Walk-classes, Centrality Collisions, and Spider Donuts

Speaker: Kyle Kloster

With Co-authors
Blair Sullivan & Eric Horton

NC State University

Preliminaries

Central theme:

How do walk structure and centrality behavior relate?

Preliminaries

walk structure **centrality behavior**

$$\left[e^{\beta \boldsymbol{A}}\right]_{jj}$$ Subgraph centrality

$$[(\boldsymbol{I} - \beta \boldsymbol{P})^{-1}\mathbf{e}_S]_j$$ Personalized PageRank

$$\left[e^{\beta \boldsymbol{A}}\mathbf{1}\right]_j$$ Total communicability

Preliminaries

	walk structure	centrality behavior	
Closed L-walks	$\left[\boldsymbol{A}^{\ell}\mathbf{e}_j\right]_j$	$\left[e^{\beta\boldsymbol{A}}\right]_{jj}$	Subgraph centrality
Transition probabilities	$\left[\boldsymbol{P}^{\ell}\mathbf{e}_S\right]_j$	$[(\boldsymbol{I}-\beta\boldsymbol{P})^{-1}\mathbf{e}_S]_j$	Personalized PageRank
All L-walks	$\left[\boldsymbol{A}^{\ell}\mathbf{1}\right]_j$	$\left[e^{\beta\boldsymbol{A}}\mathbf{1}\right]_j$	Total communicability

Preliminaries

	walk structure	centrality behavior	
Closed *L*-walks	$\left[\boldsymbol{A}^{\ell}\mathbf{e}_j\right]_j$	$\left[e^{\beta\boldsymbol{A}}\right]_{jj}$	Subgraph centrality
Transition probabilities	$\left[\boldsymbol{P}^{\ell}\mathbf{e}_S\right]_j$	$[(\boldsymbol{I}-\beta\boldsymbol{P})^{-1}\mathbf{e}_S]_j$	Personalized PageRank
All *L*-walks	$\left[\boldsymbol{A}^{\ell}\mathbf{1}\right]_j$	$\left[e^{\beta\boldsymbol{A}}\mathbf{1}\right]_j$	Total communicability
General	$\left[\boldsymbol{M}^{\ell}\mathbf{s}(j)\right]_j$	$[f(\beta\boldsymbol{M})\mathbf{s}(j)]_j$	

Preliminaries

How do walk structure and centrality behavior relate?

Fix a graph matrix	M	diagonalizable
a vector function	$\mathbf{s}(\cdot)$	e.g. $\mathbf{1}, \mathbf{e}_j, \mathbf{e}_S$
and suitable function	$f(\cdot)$	pos power series

Preliminaries

How do walk structure and centrality behavior relate?

Fix a graph matrix	M	diagonalizable
a vector function	$\mathbf{s}(\cdot)$	e.g. $\mathbf{1}, \mathbf{e}_j, \mathbf{e}_S$
and suitable function	$f(\cdot)$	pos power series

Positive Power Series:

An analytic function with a power series with all positive coefficients: $(1 - \beta x)^{-1}, \quad e^{\beta x}$

Preliminaries

How do walk structure and centrality behavior relate?

Fix a graph matrix	$\boldsymbol{M}$	diagonalizable
a vector function	$\mathbf{s}(\cdot)$	e.g. $\mathbf{1}, \mathbf{e}_j, \mathbf{e}_S$
and suitable function	$f(\cdot)$	pos power series

Walk centrality

$$C_f(j, \beta) = [f(\beta \boldsymbol{M})\mathbf{s}(j)]_j$$

$C(\cdot, \beta)$ when context is clear

Questions

How do walk structure and centrality behavior relate?

$$C_f(j,\beta) = \left[f(\beta \boldsymbol{M})\mathbf{s}(j)\right]_j$$

- when can *i,j* have same score?
  - what does it mean if they do?
  - when can we say *i,j* never have same score?
- If *i,j* "walk the same" how do scores compare?

Questions

How do walk structure and centrality behavior relate?

$$C_f(j, \beta) = \left[f(\beta \boldsymbol{M})\mathbf{s}(j)\right]_j$$

- when can *i,j* have same score?
  - what does it mean if they do?
  - when can we say *i,j* never have same score?
- If *i,j* “walk the same” how do scores compare?

