don't care about edges inside each bipartition)



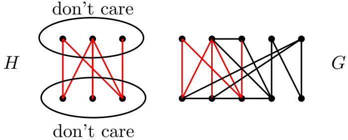
Bounded VC dimension $\iff$ forbidding a bi-induced subgraph

H as a subgraph of G (all edges of H are present in G)

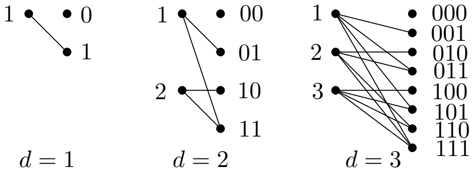
H as an induced subgraph of G (all edges of H are present in G and non-edges are not present)



Bipartite H as a bi-induced subgraph (similar to induced but don't care about edges inside each bipartition)

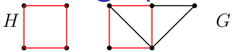


$\dim_{\mathrm{VC}} G < d \iff G$ forbids the following as a bi-induced subgraph:

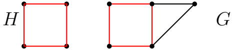


Bounded VC dimension $\iff$ forbidding a bi-induced subgraph

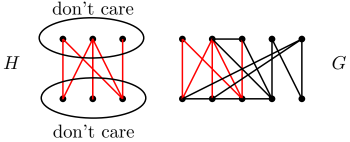
H as a subgraph of G (all edges of H are present in G)



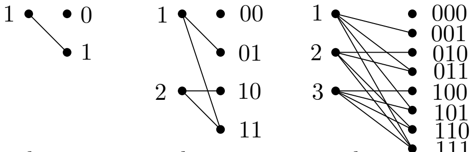
H as an induced subgraph of G (all edges of H are present in G and non-edges are not present)



Bipartite H as a bi-induced subgraph (similar to induced but don't care about edges inside each bipartition)



$\dim_{\mathrm{VC}} G < d \iff G$ forbids the following as a bi-induced subgraph:



$d = 1$ $d = 2$ $d = 3$

Conversely, if G is bi-induced-H-free, then $\dim_{\mathrm{VC}} G = O_H(1)$

When can you guarantee poly($1/\epsilon$) bounds?

Hereditary family – any family of graphs closed under deletion of vertices.

- E.g., 3-colorable, planar, bipartite, triangle-free, chordal, perfect
- Equivalent to being induced-H-free for all H in some (possibly infinite) family $\mathcal{H}$

When can you guarantee poly($1/\epsilon$) bounds?

Hereditary family – any family of graphs closed under deletion of vertices.

- E.g., 3-colorable, planar, bipartite, triangle-free, chordal, perfect
- Equivalent to being induced-H-free for all H in some (possibly infinite) family $\mathcal{H}$

For any given hereditary family $\mathcal{F}$ for graphs:

- If graphs in $\mathcal{F}$ have bounded VC-dimension, then every graph has an ϵ-regular partition into $\epsilon^{-O(1)}$ parts.

When can you guarantee poly($1/\epsilon$) bounds?

Hereditary family – any family of graphs closed under deletion of vertices.

- E.g., 3-colorable, planar, bipartite, triangle-free, chordal, perfect
- Equivalent to being induced-H-free for all H in some (possibly infinite) family $\mathcal{H}$

For any given hereditary family $\mathcal{F}$ for graphs:

- If graphs in $\mathcal{F}$ have bounded VC-dimension, then every graph has an ϵ-regular partition into $\epsilon^{-O(1)}$ parts.
- If graphs in $\mathcal{F}$ do not have bounded VC-dimension, then there exist graphs in $\mathcal{F}$ whose ϵ-regular partition whose number of parts is necessarily at least $2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}$ (tower height ϵ^{-c})

When can you guarantee $\text{poly}(1/\epsilon)$ bounds?

Regularity lemma for graphs of bounded VC dimension

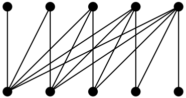
For a fixed bipartite H, if G is bi-induced-H-free, then G has a vertex partition into $\epsilon^{-O(1)}$ parts so that all but $\leq \epsilon$-fraction of pairs have edge-densities $\leq \epsilon$ or $\geq 1 - \epsilon$.

When can you guarantee $\mathrm{poly}(1/\epsilon)$ bounds?

