Typical Support of Closed Walks and Eigenvalue Multiplicity

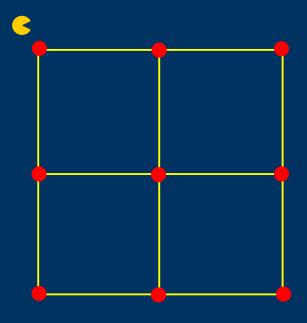
Theo McKenzie

with Peter Rasmussen (University of Copenhagen) and Nikhil Srivastava (University of California, Berkeley)

Delaware Discrete Mathematics/Algebra Seminar 2/24/2021

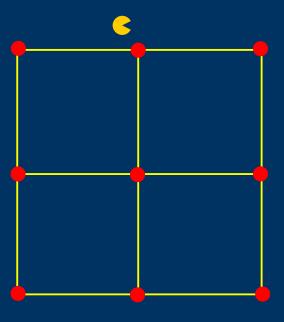


- Our goal is to understand the behavior of walks in large graphs.
- A walk is performed by choosing an adjacent node in the graph.





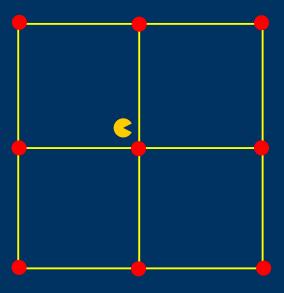
- Our goal is to understand the behavior of walks in large graphs.
- A walk is performed by choosing an adjacent node in the graph.



3

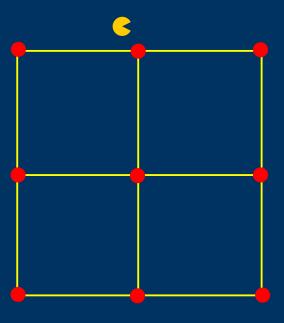


- Our goal is to understand the behavior of walks in large graphs.
- A walk is performed by choosing an adjacent node in the graph.





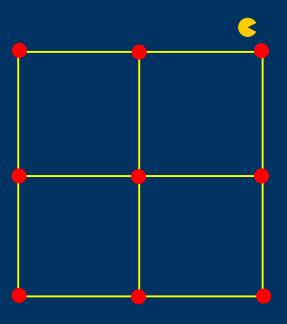
- Our goal is to understand the behavior of walks in large graphs.
- A walk is performed by choosing an adjacent node in the graph.





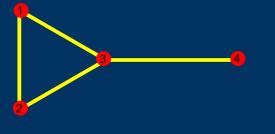


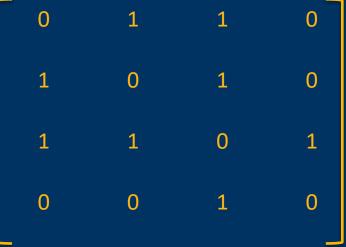
- Our goal is to understand the behavior of walks in large graphs.
- A walk is performed by choosing an adjacent node in the graph.



Adjacency Matrix

 Encode the walk through an "adjacency matrix" A, with rows/columns corresponding to the vertices, and putting a 1 between connected vertices.





- Note that as the matrix is symmetric, the eigenvalues are real and can be ordered $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$, where *n* is the number of vertices.
- Multiplying by the matrix can be thought of as a step in the walk.
- The entry $(A^k)_{uv}$ corresponds to walks of length k between u and v.

Question (van Lint and Seidel 1966):

What is the maximum number of lines in \mathbb{R}^d that all share the same angle θ , for $0 < \theta < \pi/2$?

Theorem (Jiang, Tidor, Yao, Zhang and Zhao '19): If there exists a minimal k such that there exists a graph on kvertices with spectral radius exactly $(1 - \alpha)/(2\alpha)$, then the maximum is $\lfloor k(d - 1)/(k - 1) \rfloor$ for $\alpha = \arccos(\theta)$ and large enough d.

Otherwise, the maximum is d + o(d).

Question (van Lint and Seidel 1966):

What is the maximum number of lines in \mathbb{R}^d that all share the same angle θ , for $0 < \theta < \pi/2$?

Theorem (Jiang, Tidor, Yao, Zhang and Zhao '19):

If there exists a minimal k such that there exists a graph on k vertices with spectral radius exactly $(1 - \alpha)/(2\alpha)$, then the maximum is $\lfloor k(d - 1)/(k - 1) \rfloor$ for $\alpha = \arccos(\theta)$ and large enough d. Otherwise, the maximum is d + o(d).

The key ingredient was a property of graphs. Namely, showing that for every bounded degree matrix, the multiplicity of the second eigenvalue is o(n).

Theorem (Jiang et al. '19):

For every bounded degree connected graph with *n* vertices, the second eigenvalue of the adjacency matrix has multiplicity $O(\frac{n}{\log \log n})$.

The key ingredient was a property of graphs. Namely, showing that for every bounded degree matrix, the multiplicity of the second eigenvalue is o(n).

Theorem (Jiang et al. '19):

For every bounded degree connected graph with *n* vertices, the second eigenvalue of the adjacency matrix has multiplicity $O(\frac{n}{\log \log n})$.