Rank-trajectory plot

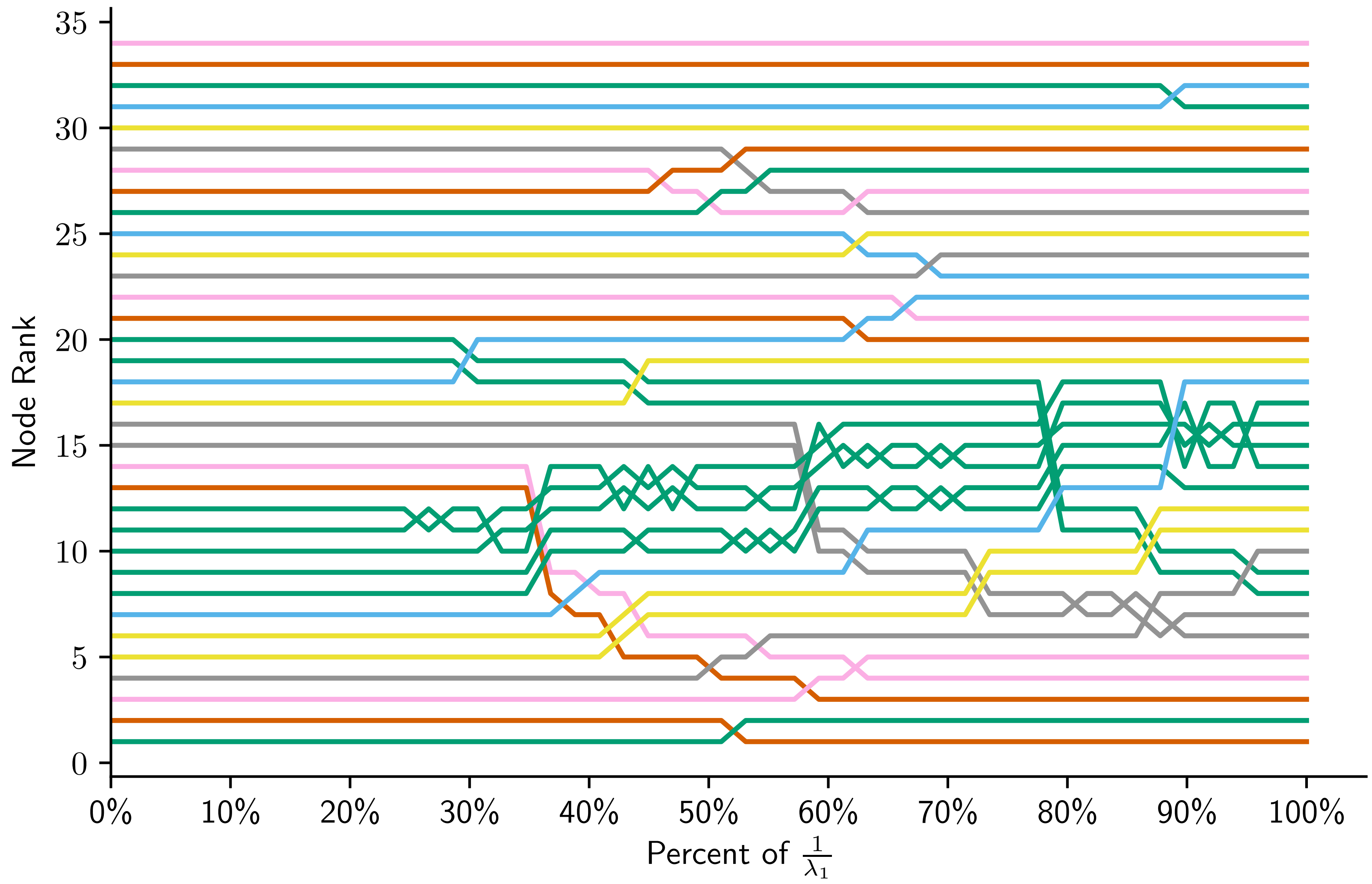
Each curve represents a node in Zachary's Karate Club

x : Katz parameter

$$\beta \in (0, \frac{1}{\lambda_1})$$

y : *n* - node's rank according to

$$C_{Katz}(\cdot, \beta)$$



Rank-trajectory plot

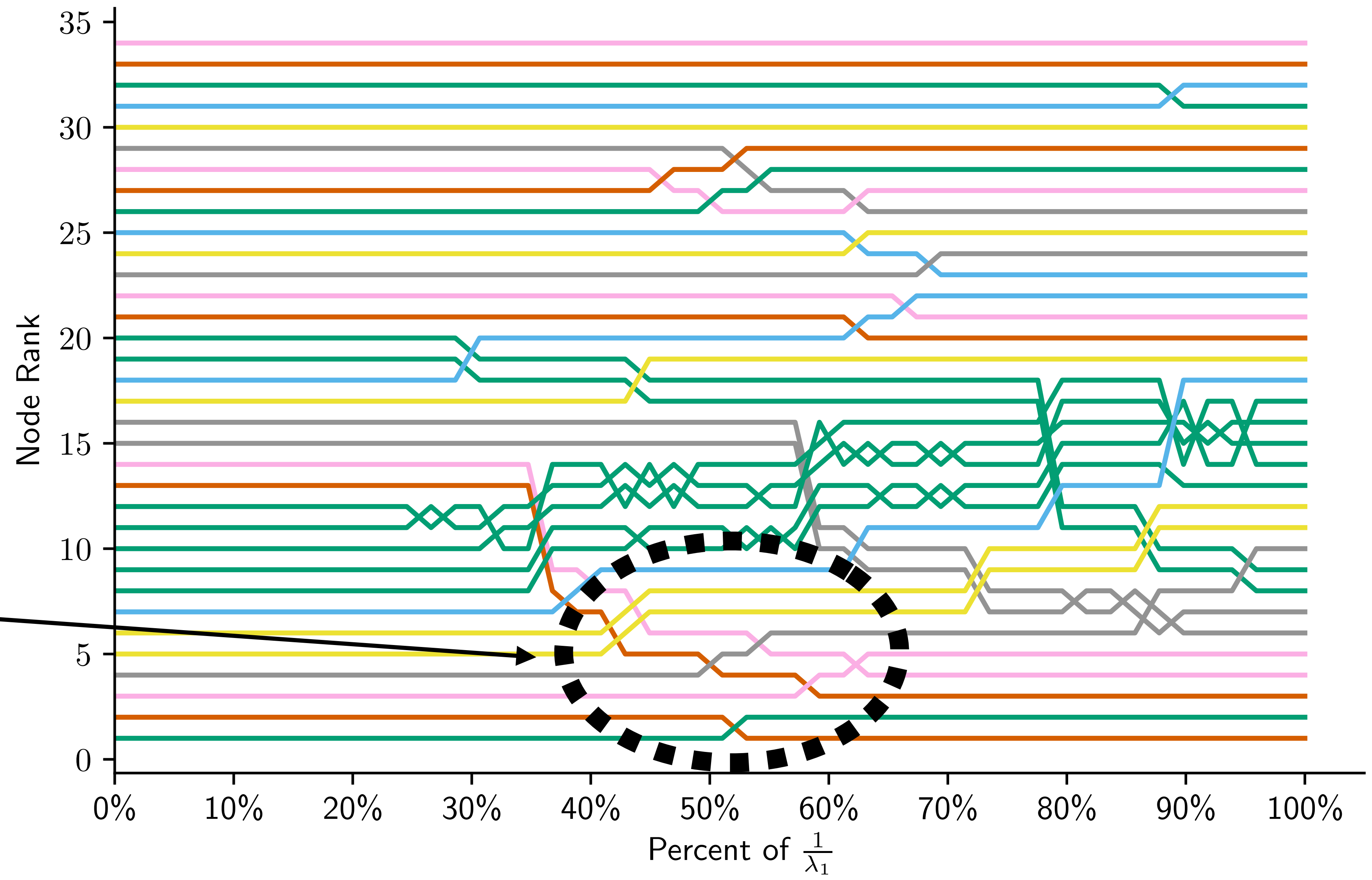
As β varies, node rankings vary. When nodes *i,j* swap rank,

by continuity,

$$C_f(j, \beta) = C_f(i, \beta)$$

at some value β.