Regularity lemma for graphs of bounded VC dimension

For a fixed bipartite H, if G is bi-induced-H-free, then G has a vertex partition into $\epsilon^{-O(1)}$ parts so that all but $\leq \epsilon$-fraction of pairs have edge-densities $\leq \epsilon$ or $\geq 1 - \epsilon$.

A graph is *k-stable* if it does not contain a bi-induced half-graph on $2k$ vertices.

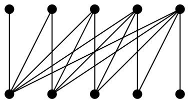


When can you guarantee $\mathrm{poly}(1/\epsilon)$ bounds?

Regularity lemma for graphs of bounded VC dimension

For a fixed bipartite H, if G is bi-induced-H-free, then G has a vertex partition into $\epsilon^{-O(1)}$ parts so that all but $\leq \epsilon$-fraction of pairs have edge-densities $\leq \epsilon$ or $\geq 1 - \epsilon$.

A graph is k-stable if it does not contain a bi-induced half-graph on $2k$ vertices.



Stable regularity lemma [Malliaris–Shelah]

If the graph is k-stable, then we can furthermore guarantee that **every** pair of parts has density $\leq \epsilon$ or $\geq 1 - \epsilon$.

Arithmetic setting

G abelian group, $A \subset G$

$$\dim_{\mathrm{VC}} A := \dim_{\mathrm{VC}} \{A + x : x \in G\} = \dim_{\mathrm{VC}} \mathrm{CayleyGraph}(G, A)$$

Arithmetic setting

G abelian group, $A \subset G$

$$\dim_{\mathrm{VC}} A := \dim_{\mathrm{VC}} \{A + x : x \in G\} = \dim_{\mathrm{VC}} \mathrm{CayleyGraph}(G, A)$$

We say that A contains a bi-induced copy of a bipartite graph H if the same is true for $\mathrm{CayleyGraph}(G, A)$

Arithmetic regularity lemma

- **Szemerédi's graph regularity lemma**: $\forall \epsilon > 0 \ \exists M$: every graph has a vertex partition into $\le M$ parts so that all but $< \epsilon$ fraction of pairs are ϵ-regular

Arithmetic regularity lemma

- **Szemerédi's graph regularity lemma**: $\forall \epsilon > 0 \; \exists M$: every graph has a vertex partition into $\leq M$ parts so that all but $< \epsilon$ fraction of pairs are ϵ-regular
- **Arithmetic regularity lemma** [Green]: $\forall \epsilon > 0, p \; \exists M$: for every $A \subset \mathbb{F}_p^n$, there is some $V \leq \mathbb{F}_2^n$ of codimension $\leq M$ such that all but $\leq \epsilon$ fraction of translates of V meet A in an ϵ-Fourier uniform way

Arithmetic regularity lemma

- **Szemerédi's graph regularity lemma**: $\forall \epsilon > 0\ \exists M$: every graph has a vertex partition into $\leq M$ parts so that all but $< \epsilon$ fraction of pairs are ϵ-regular
- **Arithmetic regularity lemma** [Green]: $\forall \epsilon > 0, p\ \exists M$: for every $A \subset \mathbb{F}_p^n$, there is some $V \leq \mathbb{F}_2^n$ of codimension $\leq M$ such that all but $\leq \epsilon$ fraction of translates of V meet A in an ϵ-Fourier uniform way
- Corollary: removal lemma for arithmetic patterns

Arithmetic regularity lemma

- **Szemerédi's graph regularity lemma**: $\forall \epsilon > 0\ \exists M$: every graph has a vertex partition into $\leq M$ parts so that all but $< \epsilon$ fraction of pairs are ϵ-regular
- **Arithmetic regularity lemma** [Green]: $\forall \epsilon > 0, p\ \exists M$: for every $A \subset \mathbb{F}_p^n$, there is some $V \leq \mathbb{F}_2^n$ of codimension $\leq M$ such that all but $\leq \epsilon$ fraction of translates of V meet A in an ϵ-Fourier uniform way
- Corollary: removal lemma for arithmetic patterns (later shown to follow from graph removal lemma [Král'–Serra–Vena / Shapira])