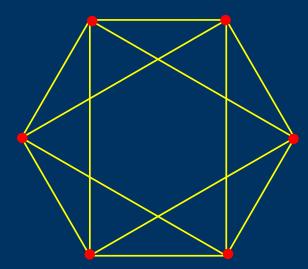
Compare this to the best lower bound, which says that the second eigenvalue of Cayley graphs of PSL(2, \mathbb{Z}_p) has multiplicity $\Omega(n^{1/3})$.

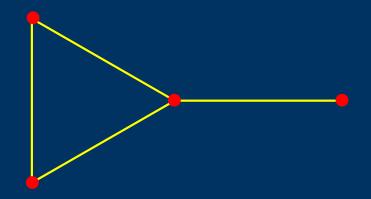
Theorem

• **Question**: Can we close this gap at all?

Theorem A (M.-Rasmussen-Srivastava): For any bounded degree **regular** connected graph with *n* vertices, the multiplicity of the second eigenvalue of *A* is at most $O(\frac{n}{\log^{1/5-o_n(1)} n})$.

Regularity





regular

non-regular

Theorem

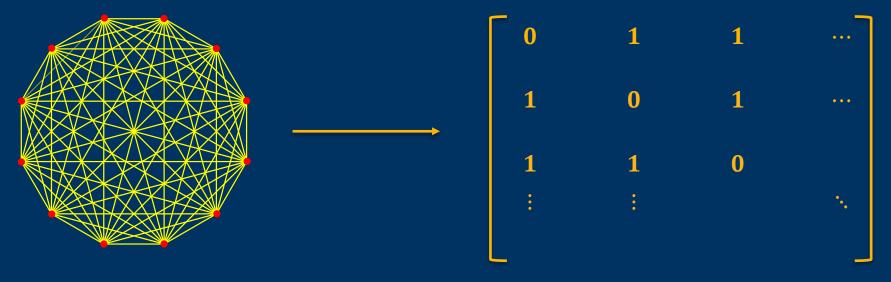
• **Question**: Can we close this gap at all?

• **Theorem A** (M.-Rasmussen-Srivastava): For any bounded degree **regular** connected graph with *n* vertices, the multiplicity of the second eigenvalue of *A* is at most $O(\frac{n}{\log^{1/5} - o_n(1)n})$.

Bounded Degree

There are examples of graphs with non-bounded degree that have high eigenvalue multiplicity.

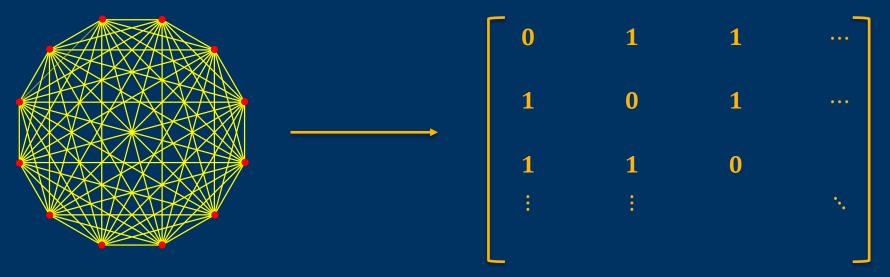
• For example, the complete graph where every vertex is connected to every other vertex.



Bounded Degree

There are examples of graphs with non-bounded degree that have high eigenvalue multiplicity.

• For example, the complete graph where every vertex is connected to every other vertex.

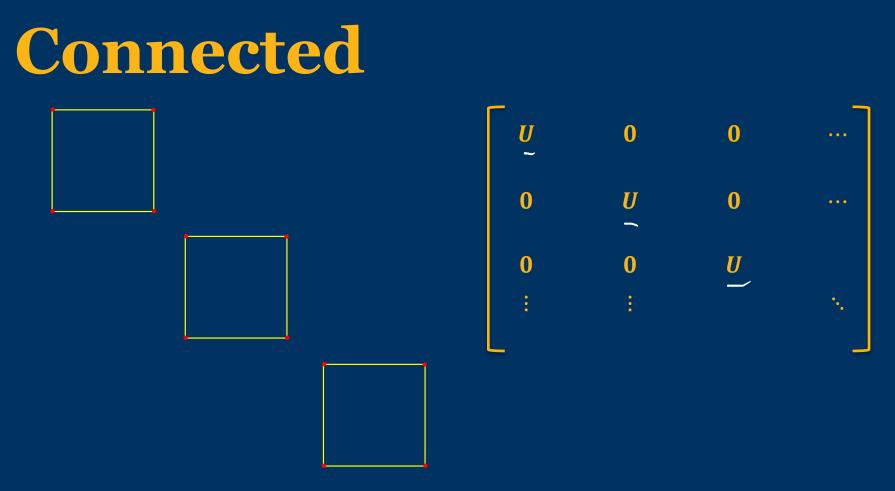


If J is the all-ones matrix and I the identity, then A = J - I. The second eigenvalue has multiplicity n - 1.

Theorem

• **Question**: Can we close this gap at all?

