Nodal domains of $G(n, p)$ graphs]

Mark Rudelson
joint work with Han Huang

Department of Mathematics
University of Michigan

What are nodal domains?

Let M be a compact manifold. Let Δ be the Laplacian on M, and consider an eigenfunction f_k of $-\Delta$ corresponding to the eigenvalue λ_k.

Definition

A connected component of the set where f_k is positive or negative is called a nodal domain of f_k.

Basic questions

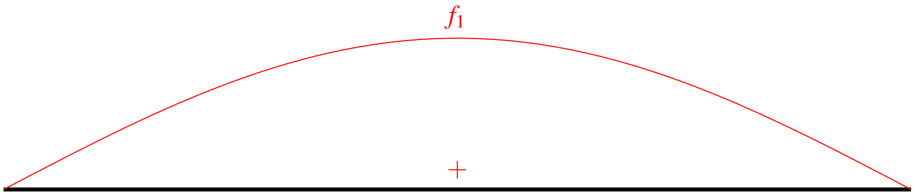
- Number of nodal domains.
- Geometry.

Typical behavior

Example

Let $M = [0, 1]$ and $\Delta f = f''$. Consider the Dirichlet boundary conditions $f(0) = f(1) = 0$.

$$\lambda_k = (\pi k)^2 \quad f_k(x) = \sin(\pi k x)$$

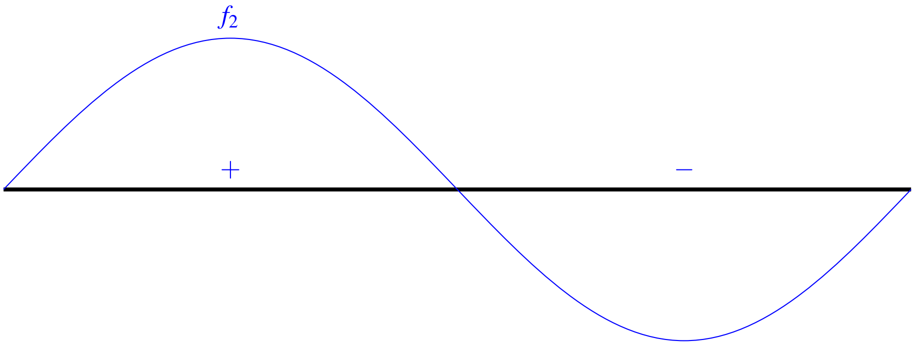


Typical behavior

Example

Let $M = [0, 1]$ and $\Delta f = f''$. Consider the Dirichlet boundary conditions $f(0) = f(1) = 0$.

$$\lambda_k = (\pi k)^2 \quad f_k(x) = \sin(\pi k x)$$

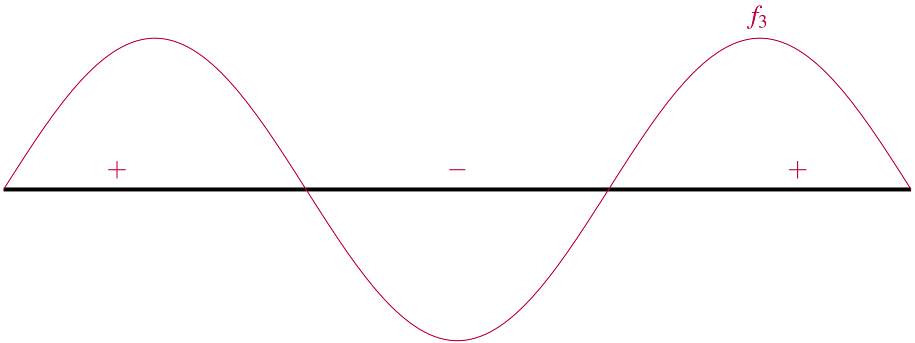


Typical behavior

Example

Let $M = [0, 1]$ and $\Delta f = f''$. Consider the Dirichlet boundary conditions $f(0) = f(1) = 0$.

$$\lambda_k = (\pi k)^2 \quad f_k(x) = \sin(\pi k x)$$



Manifolds vs graphs

Theorem (Courant)

Let M be a compact manifold. Then the number of nodal domains of k-th eigenfunction does not exceed k.