We call these **collisions.**



Observations about questions…

- when can *i,j* have same score?
    - what does it mean if they do?
    - when can we say *i,j* never have same score?

Observations about questions…

- when can *i,j* have same score?
  - what does it mean if they do?
  - when can we say *i,j* never have same score?

If *i,j* never collide, then $C_f(j,\beta) > C_f(i,\beta)$ for all β , i.e., *j* is always ranked above *i.*

Sufficient Condition: If for $\ell \in \mathbb{N}$ $\left[\boldsymbol{M}^{\ell}\mathbf{s}(j)\right]_j > \left[\boldsymbol{M}^{\ell}\mathbf{s}(i)\right]_i$
Then for all *f* and β , *j* ranks above *i*. We say node *j* **majorizes** node *i*.

Nodes that majorize always rank higher

Sufficient Condition: If for $\ell \in \mathbb{N}$ $\left[\boldsymbol{M}^{\ell}\mathbf{s}(j)\right]_j > \left[\boldsymbol{M}^{\ell}\mathbf{s}(i)\right]_i$

Then for all *f* and β, *j* ranks above *i*. We say node *j* **majorizes** node *i*.

Proof sketch:

Nodes that majorize always rank higher

Sufficient Condition: If for $\ell \in \mathbb{N}$ $\left[\boldsymbol{M}^{\ell}\mathbf{s}(j)\right]_j > \left[\boldsymbol{M}^{\ell}\mathbf{s}(i)\right]_i$

Then for all *f* and β, *j* ranks above *i*. We say node *j* **majorizes** node *i*.

Proof sketch:

$$C_f(j,\beta) = \sum_{\ell=0}^{\infty} f_\ell \beta^\ell \left[\boldsymbol{M}^{\ell}\mathbf{s}(j)\right]_j$$

$$C_f(i,\beta) = \sum_{\ell=0}^{\infty} f_\ell \beta^\ell \left[\boldsymbol{M}^{\ell}\mathbf{s}(i)\right]_i$$

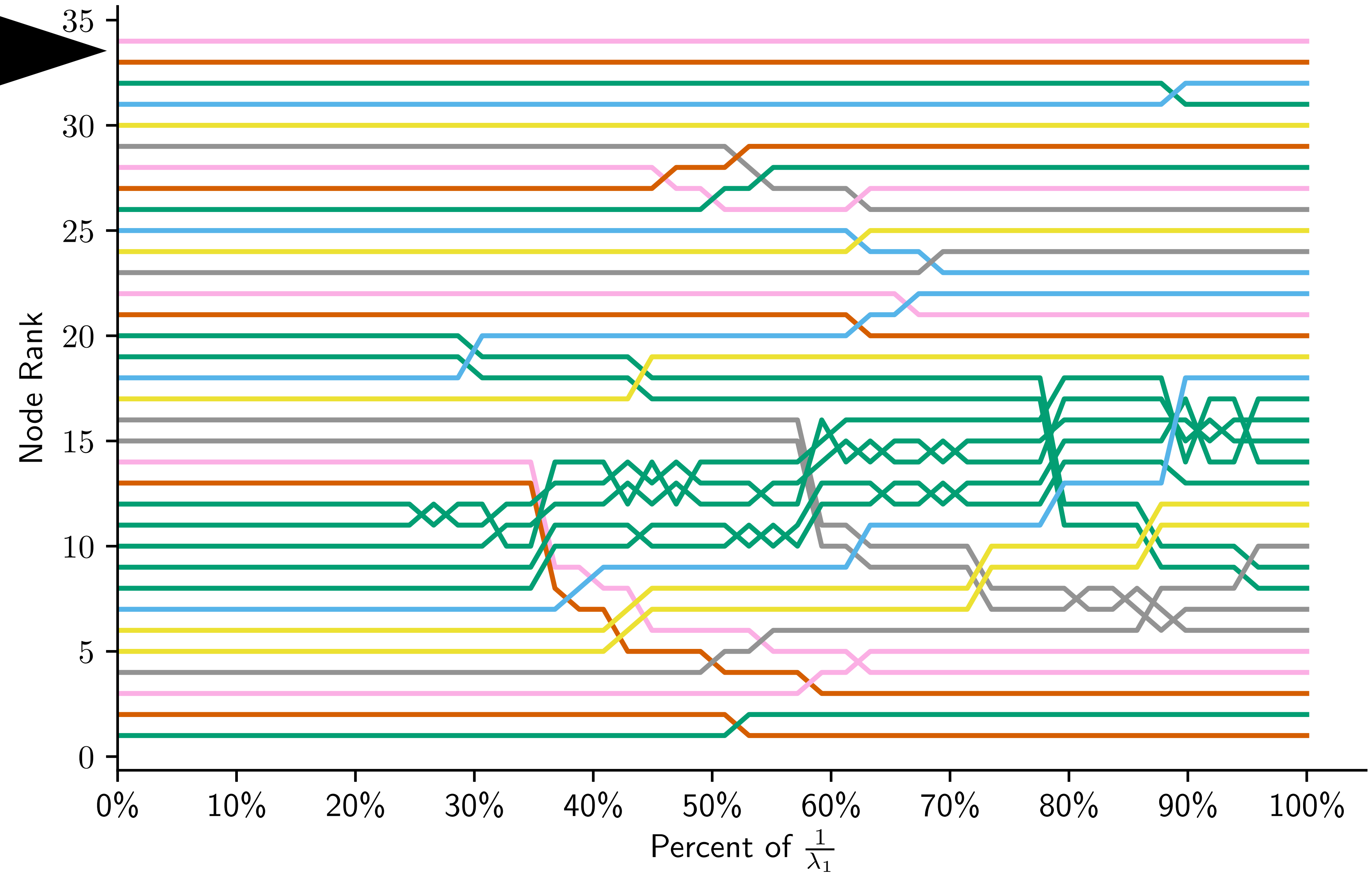
Nodes that majorize always rank higher

Sufficient Condition: If for $\ell \in \mathbb{N}$ $\left[\boldsymbol{M}^\ell \mathbf{s}(j)\right]_j > \left[\boldsymbol{M}^\ell \mathbf{s}(i)\right]_i$

Then for all *f* and β, *j* ranks above *i*. We say node *j* **majorizes** node *i*.