• **Theorem A** (M.-Rasmussen-Srivastava): For any bounded degree **regular** connected graph with *n* vertices, the multiplicity of the second eigenvalue of *A* is at most $O(\frac{n}{\log^{1/5-o_n(1)} n})$.



If there are k copies of an adjacency matrix U, then the multiplicity of the second eigenvalue is at least k.

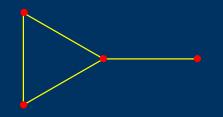
Theorem

• **Question**: Can we close this gap at all?

• **Theorem A** (M.-Rasmussen-Srivastava): For any bounded degree **regular** connected graph with *n* vertices, the multiplicity of the second eigenvalue of *A* is at most $O(\frac{n}{\log^{1/5-o_n(1)} n})$.

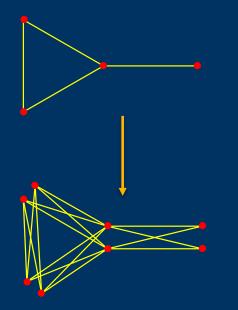
Second eigenvalue

Arbitrary eigenvalues can have linear multiplicity.



Second eigenvalue

Arbitrary eigenvalues can have linear multiplicity.



By creating two copies of each vertex, at least half of the eigenvalues are 0.

Theorem

• **Question**: Can we close this gap at all?

• **Theorem A** (M.-Rasmussen-Srivastava): For any bounded degree **regular** connected graph with *n* vertices, the multiplicity of the second eigenvalue of *A* is at most $O(\frac{n}{\log^{1/5} - o_n(1)})$.

• This beats the bound from Jiang et al. if the graph is regular.

Theorem

• **Question**: Can we close this gap at all?

• **Theorem A** (M.-Rasmussen-Srivastava): For any bounded degree **regular** connected graph with *n* vertices, the multiplicity of the second eigenvalue of *A* is at most $O(\frac{n}{\log^{1/5-o_n(1)} n})$.

• This beats the bound from Jiang et al. if the graph is regular.

• trace
$$(A^{2k}) = \sum_{u} (A^{2k})_{uu} = \sum_{i=1}^{n} \lambda_i^{2k}$$

- We use the trace as a proxy for the multiplicity of the eigenvalue.
- We therefore need to bound the trace of A^{2k} .
- The trace of A^{2k} is $\sum_{u} (A^{2k})_{uu} = \sum_{u} e_v^T A^{2k} e_v$.

• trace
$$(A^{2k}) = \sum_{u} (A^{2k})_{uu} = \sum_{i=1}^{n} \lambda_i^{2k}$$

- We use the trace as a proxy for the multiplicity of the eigenvalue.
- We therefore need to bound the trace of A^{2k} .
- The trace of A^{2k} is $\sum_{u} (A^{2k})_{uu} = \sum_{u} e_u^T A^{2k} e_u$
 - The walk starts at *u*.

• trace
$$(A^{2k}) = \sum_{u} (A^{2k})_{uu} = \sum_{i=1}^{n} \lambda_i^{2k}$$

- We use the trace as a proxy for the multiplicity of the eigenvalue.
- We therefore need to bound the trace of A^{2k} .
- The trace of A^{2k} is $\sum_{u} (A^{2k})_{uu} = \sum_{u} e_u^T A^{2k} e_u$
 - The walk starts at *u*.
 - We walk on our graph for 2k steps.

26

• trace
$$(A^{2k}) = \sum_{u} (A^{2k})_{uu} = \sum_{i=1}^{n} \lambda_i^{2k}$$

- We use the trace as a proxy for the multiplicity of the eigenvalue.
- We therefore need to bound the trace of A^{2k} .
- The trace of A^{2k} is $\sum_{u} (A^{2k})_{uu} = \sum_{u} e_u^T A^{2k} e_u$
 - The walk starts at *u*.
 - We walk on our graph for 2k steps.
 - The walk ends at u.

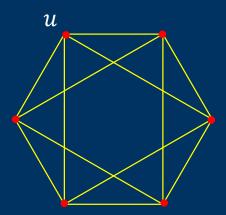
• trace
$$(A^{2k}) = \sum_{u} (A^{2k})_{uu} = \sum_{i=1}^{n} \lambda_i^{2k}$$

- We use the trace as a proxy for the multiplicity of the eigenvalue.
- We therefore need to bound the trace of A^{2k} .
- The trace of A^{2k} is $\sum_{u} (A^{2k})_{uu} = \sum_{u} e_u^T A^{2k} e_u$
 - The walk starts at *u*.
 - We walk on our graph for 2k steps.
 - The walk ends at u.
- **Definition**: A walk is closed if it ends where it starts.
- Our goal is to count the number of closed walks.