This number typically increases as k increases.

Manifolds vs graphs

Theorem (Courant)

Let M be a compact manifold. Then the number of nodal domains of k-th eigenfunction does not exceed k.

This number typically increases as k increases.

Let $G = (V, E)$ be a graph with the adjacency matrix A:

$$A_{u,v} = 1 \quad \text{if and only if } (u, v) \in E.$$

The graph Laplacian is defined as

$$\Delta f(u) = \sum_{v \sim u} f(v) - \deg(u) \cdot f(u) = (A - D)f,$$

where $D = \deg(v_1, \ldots, v_n)$.

Nodal domains of $G(n, p)$ graphs

Theorem (Dekel, Lee, Linial, $\sim$ 2007)

Let $p \in (0, 1)$. Then with high probability, any eigenvector of the adjacency matrix of a $G(n, p)$ graph has $C(p)$ nodal domains. Here $C(p)$ is independent of n and uniform over the eigenvectors.

Nodal domains of $G(n,p)$ graphs

Theorem (Dekel, Lee, Linial, $\sim$ 2007)

Let $p \in (0,1)$. Then with high probability, any eigenvector of the adjacency matrix of a $G(n,p)$ graph has $C(p)$ nodal domains. Here $C(p)$ is independent of n and uniform over the eigenvectors.

Theorem (Dekel, Lee, Linial + Arora, Bhaskara)

Let $p \in (n^{-\delta}, 1-n^{-\delta})$. Then with high probability, any non-leading eigenvector of the adjacency matrix of a $G(n,p)$ graph has 2 nodal domains.

Nodal domains of $G(n,p)$ graphs

Theorem (Dekel, Lee, Linial, $\sim$ 2007)

Let $p \in (0,1)$. Then with high probability, any eigenvector of the adjacency matrix of a $G(n,p)$ graph has $C(p)$ nodal domains. Here $C(p)$ is independent of n and uniform over the eigenvectors.

Theorem (Dekel, Lee, Linial + Arora, Bhaskara)

Let $p \in (n^{-\delta}, 1-n^{-\delta})$. Then with high probability, any non-leading eigenvector of the adjacency matrix of a $G(n,p)$ graph has 2 nodal domains.

Linial's questions

Geography of nodal domains. View the positive nodal domain as earth, and the negative one as water.

- What is the distribution of heights?
- What is the length of the shoreline?

Delocalization of eigenvectors

Let $x \in S^{n-1}$ be a unit vector of an $n \times n$ Gaussian matrix (i.i.d. or i.i.d. symmetric). Then x is uniformly distributed over S^{n-1}.

Delocalization of eigenvectors

Let $x \in S^{n-1}$ be a unit vector of an $n \times n$ random matrix (i.i.d. or i.i.d. symmetric). Then x is approximately uniformly distributed over S^{n-1}.

Delocalization of eigenvectors

Let $x \in S^{n-1}$ be a unit vector of an $n \times n$ random matrix (i.i.d. or i.i.d. symmetric). Then x is approximately uniformly distributed over S^{n-1}.

ℓ_∞ delocalization – rules out large coordinates

- $\|x\|_\infty = O\left(\frac{\sqrt{\log n}}{\sqrt{n}}\right)$ with high probability.
- Erdős, Knowles, Yau, Yin for i.i.d. symmetric matrices,
  R'-Vershynin for i.i.d. matrices

$$\|x\|_\infty = O\left(\frac{\log^c n}{\sqrt{n}}\right) \quad \text{with probability at least } 1 - n^{-C}$$

Delocalization of eigenvectors

No-gaps delocalization – rules out small coordinates

- With high probability, any $J \subset [n]$ supports a non-negligible mass:

$$\|x|_J\|_2 \geq c \left(\frac{|J|}{n} \right)^{\alpha}$$

Delocalization of eigenvectors

No-gaps delocalization – rules out small coordinates

- With high probability, any $J \subset [n]$ supports a non-negligible mass:

$$\|x|_J\|_2 \geq c \left(\frac{|J|}{n} \right)^{\alpha}$$

- R'-Vershynin for i.i.d. i.i.d. symmetric, i.i.d. skew-symmetric etc. matrices
  If $|J| \geq n^{6/7}$ then

$$\|x|_J\|_2 \geq c \left(\frac{|J|}{n} \right)^{6} \quad \text{with probability at least } 1 - \exp(-|J|).$$

The two types for delocalization are complementary and are frequently used in combination.