Nodes ranked 1,2 in Karate each majorize nodes 4 through *n*

These nodes rank in top 3 for all *f*, β



“Partial” majorizing

What if *j* doesn’t have more *k*-walks for **all** *k*, just some?

“Partial” majorizing

What if *j* doesn’t have more *k*-walks for **all** *k*, just some?

[Horton, K., Sullivan]

Theorem. *Let G be undirected and unweighted, and suppose nodes i and j satisfy $[\boldsymbol{A}^{\ell}\boldsymbol{1}]_j > [\boldsymbol{A}^{\ell}\boldsymbol{1}]_i$ for $\ell = 1, \cdots, t$. Then the Katz Centrality for nodes i and j satisfies $C(j, \beta) > C(i, \beta)$ for all $\beta \in \left(0, \frac{1}{\lambda_1} \cdot \frac{1}{(2\sqrt{n}\lambda_1)^{1/t}}\right)$.*

Recall: Katz domain is $\left(0, \frac{1}{\lambda_1}\right)$

“Partial” majorizing

What if *j* doesn’t have more *k*-walks for **all** *k,* just some?

[Horton, K., Sullivan]

Theorem. *Let G be undirected and unweighted, and suppose nodes i and j satisfy $[\boldsymbol{A}^{\ell}\boldsymbol{1}]_j > [\boldsymbol{A}^{\ell}\boldsymbol{1}]_i$ for $\ell = 1, \cdots, t$. Then the Katz Centrality for nodes i and j satisfies $C(j,\beta) > C(i,\beta)$ for all $\beta \in \left(0, \frac{1}{\lambda_1} \cdot \frac{1}{(2\sqrt{n}\lambda_1)^{1/t}}\right)$.*

Corollary: if d(*j*) > d(*i*) then *j* ranks above *i* for β near *0*

"Partial" majorizing

What if *j* doesn't have more *k*-walks for **all** *k,* just some?

[Horton, K., Sullivan]

Theorem. *Let G be undirected and unweighted, and suppose nodes i and j satisfy* $[\boldsymbol{A}^\ell \boldsymbol{1}]_j > [\boldsymbol{A}^\ell \boldsymbol{1}]_i$ *for* $\ell = 1, \cdots, t$. *Then the Katz Centrality for nodes i and j satisfies* $C(j,\beta) > C(i,\beta)$ *for all* $\beta \in \left(0, \frac{1}{\lambda_1} \cdot \frac{1}{(2\sqrt{n}\lambda_1)^{1/t}}\right)$.

Corollary: if d(*j*) > d(*i*) then *j* ranks above *i* for β near *0*

Theorem: [Klymko, Benzi 2015] "On the Limiting behavior…"

For *f*, **M**, nonnegative **s**,

$$C_f(\cdot,\beta) \to \begin{cases} \text{degree centrality, as } \beta \to 0^+ \\ \text{eigenvector centrality, as } \beta \to \beta_{max}^- \end{cases}$$

"Partial" majorizing

What if *j* doesn't have more *k*-walks for **all** *k,* just some?

[Horton, K., Sullivan]

Theorem. *Let G be undirected and unweighted, and suppose nodes i and j satisfy* $[\boldsymbol{A}^{\ell}\boldsymbol{1}]_j > [\boldsymbol{A}^{\ell}\boldsymbol{1}]_i$ *for* $\ell = 1, \cdots, t$*. Then the Katz Centrality for nodes i and j satisfies* $C(j,\beta) > C(i,\beta)$ *for all* $\beta \in \left(0, \frac{1}{\lambda_1} \cdot \frac{1}{(2\sqrt{n}\lambda_1)^{1/t}}\right)$.

Our bound makes explicit the rate of convergence here:

Theorem: [Klymko, Benzi 2015] "On the Limiting behavior…"

For *f*, **M**, nonnegative **s**,

$$C_f(\cdot,\beta) \to \begin{cases} \text{degree centrality, as } \beta \to 0^+ \\ \text{eigenvector centrality, as } \beta \to \beta_{max}^- \end{cases}$$