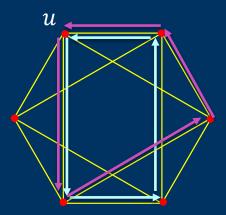
Cauchy's Interlacing Theorem

- **Cauchy's Interlacing Theorem**: for an $n 1 \times n 1$ principal submatrix with eigenvalues μ_i of an $n \times n$ matrix A with eigenvalues λ_i , the eigenvalues interlace: $\lambda_1 \leq \mu_1 \leq \lambda_2 \leq \cdots \leq \mu_{n-1} \leq \lambda_n$.
- Cauchy's Interlacing Theorem tells us that if the multiplicity of λ_2 in any submatrix of size n s is m, then in the original graph it is at most m + s.
- Idea: Delete vertices such that most closed walks are deleted. As there are now few closed walks, then the trace and the multiplicity of λ₂ is low on this subgraph. Then the multiplicity of λ₂ in the full graph is that of the subgraph plus the number of vertices deleted.

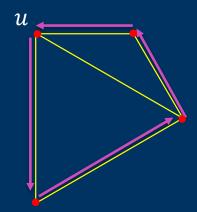
• The diagonal entries of *A*^{2k} correspond to the walks that are closed. We want to show most walks are deleted.



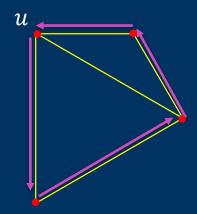
• The diagonal entries of *A*^{2k} correspond to the walks that are closed. We want to show most walks are deleted.



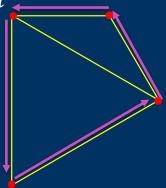
• The diagonal entries of *A*^{2*k*} correspond to the walks that are closed. We want to show most walks are deleted.



- The diagonal entries of *A*^{2*k*} correspond to the walks that are closed. We want to show most walks are deleted.
- Most walks will be deleted if most walks hit many different vertices.



- The diagonal entries of *A*^{2*k*} correspond to the walks that are closed. We want to show most walks are deleted.
- Most walks will be deleted if most walks hit many different vertices.
- If we delete most walks, then the trace is low, meaning the multiplicity in the subgraph is low, meaning the original multiplicity is low.



Support Question

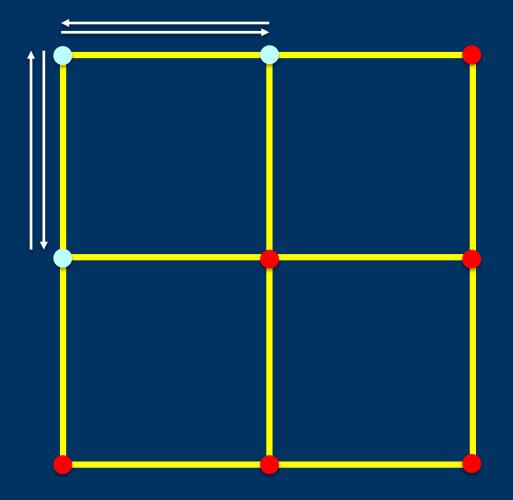
New Goal: Show that most closed walks of length *k* visit many vertices.

Support Question

New Goal: Show that most closed walks of length k visit many vertices.

• If I do not care whether the walk is closed, the problem is much easier. The hard part is the closed condition.

Example



High Support

Theorem B: With high probability, the support of a closed walk of length 2k on a regular graph has support $\Omega(k^{1/5})$.

High Support

Theorem B: With high probability, the support of a closed walk of length 2k on a regular graph has support $\Omega(k^{1/5})$.

Theorem B (detailed):

The probability a closed walk of length 2k on a regular graph has support at most s is $o_k(1)$ of the probability of having support at most 2s, for $s = O(k^{1/5})$.

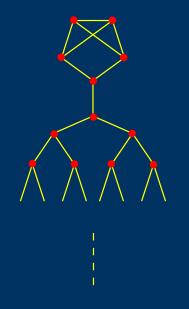
The path (warmup)

• For the path, on average we travel $\Theta(\sqrt{k})$ from the starting vertex on a closed walk of length 2k.



The bulb tree (warmup)

 For the tree with a bulb at the root, on average we do not travel more than Θ(log k) from the bulb.



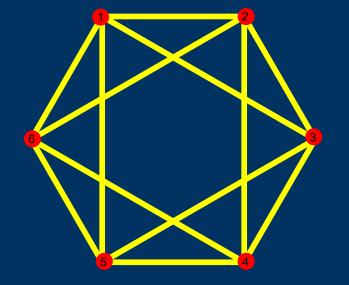
- How many walks are closed, starting at *u* and stay within the set *S*?
- If a walk stays within the set *S*, then it is counted in the quadratic form of the submatrix $e_u^T A_s^{2k} e_u$.

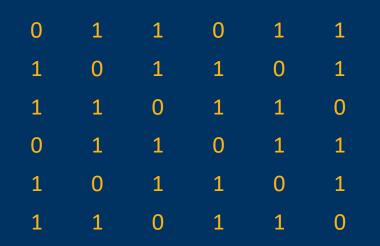
- How many walks are closed, starting at *u* and stay within the set *S*?
- If a walk stays within the set *S*, then it is counted in the quadratic form of the submatrix $e_u^T A_S^{2k} e_u$.
- We start our walk at *u*.

