

On the limit approach to random regular graphs

Balázs Szegedy

Joint work with Ágnes Backhausz

Two papers

- *On large girth regular graphs and random processes on trees, ArXiv*
- *On the almost eigenvectors of random regular graphs, In preparation*

These papers fit into a longer research project: We examine random d -regular graphs using graph limit techniques.

On random regular graphs

- Let d be a fixed number and let $G = G(n, d)$ be a random regular graph on n vertices. We will think of n as a big number.
- G looks like a tree locally but it has an interesting and mysterious global geometry. Later in this talk I will make it precise what we mean by global structure.
- G a good expander: J. Friedman: G is almost Ramanujan, $|\lambda_2| \sim 2\sqrt{d-1}$.

Random matrix theory

- A Wigner matrix is a random symmetric matrix with i.i.d entries in the upper half.
- Wigner conjectured that spacings between the lines in the spectrum of a complicated quantum system (say a heavy atom) should resemble the spacings between the eigenvalues of a random matrix. Later it was conjectured that the zeros of the Riemann zeta function on the $1/2$ line have a similar distribution.
- Random matrix theory is a very active area recently: Erdős, Knowles, Tao, Vu, Yau, Yin, etc...

- Two major questions: Eigenvalue distribution (globally and locally) and structure of eigenvectors.

Random regular graphs as random matrices

- We can think of $G = G(n, d)$ as a random symmetric $0 - 1$ matrix in which row sums and column sums are conditioned to be d . It is interesting to compare the spectral properties of Wiegner matrices (general random matrices) and the spectral properties of G .
- Wiegner semicircle Law \longleftrightarrow Kesten MacKay measure (uses only local structure of G)
- Eigenvalue spacing: Not known for random regular graphs

- **Main conjecture about eigenvectors:** If v is an eigenvector of G and normalized such that $\|v\|_2 = \sqrt{n}$ then the entry distribution of v is close to $N(0, 1)$ in the weak topology
- Despite of significant effort very little was known about the entry distribution of eigenvectors
- The most interesting results are due to Brooks and Lindenstrauss (some weak localization), Anantharaman (quantum unique ergodicity)

Main result

- **Corollary of our main theorem:** If v is an eigenvector of G and normalized such that $\|v\|_2 = \sqrt{n}$ then the entry distribution of v is close to $N(0, \sigma)$ with $0 \leq \sigma \leq$ in the weak topology.
- Our result is stronger in two ways:
 - 1.) It holds for almost eigenvectors: If $\|v\|_2 = 1$ then $\|(A - \lambda I)v\|_2 \sim 0$
 - 2.) It implies Gaussianity of joint distributions of entries in small neighborhoods

- Nice fact: Our result is best possible for almost eigenvectors! All $0 \leq \sigma \leq 1$ can occur!

Motivation

- Spectral properties of $G(n, d)$ are interesting on their own right but the problem about the eigenvectors falls into a similar class of problems as maximal independent sets, maximal cuts etc... (See the talk of N. Wormald!)
- **In all of these problems we try to construct a labeling of the vertices of G satisfying some local rule!!!**
- We can hope that the solution of any of these problems will lead to the development of general methods to study random regular graphs. For example:

Differential equation method (Wormald). In our case we develop a graph limit (ergodic theoretic) approach combined with information theory. We believe that this method will find other applications in random regular graphs! (Part of our proof is rather general)

Limits of random regular graphs

- Let d be a fixed number. Is there some nice and interesting infinite structure which is the limit object of random d -regular graphs on a growing number of vertices?
- Note that random regular graphs converge to the the infinite d -regular tree T_d in the Benjamini-Schramm metric but this is a boring fact.
- Local-global metric (Hatami-Lovász-Sz): Two graphs G and H are similar if 1.) they are similar in the

Benjamini-Schramm metric 2.) If we put any extra structure on one of the graphs (which can be expressed by a coloring of the vertices) then there is a similar structure on the other graph.

- Example: If G and H are close in this metric and G is close to be bipartite then so is H . (This similarity can not be seen from local statistics.)

Convergence of random d -regular graphs

- Is it true that a growing sequence of random d -regular graphs $\{G_i\}_{i=1}^{\infty}$ (where G_i is defined on i vertices) is local-global convergent with probability 1? We don't know.
- Theorem: From every growing sequence $\{n_i\}_{i=1}^{\infty}$ of natural numbers we can choose a sub sequence such that the above statement becomes true.
- Reason: If n is large enough then random d -regular graphs on n -vertices are highly concentrated in the local-global metric

- Note that limit objects are certain measurable graphs called graphings.

How does ergodic theory enter the picture?