Arithmetic regularity lemma

- **Szemerédi's graph regularity lemma**: $\forall \epsilon > 0\ \exists M$: every graph has a vertex partition into $\leq M$ parts so that all but $< \epsilon$ fraction of pairs are ϵ-regular
- **Arithmetic regularity lemma** [Green]: $\forall \epsilon > 0, p\ \exists M$: for every $A \subset \mathbb{F}_p^n$, there is some $V \leq \mathbb{F}_2^n$ of codimension $\leq M$ such that all but $\leq \epsilon$ fraction of translates of V meet A in an ϵ-Fourier uniform way
- Corollary: removal lemma for arithmetic patterns (later shown to follow from graph removal lemma [Král'–Serra–Vena / Shapira])
- optimal M is $2^{2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}}$ tower of height $\epsilon^{-O(1)}$

Regularity lemmas with constraints

Graph regularity:

- Bounded VC-dimension (equiv. forbidding a bi-induced subgraph): a vertex partition into $\leq \epsilon^{-O(1)}$ parts so that all but $\leq \epsilon$-fraction of pairs of parts have densities $\leq \epsilon$ or $\geq 1 - \epsilon$
- Stability (i.e., forbidding a fixed-size half-graph): furthermore **every** pair of parts has density $\leq \epsilon$ or $\geq 1 - \epsilon$

Regularity lemmas with constraints

Graph regularity:

- Bounded VC-dimension (equiv. forbidding a bi-induced subgraph): a vertex partition into $\leq \epsilon^{-O(1)}$ parts so that all but $\leq \epsilon$-fraction of pairs of parts have densities $\leq \epsilon$ or $\geq 1-\epsilon$
- Stability (i.e., forbidding a fixed-size half-graph): furthermore **every** pair of parts has density $\leq \epsilon$ or $\geq 1-\epsilon$

Arithmetic regularity: $A \subset G = \mathbb{F}_p^n$, p fixed

- Stability [Terry–Wolf]: there exists a subspace $H \leq G$ with $[G:H] \leq e^{\epsilon^{-O(1)}}$ such that for all $x \in G$,

$$|A \cap (H+x)| \leq \epsilon |H| \text{ or } \geq (1-\epsilon)|H| \qquad (\dagger)$$

Regularity lemmas with constraints

Graph regularity:

- Bounded VC-dimension (equiv. forbidding a bi-induced subgraph): a vertex partition into $\le \epsilon^{-O(1)}$ parts so that all but $\le \epsilon$-fraction of pairs of parts have densities $\le \epsilon$ or $\ge 1 - \epsilon$
- Stability (i.e., forbidding a fixed-size half-graph): furthermore **every** pair of parts has density $\le \epsilon$ or $\ge 1 - \epsilon$

Arithmetic regularity: $A \subset G = \mathbb{F}_p^n$, p fixed

- Stability [Terry–Wolf]: there exists a subspace $H \le G$ with $[G : H] \le e^{\epsilon^{-O(1)}}$ such that for all $x \in G$,

$$|A \cap (H + x)| \le \epsilon |H| \text{ or } \ge (1 - \epsilon)|H| \qquad (\dagger)$$

- Bounded VC-dimension [Alon–Fox–Z.]: there exists a subspace $H \le G$ with $[G : H] \le \epsilon^{-O(1)}$ such that ($\dagger$) holds for all but an $\le \epsilon$-fraction of $x \in G$

Regularity lemmas for groups with constraints

Theorem prototype: If $A \subset G$ has (stability | bounded VC dimension), then one can find a subgroup of G with bounded index so that A has density close to 0 or 1 in (every | almost every) coset.

Regularity lemmas for groups with constraints

Theorem prototype: If $A \subset G$ has (stability | bounded VC dimension), then one can find a subgroup of G with bounded index so that A has density close to 0 or 1 in (every | almost every) coset.

$G = \mathbb{F}_p^n$:

- (Stability) [Terry–Wolf]: there exists subspace $H \leq G$ with $[G : H] \leq e^{\epsilon^{-O(1)}}$ ...
- (Bounded VC dimension) [Alon–Fox–Z.]: there exists subspace $H \leq G$ with $[G : H] \leq \epsilon^{-O(1)}$ ...

Regularity lemmas for groups with constraints

Theorem prototype: If $A \subset G$ has (stability | bounded VC dimension), then one can find a subgroup of G with bounded index so that A has density close to 0 or 1 in (every | almost every) coset.