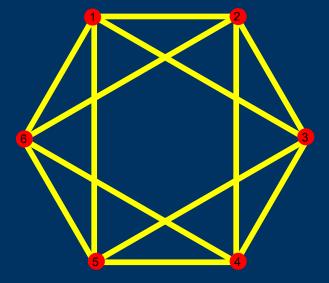
- How many walks are closed, starting at *u* and stay within the set *S*?
- If a walk stays within the set *S*, then it is counted in the quadratic form of the submatrix $e_u^T A_S^{2k} e_u$.
- We start our walk at *u*.
- We remain within *S* for 2*k* steps.

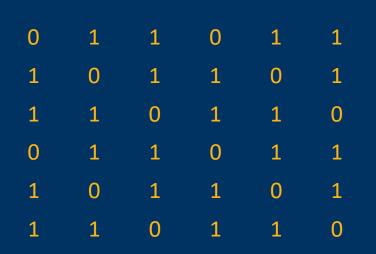
- How many walks are closed, starting at *u* and stay within the set *S*?
- If a walk stays within the set *S*, then it is counted in the quadratic form of the submatrix $e_u^T A_S^{2k} e_u$.
- We start our walk at *u*.
- We remain within *S* for 2*k* steps.
- We end at *u*.

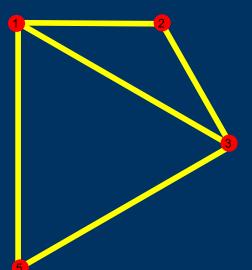
- How many walks are closed, starting at *u* and stay within the set *S*?
- If a walk stays within the set *S*, then it is counted in the quadratic form of the submatrix $e_u^T A_s^{2k} e_u$.
- We start our walk at *u*.
- We remain within *S* for 2*k* steps.
- We end at *u*.
- This quadratic form is upper bounded by the top eigenvalue of A_S , λ_S .

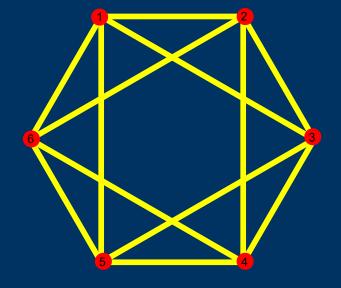


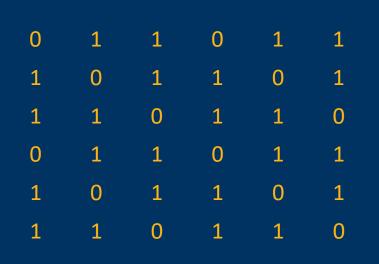


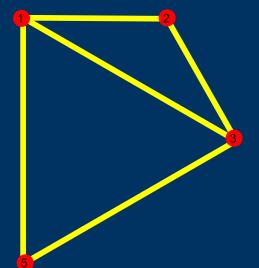


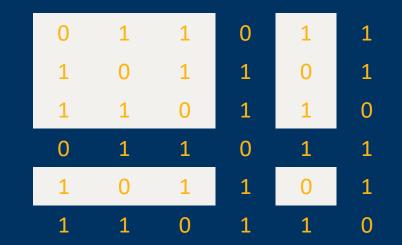


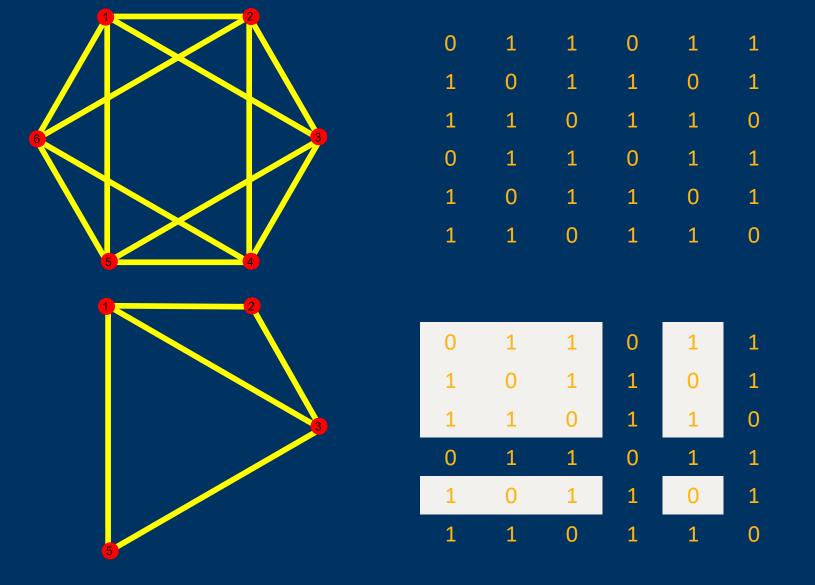












We can count walks on this support using the submatrix A_S .

Resolution

- The number of closed walks remaining on a set S after 2k steps is at most λ_S^{2k} .
- Moreover, we know that for large k, $e_v^T A_S^{2k} e_v \approx \langle \psi_S, e_v \rangle^2 \lambda_S^{2k}$.

Claim: To show that there are many more walks on *T* than *S*, it is sufficient to show that $\lambda_T^{2k} \gg \lambda_S^{2k}$.