Questions

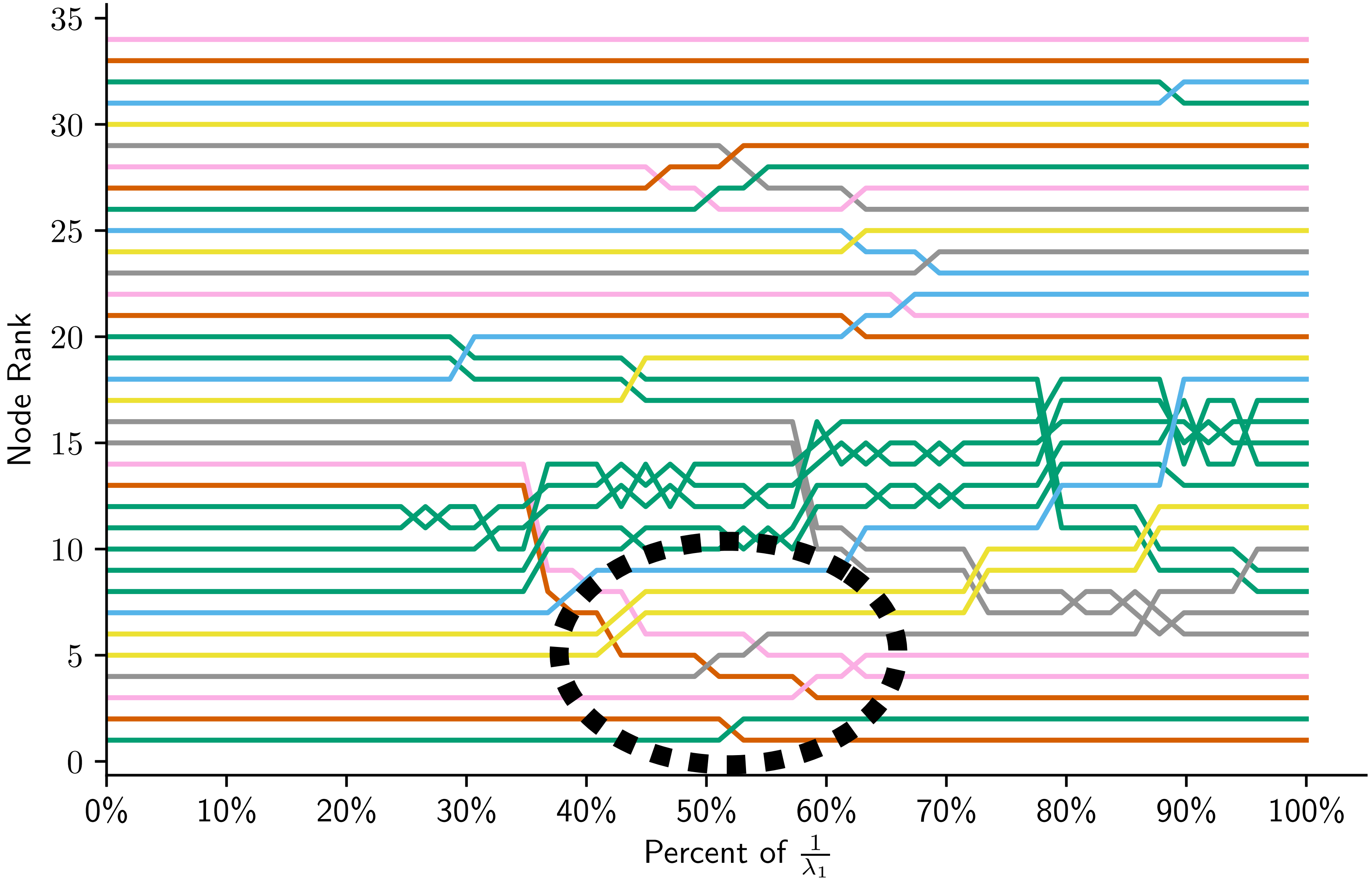
- when can *i,j* have same score?
  - what does it mean if they do?
  - when can we say *i,j* never have same score?

Back to collisions…

Values β where

$$C_f(j, \beta) = C_f(i, \beta)$$

How many times can this occur?



Number of collisions is bounded

Theorem: [Benzi, 2014]
If $e^{\beta \boldsymbol{A}}_{jj} = e^{\beta \boldsymbol{A}}_{ii}$ for all β in a set with an accumulation point, then it holds for all β .

Number of collisions is bounded

Theorem: [Benzi, 2014]

If $e^{\beta \boldsymbol{A}}_{jj} = e^{\beta \boldsymbol{A}}_{ii}$ for all β in a set with an accumulation point, then it holds for all β .

Theorem: [Horton, K., Sullivan]

If $C_f(j, \beta) = C_f(i, \beta)$ for an infinite number of β, then it holds for all β in the domain of *f*.

Number of collisions is bounded

Theorem: [Benzi, 2014]
If $e^{\beta \boldsymbol{A}}_{jj} = e^{\beta \boldsymbol{A}}_{ii}$ for all β in a set with an accumulation point,
then it holds for all β .

Theorem: [Horton, K., Sullivan]
If $C_f(j, \beta) = C_f(i, \beta)$ for an infinite number of β,
then it holds for all β in the domain of *f*.

Proof: We show *C* is analytic and use the Identity Theorem.

Number of collisions is bounded

Theorem: [Benzi, 2014]

If $e^{\beta \boldsymbol{A}}_{jj} = e^{\beta \boldsymbol{A}}_{ii}$ for all β in a set with an accumulation point,

then it holds for all β .

Theorem: [Horton, K., Sullivan]

If $C_f(j, \beta) = C_f(i, \beta)$ for an infinite number of β,

then it holds for all β in the domain of *f*.

Proof: We show *C* is analytic and use the Identity Theorem.

Corollary:

This proves that the number of distinct rankings $C_f(\cdot, \cdot)$

Produces is finite!

Number of collisions is bounded

Degree of min poly

Theorem: [Horton, K., Sullivan]

If *f* is the resolvent, $(1-\beta x)^{-1}$, then *i,j* collide <= *m-1* times, unless they collide for all β .

Number of collisions is bounded

Degree of min poly

Theorem: [Horton, K., Sullivan]

If *f* is the resolvent, $(1-\beta x)^{-1}$, then *i,j* collide <= *m-1* times, unless they collide for all β .

Corollary: [Horton, K., Sullivan]

PageRank and Katz induce at most O(n^3) distinct rankings.