Dekel, Lee, Linial, Arora, Bhaskara theorem

Theorem

Let $p \in (n^{-\delta}, 1 - n^{-\delta})$. Then with high probability, any non-leading eigenvector of the adjacency matrix of a $G(n,p)$ graph has 2 nodal domains.

Proof (Step 1).

Denote by P and N the largest positive and negative nodal domains. Set $W = [n] \setminus (P \cup N)$. Then

$$|W| = O\left(\frac{\log^2 n}{p^2}\right).$$

Dekel, Lee, Linial, Arora, Bhaskara theorem

Theorem

Let $p \in (n^{-\delta}, 1 - n^{-\delta})$. Then with high probability, any non-leading eigenvector of the adjacency matrix of a $G(n,p)$ graph has 2 nodal domains.

Proof (Step 1).

Denote by P and N the largest positive and negative nodal domains. Set $W = [n] \setminus (P \cup N)$. Then

$$|W| = O\left(\frac{\log^2 n}{p^2}\right).$$

- If $V_1, V_2 \subset [n]$ are not connected, then one of them has $O\left(\frac{\log n}{p}\right)$ elements.
- Take one vertex from each positive (negative) nodal domain. They form an independent set. The cardinality of an independent set is $O\left(\frac{\log n}{p}\right)$.

□

Dekel, Lee, Linial, Arora, Bhaskara theorem

Assume that $p = \text{const}$.
P and N are the largest positive and negative nodal domains, $W = [n] \setminus (P \cup N)$.

$$|W| = O\left(\log^2 n\right).$$

Dekel, Lee, Linial, Arora, Bhaskara theorem

Assume that $p = \text{const}$.
P and N are the largest positive and negative nodal domains, $W = [n] \setminus (P \cup N)$.

$$|W| = O\left(\log^2 n\right).$$

Assume that $W \neq \varnothing$. Let $w \in W, x_w < 0$. Then $w \not\sim N$, so

$$\lambda x_w = (Ax)_w = \sum_{\substack{v \in P \\ v \sim w}} x_v + \sum_{\substack{v \in W \\ v \sim w}} x_v$$

Dekel, Lee, Linial, Arora, Bhaskara theorem

Assume that $p = \text{const}$.
P and N are the largest positive and negative nodal domains, $W = [n] \setminus (P \cup N)$.

$$|W| = O\left(\log^2 n\right).$$

Assume that $W \neq \varnothing$. Let $w \in W, x_w < 0$. Then $w \not\sim N$, so

$$\lambda x_w = (Ax)_w = \sum_{\substack{v \in P \\ v \sim w}} x_v + \sum_{\substack{v \in W \\ v \sim w}} x_v$$

$$\Rightarrow \quad \sum_{\substack{v \in P \\ v \sim w}} x_v \leq (|\lambda| + |W|) \cdot \|x\|_\infty \overset{\ell_\infty \text{ delocalization}}{=} \log^C n$$

Dekel, Lee, Linial, Arora, Bhaskara theorem

Assume that $p = \text{const}$.
P and N are the largest positive and negative nodal domains, $W = [n] \setminus (P \cup N)$.

$$|W| = O\left(\log^2 n\right).$$

Assume that $W \neq \varnothing$. Let $w \in W, x_w < 0$. Then $w \not\sim N$, so

$$\lambda x_w = (Ax)_w = \sum_{\substack{v \in P \\ v \sim w}} x_v + \sum_{\substack{v \in W \\ v \sim w}} x_v$$

$$\Rightarrow \quad \sum_{\substack{v \in P \\ v \sim w}} x_v \leq (|\lambda| + |W|) \cdot \|x\|_\infty \overset{\ell_\infty \text{ delocalization}}{=} \log^C n$$

$$\Rightarrow \quad \sum_{\substack{v \in P \\ v \sim w}} x_v^2 \leq \sum_{\substack{v \in P \\ v \sim w}} x_v \cdot \|x\|_\infty \overset{\ell_\infty \text{ delocalization}}{=} O(n^{-1/4} \cdot \log^C n).$$