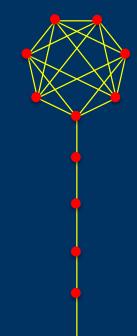
• If I find a set T such that $\lambda_T \ge (1 + \epsilon)\lambda$, then even if ϵ is small, for large enough k, $\lambda_T^{2k} \gg \lambda_S^{2k}$.

Lemma A: There exists a set $T = S \cup \{v\}$ such that $\lambda_T > \lambda_S(1 + \frac{c}{|S|^5})$. Therefore, for $k \gg |S|^5$, $\lambda_T^{2k} \gg \lambda_S^{2k}$.

How to increase eigenvalue?

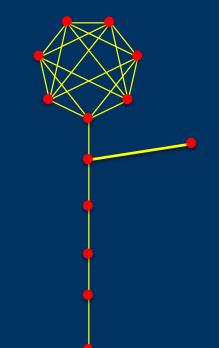
- **Lemma B:** For any vertex $v \notin S$ that neighbors a vertex $u \in S$, the top eigenvalue of $T = S \cup \{v\}$ satisfies $\lambda_T \ge \lambda_S + \Omega(\psi_S(u)^2)$.
- **Remark:** As we are looking at a subset of a regular graph, we can extend the graph at any vertex which is not of maximal degree in *S*.
- **Theorem C**: For any connected graph of bounded degree, there is a vertex u of non-maximal degree such that $\psi(u) = \Omega(n^{-\frac{5}{2}})$.
- Note that we only need one vertex with non-maximal degree to have large value in ψ. Cioabă and Gregory give a lower bound on the minimum value in the principal eigenvector, but that bound is exponentially small.

Example of the Lollipop



• The lollipop has an exponentially small value at the end of its tail, but vertices near the bulb still have polynomially large value.

Example of the Lollipop



• The lollipop has an exponentially small value at the end of its tail, but vertices near the bulb still have polynomially large value.

Result

•

- We know that for some vertex t, we have $\psi_S(t) \ge 1/\sqrt{|S|}$
 - <u>If we can bound the ratio</u>

 $\psi_S(u)/\psi_S(t)$

for this t, that is sufficient.

• Lemma C: There exists a u of non-maximal degree for which $\frac{\psi_s(u)}{\psi_s(t)} = \Omega(\frac{1}{D|\partial S|})$

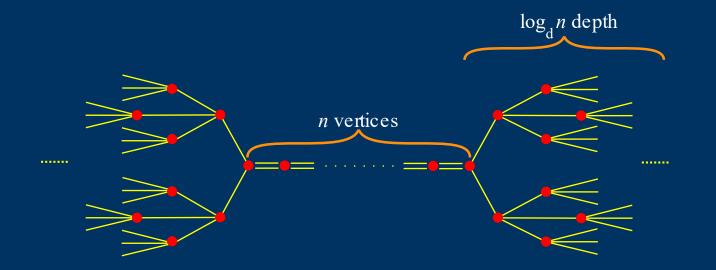
where D is the diameter of the graph, and $|\partial S|$ is the number of vertices of non-maximal degree.

 Both these quantities are at most |S|, which translates into a bound of

 $\psi_S(u) \ge 1/|S|^{\frac{5}{2}}$

Mangrove

• The worst case is when both the boundary and diameter are the order of |S|. In the below example, for all u on the boundary $\psi(u) = \Theta(1/|S|^{\frac{5}{2}})$, which is the bound given by our work.



Full Results

• Our full result bounds the number of eigenvalues in an interval.

• **Theorem A** (full): The number of eigenvalues of *A* in the interval $\left[\left(1 - \frac{\log \log n}{\log n}\right)\lambda_2, \lambda_2\right]$ is $O\left(\frac{n}{\log^{1/5 - o_n(1)}n}\right)$.

• For bipartite Ramanujan graphs, the number of eigenvalues in this interval is $\Omega\left(\frac{n}{\log^{3/2}n}\right)$, meaning our result is tight except for potentially the exponent.

Full Increase

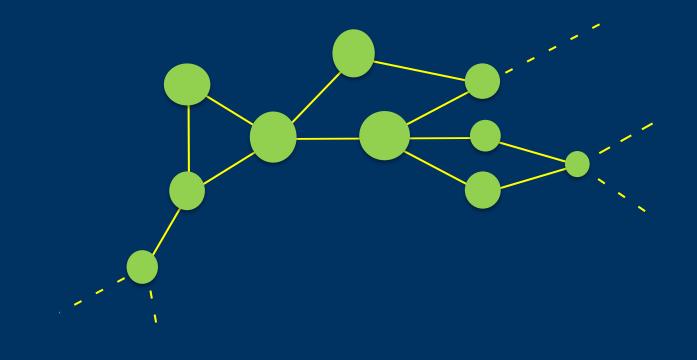
Lemma A (full): For every connected subset of vertices *S* of a regular graph there is a subset of vertices $T \supset S$, |T| = 2|S|, such that $\lambda_1(A_T) \ge \lambda_1(A_S) + \frac{c}{|S|^4}$

for some constant c.