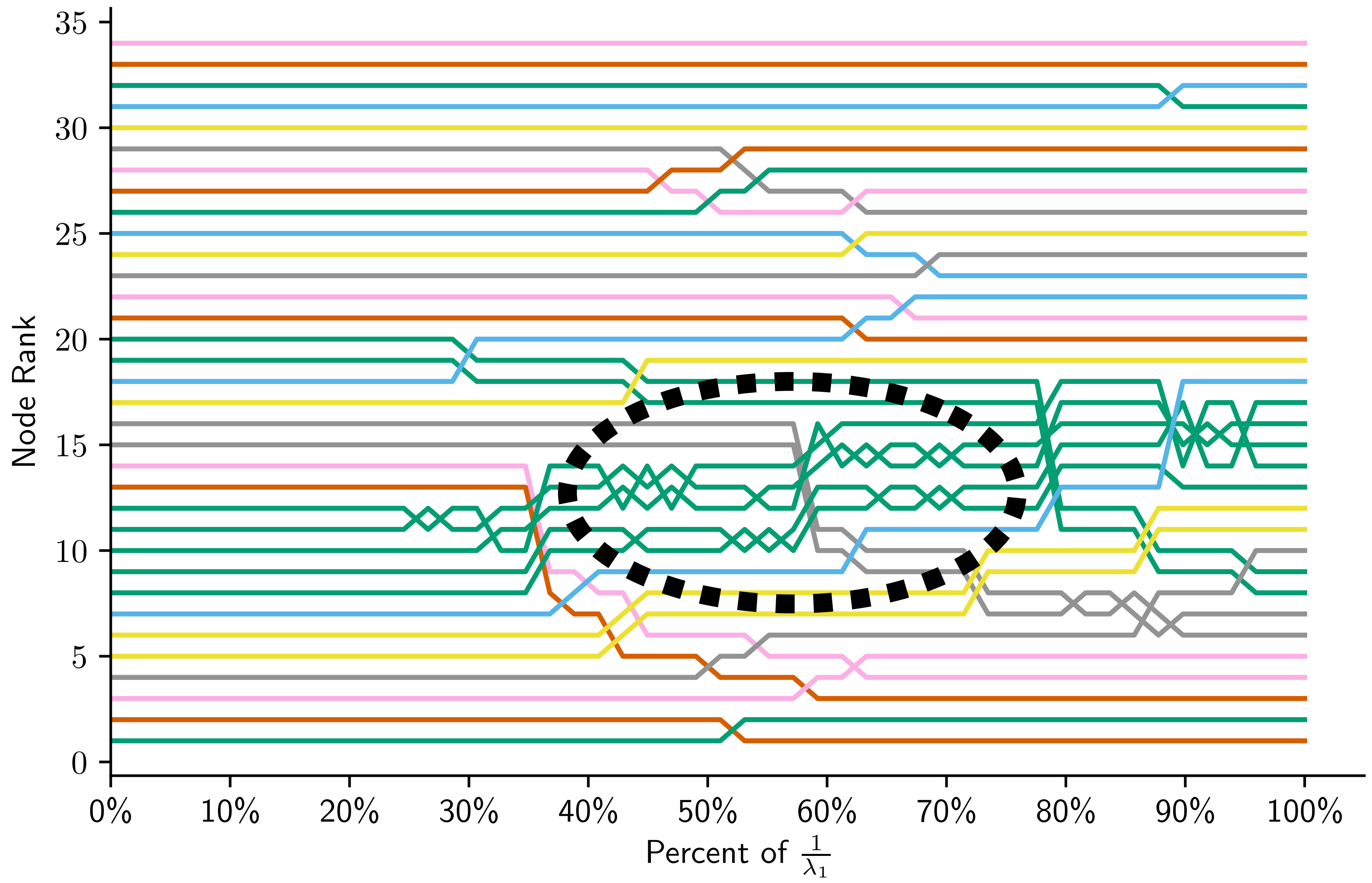
- Trivial upper bound is (n!)
- Do other *f* induce more rankings? Is that better?
- Find a bound for f(x) = e^x

Too many collisions…

What causes *i,j* to have "too many" collisions?

$$C_f(j, \beta) = C_f(i, \beta)$$

Holds for all β …



Questions

How do walk structure and centrality behavior relate?

$$C_f(j,\beta) = \left[f(\beta \boldsymbol{M})\mathbf{s}(j)\right]_j$$

- when can *i,j* have same score?
  - what does it mean if they do?
  - when can we say *i,j* never have same score?
- If *i,j* “walk the same” how do scores compare?

Questions

How do walk structure and centrality behavior relate?

$$C_f(j,\beta) = [f(\beta \boldsymbol{M})\mathbf{s}(j)]_j$$

- when can *i,j* have same score?
  - what does it mean if they do?
  - when can we say *i,j* never have same score?
- If *i,j* "walk the same" how do scores compare?

Nodes *i,j* have "too many" collisions *iff* same walk structure

Walk structure determines centrality behavior

$$C(j, \beta) = \sum_{k=0}^{\infty} f_k \beta^k \left[\boldsymbol{M}^k \mathbf{s}(j) \right]_j$$

Walk structure determines centrality behavior

$$C(j, \beta) = \sum_{k=0}^{\infty} f_k \beta^k \left[\boldsymbol{M}^k \mathbf{s}(j) \right]_j$$

Walk structure determines centrality behavior

$$C(j,\beta) = \sum_{k=0}^{\infty} f_k \beta^k \left[\boldsymbol{M}^k \mathbf{s}(j)\right]_j$$

degree of
min poly of **M**

Proposition:

If for k = 0, … , m-1

$$\left[\boldsymbol{M}^k \mathbf{s}(j)\right]_j = \left[\boldsymbol{M}^k \mathbf{s}(i)\right]_i$$

then for all f and β

$$C_f(j,\beta) \equiv C_f(i,\beta)$$

Walk structure determines centrality behavior

Identical **walk structure**…

→ Identical **walk centrality**

Proposition:

degree of min poly of **M**

If for $k = 0, \ldots, m-1$

$$\left[\boldsymbol{M}^k \mathbf{s}(j)\right]_j = \left[\boldsymbol{M}^k \mathbf{s}(i)\right]_i$$

then for all f and β

$$C_f(j,\beta) \equiv C_f(i,\beta)$$

Walk structure determines centrality behavior

Identical **walk structure**…

⟷ Identical **walk centrality**

Theorem: [Horton, K., Sullivan]

If $C_f(j,\beta) = C_f(i,\beta)$ for an infinite number of β,

then it holds for all β in the domain of *f*.