Dekel, Lee, Linial, Arora, Bhaskara theorem

Assume that $p = \text{const}$.
P and N are the largest positive and negative nodal domains, $W = [n] \setminus (P \cup N)$.

$$|W| = O\left(\log^2 n\right).$$

Assume that $W \neq \varnothing$. Let $w \in W, x_w < 0$. Then $w \not\sim N$, so

$$\lambda x_w = (Ax)_w = \sum_{\substack{v \in P \\ v \sim w}} x_v + \sum_{\substack{v \in W \\ v \sim w}} x_v$$

$$\Rightarrow \quad \sum_{\substack{v \in P \\ v \sim w}} x_v \leq (|\lambda| + |W|) \cdot \|x\|_\infty \overset{\ell_\infty \text{ delocalization}}{=} \log^C n$$

$$\Rightarrow \quad \sum_{\substack{v \in P \\ v \sim w}} x_v^2 \leq \sum_{\substack{v \in P \\ v \sim w}} x_v \cdot \|x\|_\infty \overset{\ell_\infty \text{ delocalization}}{=} O(n^{-1/4} \cdot \log^C n).$$

This contradicts no-gaps delocalization since $|\{v \in P,\ v \sim w\}| \geq cn$.

Geography of nodal domains

Theorem

Let $p \geq n^{-\delta}$. Then with high probability, any vertex of the positive nodal domain of a non-leading eigenvector is connected to the negative domain.
A similar statement holds for any vertex of the negative domain.

Geography of nodal domains

Theorem

Let $p \geq n^{-\delta}$. Then with high probability, any vertex of the positive nodal domain of a non-leading eigenvector is connected to the negative domain.
A similar statement holds for any vertex of the negative domain.

Conclusion: A $G(n,p)$ graph is a swamp.

Size of nodal domains

Sizes of nodal domains

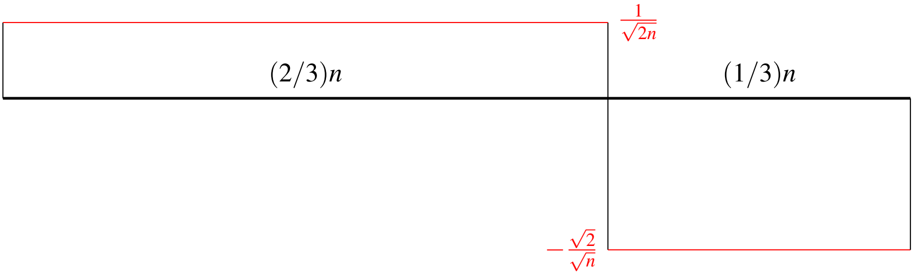
Is it true that $|P| \approx |N|$?

Size of nodal domains

Sizes of nodal domains

Is it true that $|P| \approx |N|$?

This would not follow from delocalization.
Indeed, imagine that the eigenvector x has the following coordinates:



Size of nodal domains, bulk case

Eigenvalues:

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$$

Theorem (Huang, R')

Let $\varepsilon, \kappa \in (0, 1)$, *and let* $p \in \left[n^{-c}, \frac{1}{2}\right]$.
Let u_α *be an eigenvector of a* $G(n, p)$ *graph with* $\alpha \in [\kappa n, n - \kappa n]$.
Denote by P *and* N *the nodal domains of this eigenvector. Then*

$$\mathbb{P}\left(|P| \vee |N| \geq \left(\frac{1}{2} + \varepsilon\right) n\right) \leq n^{-\eta}$$

for some $\eta = \eta(\varepsilon, \kappa) > 0$.

Size of nodal domains, edge case

Theorem (Huang, R')

Let $p \in (0, 1)$.
Let u_α *be a non-leading eigenvector of a* $G(n, p)$ *graph with*

$$\min\{\alpha, n - \alpha\} \leq \textit{small}$$

Denote by P and N the nodal domains of this eigenvector.
Then, for any $\varepsilon > 0$,

$$\mathbb{P}\left(|P| \vee |N| \geq \left(\frac{1}{2} + n^{-\frac{1}{6}+\varepsilon}\right) n\right) \leq n^{-\delta}.$$

for large n, with $\delta = \delta(\varepsilon) > 0$ *independent of n and p.*

Ideas of proof, bulk case

Our goal is to show that with high probability,

$$\sum_{i=1}^{n} \operatorname{sign}(x_i) = o(n)$$

for an eigenvector x of A.