- Classical ergodic theory (Furstenberg) deals with probability measures on $\{0,1\}^{\mathbb{Z}}$ that are invariant under the shift operation.
- More generally, if G is a group and it acts on Ω then we can study G invariant Borel measures on C^{Ω} where C is some topological space. For us the interesting case is when $\Omega = T_d$ (d -regular tree) and $G = \text{Aut}(T_d)$.
- if $\{G_i\}_{i=1}^{\infty}$ is a large girth sequence of d -regular graphs and $\{c_i : V(G_i) \rightarrow C\}_{i=1}^{\infty}$ is a B-S convergent sequence of colorings then the B-S limit object is an invariant measure on C^{T_d} . (This is similar

to Furstenberg's correspondence principle that was the starting point of the ergodic theoretic proof of Szemerédi's famous theorem)

- Informal definition: We call a process on T_d *typical* if it comes from sequences of random d -regular graphs.
- More precise definition: μ is typical if there is a sequence of natural numbers $\{n_i\}_{i=1}^{\infty}$ with the property that with probability one the sequence $\{G(n_i, d)\}_{i=1}^{\infty}$ of random d -regular graphs has a coloring with limit μ .

- Note that invariant processes on T_d can be looked at as a joint distributions $\{X_v\}_{v \in T_d}$ of random variables labeled by the vertices of T_d .
- Every factor of i.i.d process is typical, but there are typical processes that are not even in the weak closure of factor of i.i.d processes (Gamarnik-Sudan)

Basic philosophy

- We study the structure of random d -regular graphs by studying the properties of typical processes
- Final goal: give a useful characterization of typical processes (leads to some kind of structure theorem for random d -regular graphs)
- We collect necessary conditions

Entropy

- The first such necessary conditions follow from an earlier paper by Backhausz-Sz-Virág as correlation inequalities.
- The two papers, this talk is based on, study entropy inequalities. (Information theoretic approach)
- Theorem: Assume that C is a finite set and μ is a C -valued typical process then $(d/2)\mathbb{H}(\mu|_e) \geq (d - 1)\mathbb{H}(\mu|_o)$ where e is an edge in T_d and o is a vertex in T_d .

- Theorem: $\mathbb{H}(\mu|_S) \geq (d - 2)\mathbb{H}(\mu|_e)$ where S is a star in T_d
- An invariant process $\{X_v\}_{v \in T_d}$ is an eigenvector process with eigenvalue λ if $\sum_{w \in N(v)} X_w = \lambda X_v$ holds for every $v \in T_d$ with probability 1.
- Theorem: If μ is a typical process, $C \subset \mathbb{R}$ is finite then μ can not be an eigenvector process.
- The above theorem gives some weak restriction on the structure of eigenvectors of random regular graphs.

Gaussian eigenvector processes on the tree

- Let $X = \{X_v\}_{v \in T_d}$ be an invariant \mathbb{R} valued process on T_d such that it satisfies the eigenvector equation with λ . If X is jointly Gaussian then X is called a Gaussian eigenvector process or Gaussian wave.
- History: The theory of Gaussian eigenvector processes (or Gaussian waves) is rooted implicitly in the theory of spherical representations of $\text{Aut}(T_d)$. There is a unique Gaussian wave ϕ_λ for every $|\lambda| \leq d$. There is a 2009 paper by Yehonatan Elon with title "Gaussian Waves on the regular Tree" using explicit probabilistic language. The probabilistic theory

was further developed by Harangi and Virág. They proved that if $|\lambda| \leq 2\sqrt{d-1}$ then ϕ_λ is in the weak limit of factor of *i.i.d* processes (but it is not a factor of i.i.d process).

- Corollary (Harangi, Virág): If $|\lambda| \leq 2\sqrt{d-1}$ then ϕ_λ is typical. In particular every large girth d -regular graph has many completely delocalized approximative eigenvectors.

Difficulties and methods of the proof

- The proof can be broken down into parts that are interesting on their own right.
- First difficulty: Entropy does not work for infinite probability distributions. Our previous entropy inequalities become useless. We can try to discretize. However it turns out that discretization leads to unwanted, uncontrollable error terms. Solution: We formulate a finer entropy inequality that is designed to swallow the uncontrollable error terms (magic cancellation of infinities!). What we obtain is a quite nice differential entropy inequality.

- $D_{\text{sp}}(C) \geq (d/2)D(e)$
- Bad news: This is not tight for Gaussian so it seems to be useless for Gaussianity.
- Blow up trick gives: $D_{\text{sp}}(B_k(C)) \geq (d/2)D_{\text{sp}}(B_k(e))$.
- This becomes tighter and tighter for a Gaussian as $k \rightarrow \infty$!!! (another magic cancellation of infinities)
We needed a description of the eigenvalues of the covariance operator on large subsets of T_d . (Area law for Gaussian processes)

- Reduction: We use the inequality for all k to obtain a tight statement for $k = 0$. This says that among typical eigenvector processes corresponding to an eigenvalue λ the expression $D_{\text{sp}}(C) - (d/2)D(e)$ is at least as big as for the corresponding Gaussian wave .
- Last step: Above inequality+symmetries imply that μ is Gaussian. (Extremal problem for finite dimensional distributions) Method: Heat equation increases the value of $D_{\text{sp}}(C) - (d/2)D(e)$. (De Bruijn's inequality for Fisher information!)

Concluding remarks

- The proof is 35 pages long, contains methods from different areas (even some complex function theory). It has parts that go beyond the actual theorem.