Walk-classes and iso-centrality

Definition:

A **walk-class** = all nodes in G with identical "walk tuples"

$$\left(\left[\boldsymbol{M}^{0} \mathbf{s}(j) \right]_{j}, \left[\boldsymbol{M}^{1} \mathbf{s}(j) \right]_{j}, \cdots, \left[\boldsymbol{M}^{m-1} \mathbf{s}(j) \right]_{j} \right)$$

Walk-classes and iso-centrality

Definition:

A **walk-class** = all nodes in G with identical "walk tuples"

$$\left(\left[\boldsymbol{M}^0 \mathbf{s}(j) \right]_j, \left[\boldsymbol{M}^1 \mathbf{s}(j) \right]_j, \cdots, \left[\boldsymbol{M}^{m-1} \mathbf{s}(j) \right]_j \right)$$

Theorem: [Horton, K., Sullivan] The following are equivalent:

1. Nodes *i,j* are in the same walk-class
2. Nodes *i,j* have >= *m* Katz collisions
3. Nodes *i,j* have identical **walk centrality:**

$C_f(j, \beta) = C_f(i, \beta)$ for all *f* and β

Walk-classes and iso-centrality

Definition:

We say nodes satisfying these properties are **isocentral**.

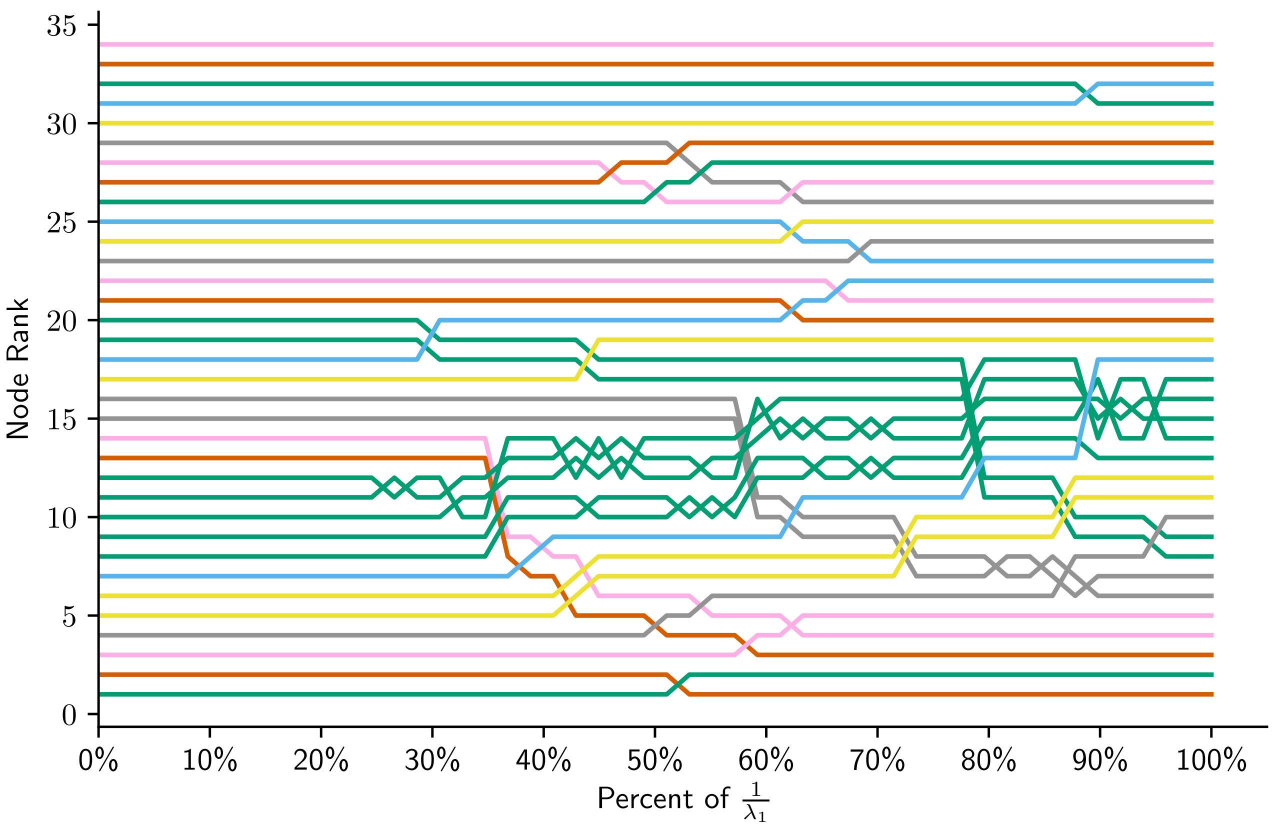
Theorem: [Horton, K., Sullivan] The following are equivalent:

1. Nodes *i,j* are in the same walk-class
2. Nodes *i,j* have >= *m* Katz collisions
3. Nodes *i,j* have identical **walk centrality:**