$G = \mathbb{F}_p^n$:

- (Stability) [Terry–Wolf]: there exists subspace $H \leq G$ with $[G : H] \leq e^{\epsilon^{-O(1)}}$ ...
- (Bounded VC dimension) [Alon–Fox–Z.]: there exists subspace $H \leq G$ with $[G : H] \leq \epsilon^{-O(1)}$ ...

General groups G: (proved via model theory; no bounds known)

- (Stability) [Conant–Pillay–Terry] there exists a normal subgroup $H \trianglelefteq G$ of bounded index ...
- (Bounded VC dimension) [Conant–Pillay–Terry] For a group G of bounded exponent, there exists a normal subgroup $H \trianglelefteq G$ of bounded index ...
  (false without bounded exponent hypothesis: e.g., interval in $\mathbb{Z}/p\mathbb{Z}$)

Applications to removal lemma

Recall the graph removal lemma: $\forall \epsilon \exists \delta$: if an n-vertex graph has $< \delta n^3$ triangles, and it can be made triangle free by removing $< \epsilon n^2$ edges.

Applications to removal lemma

Recall the graph removal lemma: $\forall\epsilon\exists\delta$: if an n-vertex graph has $< \delta n^3$ triangles, and it can be made triangle free by removing $< \epsilon n^2$ edges.

Arithmetic removal lemma for bi-induced patterns [Alon–Fox–Z.]

Fix r and bipartite graph F. Let G be a finite abelian group with exponent $\leq r$. For every $\epsilon > 0$, there exists $\delta = \epsilon^{O(|V(F)|^3)}$ such that if the bi-induced-F-density in A is $< \delta$, then A can be made bi-induced-F-free by adding/deleting $< \epsilon|G|$ elements.

Applications to removal lemma

Recall the graph removal lemma: $\forall\epsilon\exists\delta$: if an n-vertex graph has $< \delta n^3$ triangles, and it can be made triangle free by removing $< \epsilon n^2$ edges.

Arithmetic removal lemma for bi-induced patterns [Alon–Fox–Z.]

Fix r and bipartite graph F. Let G be a finite abelian group with exponent $\leq r$. For every $\epsilon > 0$, there exists $\delta = \epsilon^{O(|V(F)|^3)}$ such that if the bi-induced-F-density in A is $< \delta$, then A can be made bi-induced-F-free by adding/deleting $< \epsilon|G|$ elements.

Applications to property testing: efficient sampling algorithm to distinguish $A \subset G$ that are bi-induced-F-free from those that are far from bi-induced-F-free

Open questions

- Bounds for Conant–Pillay–Terry regularity lemmas for general groups with stability/bounded VC dimension hypotheses?
(No bounds known. Maybe $\epsilon^{-O(1)}$?)

Open questions

- Bounds for Conant–Pillay–Terry regularity lemmas for general groups with stability/bounded VC dimension hypotheses?
  (No bounds known. Maybe $\epsilon^{-O(1)}$?)
  - If $A \subset \mathbb{Z}/p\mathbb{Z}$ is k-stable, then $|A|/p = o(1)$ or $1 - o(1)$.
    How quickly does the $o(1)$ decay as $p \to 0$?

Open questions

- Bounds for Conant–Pillay–Terry regularity lemmas for general groups with stability/bounded VC dimension hypotheses?
  (No bounds known. Maybe $\epsilon^{-O(1)}$?)
  - If $A \subset \mathbb{Z}/p\mathbb{Z}$ is k-stable, then $|A|/p = o(1)$ or $1 - o(1)$.
    How quickly does the $o(1)$ decay as $p \to 0$?
- Removal lemma for bi-induced-F in a general (abelian) group?
  (No theorem known. Maybe $\epsilon^{-O(1)}$ bounds?)

Open questions

- Bounds for Conant–Pillay–Terry regularity lemmas for general groups with stability/bounded VC dimension hypotheses?
  (No bounds known. Maybe $\epsilon^{-O(1)}$?)
  - If $A \subset \mathbb{Z}/p\mathbb{Z}$ is k-stable, then $|A|/p = o(1)$ or $1 - o(1)$.
    How quickly does the $o(1)$ decay as $p \to 0$?
- Removal lemma for bi-induced-F in a general (abelian) group?
  (No theorem known. Maybe $\epsilon^{-O(1)}$ bounds?)
- Induced arithmetic pattern removal for general (abelian) groups?
  (No general theorem known)