Ideas of proof, bulk case

Our goal is to show that with high probability,

$$\sum_{i=1}^{n} \operatorname{sign}(x_i) = o(n)$$

for an eigenvector x of A.
This can be derived by Markov inequality if

$$\mathbb{E}\left(\sum_{i=1}^{n} \operatorname{sign}(x_i)\right)^2 = o\left(n^2\right).$$

Ideas of proof, bulk case

Our goal is to show that with high probability,

$$\sum_{i=1}^{n} \operatorname{sign}(x_i) = o(n)$$

for an eigenvector x of A.
This can be derived by Markov inequality if

$$\mathbb{E}\left(\sum_{i=1}^{n} \operatorname{sign}(x_i)\right)^2 = o\left(n^2\right).$$

The latter equation follows if for $i \neq j$,

$$\mathbb{E}\,\operatorname{sign}(x_i x_j) = o(1).$$

Ideas of proof, bulk case

Asymptotic normality:

Theorem (Bourgade, J. Huang, Yau)

For any polynomial $f : \mathbb{R} \to \mathbb{R}$ *for any* $n \geq n(f)$, $\alpha \in [\kappa n : n - \kappa n]$ *and any* $q \in S^{n-1}$, $q \perp (1, \dots, 1)$, *there exists an* $\nu > 0$ *such that*

$$|\mathbb{E}f(n\langle q, x\rangle^2) - \mathbb{E}f(g^2)| \leq n^{-\nu}.$$

Ideas of proof, bulk case

Asymptotic normality:

Theorem (Bourgade, J. Huang, Yau)

For any polynomial $f : \mathbb{R} \to \mathbb{R}$ for any $n \geq n(f)$, $\alpha \in [\kappa n : n - \kappa n]$ and any $q \in S^{n-1}$, $q \perp (1, \dots, 1)$, there exists an $\nu > 0$ such that

$$|\mathbb{E} f(n\langle q, x\rangle^2) - \mathbb{E} f(g^2)| \leq n^{-\nu}.$$

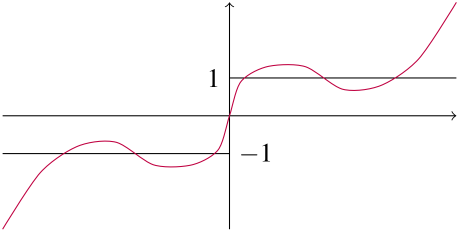
1. Pass to a polynomial of 4 variables removing the condition $q \perp (1, \dots, 1)$.
2. Pass to an even polynomial of 2 variables.
3. Pass to an odd polynomial of $x_i x_j$:

$$|\mathbb{E} f(n \cdot x_i x_j) - \mathbb{E} f(g_1 g_2)| \leq n^{-\nu} \qquad \text{and} \qquad \mathbb{E} f(g_1 g_2) = 0.$$

Ideas of proof, bulk case

$$|\mathbb{E} f(n \cdot x_i x_j)| \le n^{-\nu} \quad \stackrel{?}{\Rightarrow} \quad |\mathbb{E}\, \mathrm{sign}\,(n \cdot x_i x_j)\,| \le n^{-\nu}$$

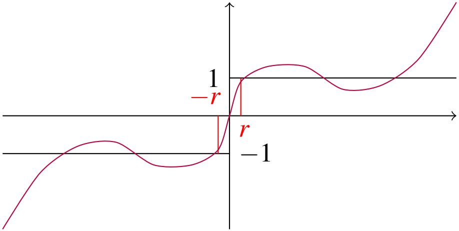
Approximate sign by an odd polynomial:



Ideas of proof, bulk case

$$|\mathbb{E}f(n \cdot x_i x_j)| \leq n^{-\nu} \quad \overset{?}{\Rightarrow} \quad |\mathbb{E}\,\mathrm{sign}\,(n \cdot x_i x_j)\,| \leq n^{-\nu}$$

Approximate sign by an odd polynomial:

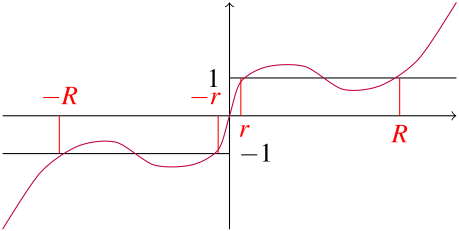


- Interval $(0, r]$ – few coordinates – no-gaps decollation.