- Method
- 1) Show $\exists u \in \partial S$ such that $\psi_S(u) \ge \frac{1}{|S|^{5/2}}$
- 2) Show that for a vertex $v \notin S$ adjacent to u, $\lambda_1(A_{S \cup v}) = \lambda_1(A_S) + \Omega(\psi_S(u)^2).$

Repeat this process |S| times to achieve a set of size 2|S|.

Visualization of 2



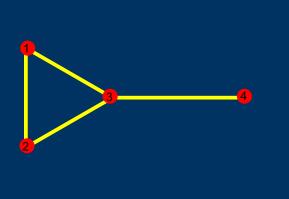
Visualization of 2

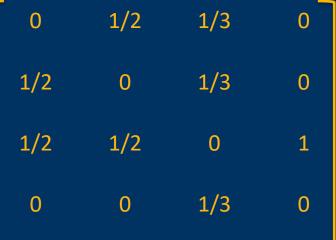


Visualization of 2

Random Walk Matrix

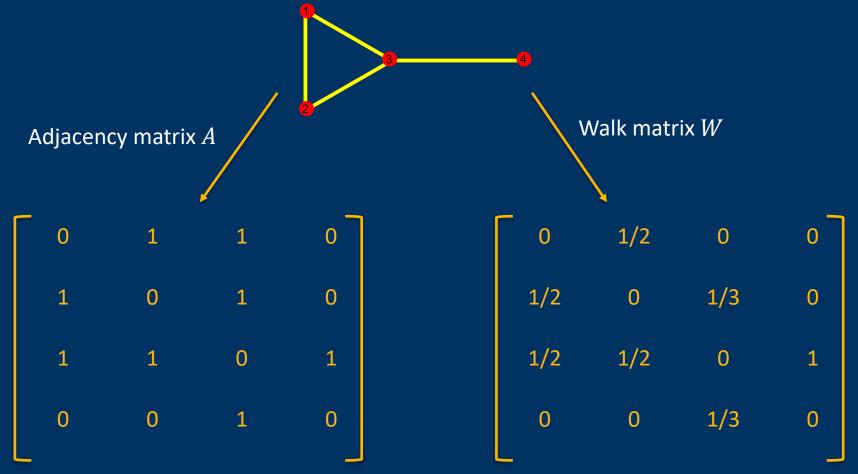
- We can encode the distribution of the random walk through its random walk matrix W, which is such that W_{vu} is the probability of transitioning from node u to node v.
- For a simple random walk, the probability of transitioning between two connected vertices *u* and *v* is the reciprocal of the degree of *u*.





Spectral Theory

The key thing about regular graphs is that their adjacency matrices are the same up to rescaling to the random walk matrix.



Theorem

We now consider any random walk matrix W of a graph with maximum degree Δ , and the boundary of $S \ \partial S$.

Lemma C (full): For any connected subgraph *S*, there is a vertex $u \in \partial S$ such that $\psi_S(u) \ge \frac{1}{\Delta |S|^2} \psi_S(t)$ for any vertex $t \in S$.

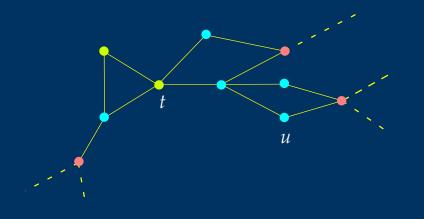
All of our results generalize to the random walk matrix.

- By the power method, $\lim_{r \to \infty} \frac{W_S^r \mathbf{1}_S}{\|W_S^r \mathbf{1}_S\|}$ approaches ψ . Therefore, $\lim_{r \to \infty} \frac{\mathbf{1}_S^T W_S^r e_u}{\mathbf{1}_S^T W_S^r e_t} = \frac{\psi(u)}{\psi(t)}$.
- $W_S^r e_u$ represents the random walk distribution for the walk remaining on *S*.
- $\frac{\mathbf{1}_{S}^{T}W_{S}^{r}e_{u}}{\mathbf{1}_{S}^{T}W_{S}^{r}e_{t}}$ is the ratio of the probabilities of walks starting at u and t remaining in our set for r steps.
- The probability that we remain in the set does not change if we contract all the points directly outside the boundary to a single point *s*.

Visualization

There's no difference in the probability we remain in the set if we contract all the points immediately outside the boundary to one point.

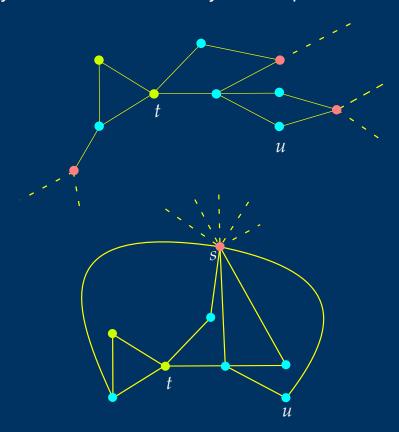
1)



Visualization

There's no difference in the probability we remain in the set if we contract all the points immediately outside the boundary to one point.

1)



• If we know the walk reaches t before it reaches s, and it takes j steps to reach t, then the probability that it stays within S is the probability a walk of length r - j starting at t stays within S.

- If we know the walk reaches t before it reaches s, and it takes j steps to reach t, then the probability that it stays within S is the probability a walk of length r j starting at t stays within S.
- Specifically, if *Y_j* is the event that the walk hits *t* before *s* AND hits *t* for the first time at step *j*, then

$$\mathbf{1}_{S}^{T}W_{S}^{r}e_{u} \geq \sum_{j=0}^{\prime} \Pr(Y_{j}) \mathbf{1}_{S}^{T}W_{S}^{r-j}e_{t} \geq \mathbf{1}_{S}^{T}W_{S}^{r}e_{t} \sum_{j=0}^{\prime} \Pr(Y_{j})$$

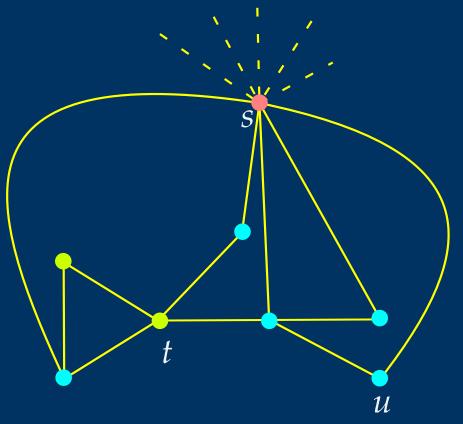
- If we know the walk reaches t before it reaches s, and it takes j steps to reach t, then the probability that it stays within S is the probability a walk of length r j starting at t stays within S.
- Specifically, if *Y_j* is the event that the walk hits *t* before *s* AND hits *t* for the first time at step *j*, then

$$\mathbf{1}_{S}^{T}W_{S}^{r}e_{u} \geq \sum_{j=0}^{\prime} \Pr(Y_{j}) \mathbf{1}_{S}^{T}W_{S}^{r-j}e_{t} \geq \mathbf{1}_{S}^{T}W_{S}^{r}e_{t} \sum_{j=0}^{\prime} \Pr(Y_{j})$$

• As the expected hitting time of t is finite, $\lim_{r \to \infty} (\sum_{j=0}^{r} \Pr(Y_j))$ is the probability that t is hit before s.

Visualization

A random walk of length r conditioned on reaching t before s that reaches t for the first time at step j has the same probability of staying within the set as a random walk starting at t of length r - j.



Electric Flow¹⁾

This probability is equivalent to the voltage of u, denoted V(u), in a flow from s to t where the voltages V(t) = 1, V(s) = 0 [e.g. Bollobás].

Electric Flow

This probability is equivalent to the voltage of u, denoted V(u), in a flow from s to t where the voltages V(t) = 1, V(s) = 0 [e.g. Bollobás].

1)

By Ohm's law, the total current from s to t is at least 1/dist(s,t). Because s has at most |S|Δ neighbors, there is some neighbor of s such that current through this vertex is at least 1/Δ|S|². Therefore, the voltage of this vertex is at least 1/Δ|S|², as is the probability of reaching t before s.

Electric Flow

This probability is equivalent to the voltage of u, denoted V(u), in a flow from s to t where the voltages V(t) = 1, V(s) = 0 [e.g. Bollobás].

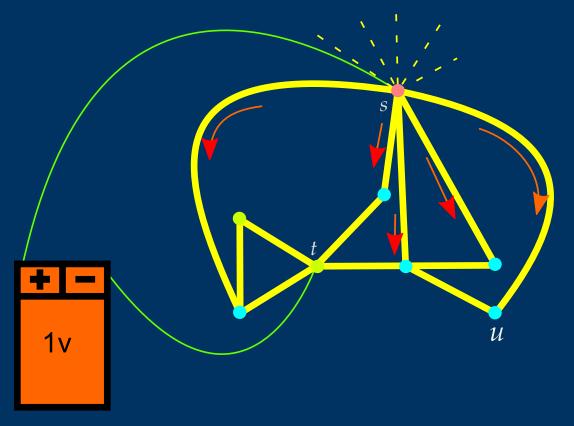
1)

- By Ohm's law, the total current from s to t is at least 1/dist(s,t). Because s has at most |S|Δ neighbors, there is some neighbor of s such that current through this vertex is at least 1/Δ|S|². Therefore, the voltage of this vertex is at least 1/Δ|S|², as is the probability of reaching t before s.
- For some vertex $u \in \partial S$, $\frac{\psi_S(u)}{\psi_S(t)} \ge \frac{1}{\Delta |S|^2}$. As we can assume $\psi_S(t) \ge 1/\sqrt{S}$,

 $\psi_S(u) \ge 1/(\Delta |S|^{5/2}).$

Electric Flow

We interpret this probability as an electrical current between s and t. There must be one vertex adjacent to s that receives a large current flow and therefore has high eigenvector value. We make this our u.



Questions

- Is Ω(k^{1/5}) tight for the typical support of a walk on a bounded degree graph.
- Is there an $\epsilon > 0$ such that the random walk matrix of every graph has second eigenvalue multiplicity $O(n^{1-\epsilon})$?
 - The most we know is that there are graphs with second eigenvalue multiplicity $\Omega(n^{1/3})$.

Thank you!