Ideas of proof, bulk case

$$|\mathbb{E}f(n \cdot x_i x_j)| \le n^{-\nu} \quad \overset{?}{\Rightarrow} \quad |\mathbb{E}\,\text{sign}\,(n \cdot x_i x_j)\,| \le n^{-\nu}$$

Approximate sign by an odd polynomial:

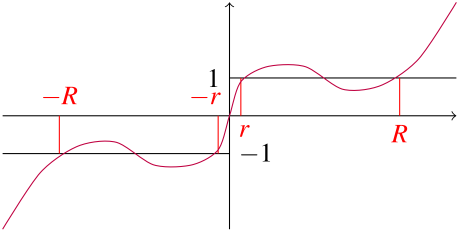


- Interval $(0, r]$ – few coordinates – no-gaps decollation.
- Interval $(r, R]$ – many coordinates – uniform approximation.

Ideas of proof, bulk case

$$|\mathbb{E} f(n \cdot x_i x_j)| \le n^{-\nu} \quad \overset{?}{\Rightarrow} \quad |\mathbb{E}\,\mathrm{sign}\,(n \cdot x_i x_j)\,| \le n^{-\nu}$$

Approximate sign by an odd polynomial:



- Interval $(0, r]$ – few coordinates – no-gaps decollation.
- Interval $(r, R]$ – many coordinates – uniform approximation.
- Interval $(R, \infty]$ – few coordinates – L_2 bound with respect to the distribution of $g_1 g_2$.

The last two requirements are conflicting.

Open problems - sparse graphs

- Multiple nodal domains. Assume that $p \ll 1$.
  The proof of DLL+AB theorem was based on

$$|W| = O\left(\frac{\log^2 n}{p^2}\right).$$

  This is vacuous if $p \leq n^{-1/2}$.

Open problems - sparse graphs

- Multiple nodal domains. Assume that $p \ll 1$.
  The proof of DLL+AB theorem was based on

$$|W| = O\left(\frac{\log^2 n}{p^2}\right).$$

  This is vacuous if $p \leq n^{-1/2}$.
  Numerical experiments (Huang, Eldan): if p is small and λ is at the negative edge then there are multiple one vertex nodal domains.
  Conjecture (Eldan):

$$p \leq C\frac{\log n}{n} \quad \text{with } C > 1.$$

Open problems - sparse graphs

- Multiple nodal domains. Assume that $p \ll 1$.
  The proof of DLL+AB theorem was based on

$$|W| = O\left(\frac{\log^2 n}{p^2}\right).$$

  This is vacuous if $p \leq n^{-1/2}$.
  Numerical experiments (Huang, Eldan): if p is small and λ is at the negative edge then there are multiple one vertex nodal domains.
  Conjecture (Eldan):

$$p \leq C\frac{\log n}{n} \quad \text{with } C > 1.$$

- Geography of nodal domains
  Numerical experiments (Huang): if p is small and λ is at the negative edge then the main nodal domains have non-trivial geography.

Open problems - dense graphs

Definition

Let $G = (V, E)$ be a graph. A partition $V = U \sqcup W$ is called

- inner if any vertex in U and W has more neighbors in the same set than in the other one.
- outer if both U and V have the opposite property.

Open problems - dense graphs

Definition

Let $G = (V, E)$ be a graph. A partition $V = U \sqcup W$ is called

- inner if any vertex in U and W has more neighbors in the same set than in the other one.
- outer if both U and V have the opposite property.

Conjecture (Linial)

- If $\lambda > 0$ then $V = P \sqcup N$ is an inner partition with high probability.
- If $\lambda < 0$ then $V = P \sqcup N$ is an outer partition with high probability.