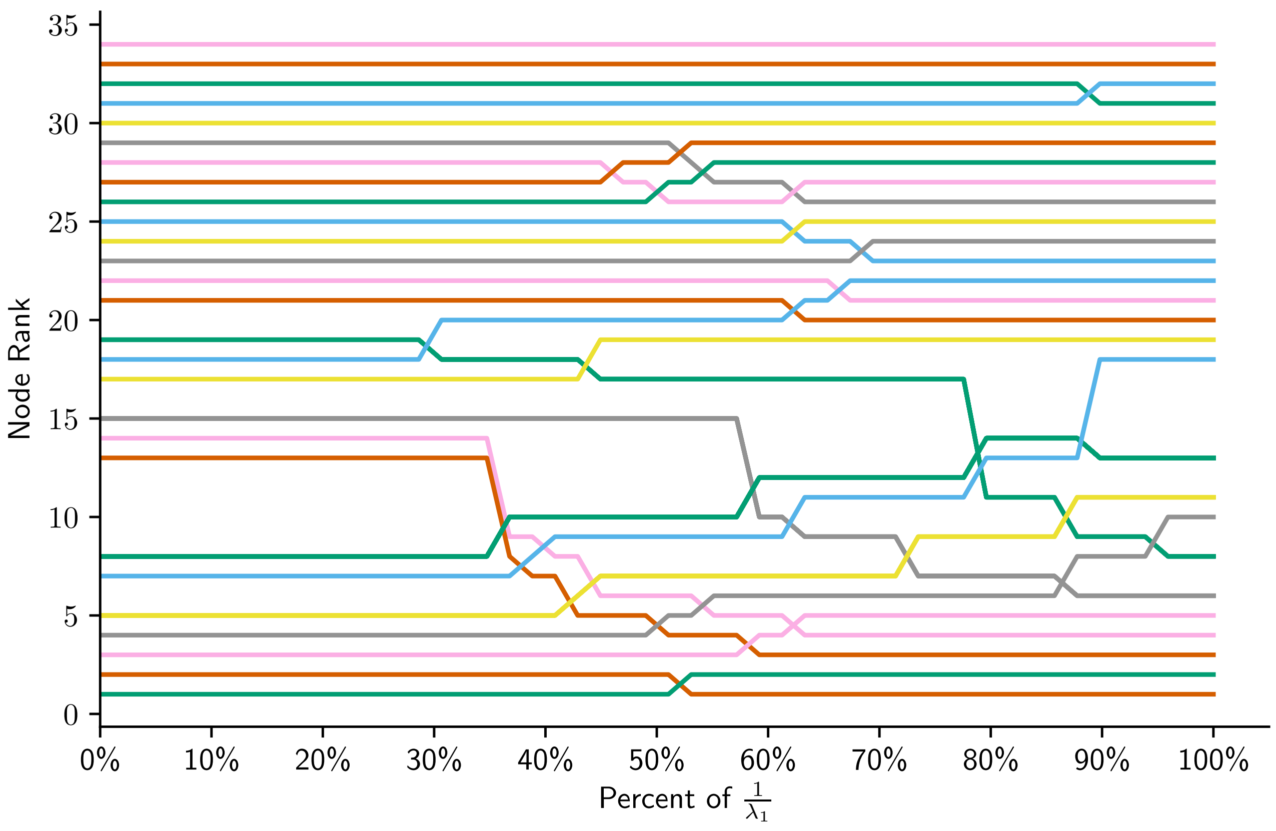
$C_f(j, \beta) = C_f(i, \beta)$ for all *f* and β

Walk-classes in Zachary's Karate Club

Rank-plot for Katz

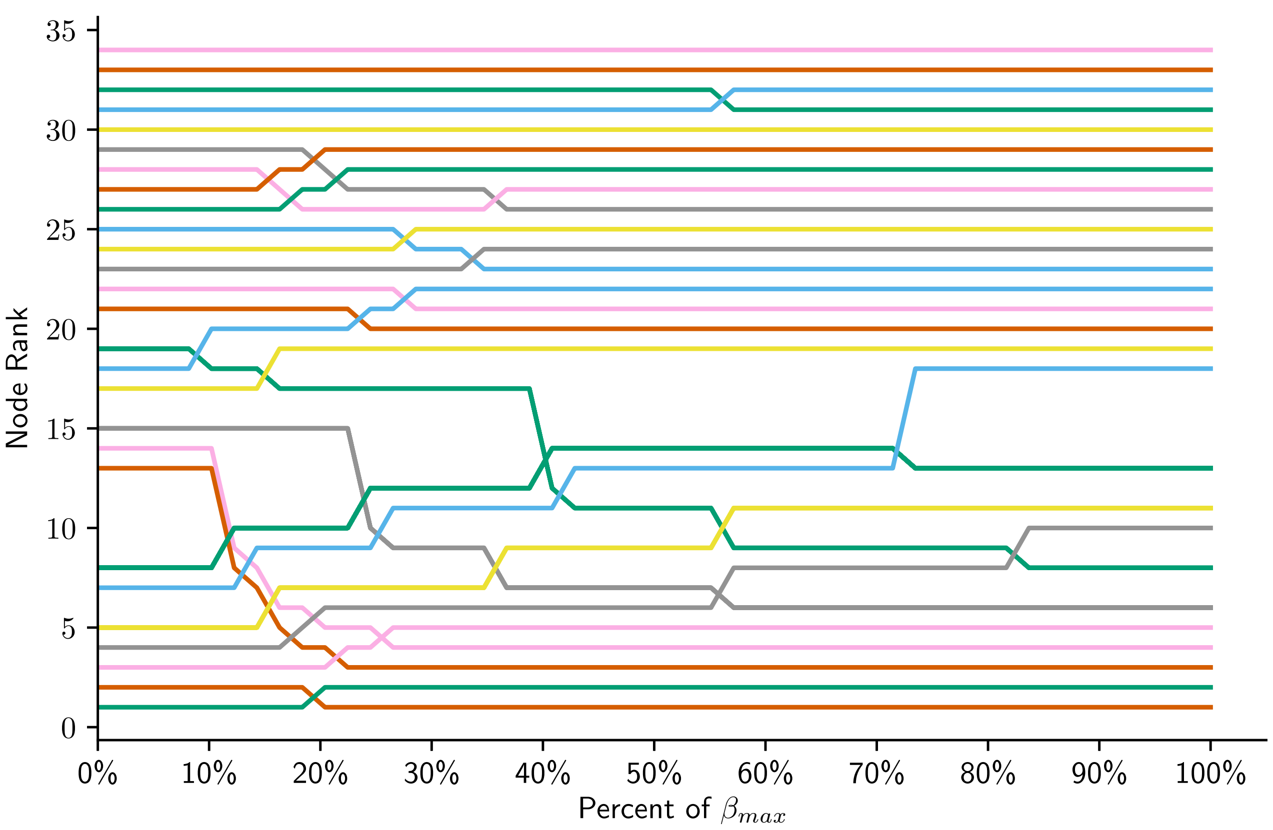


Katz, walk-classes merged

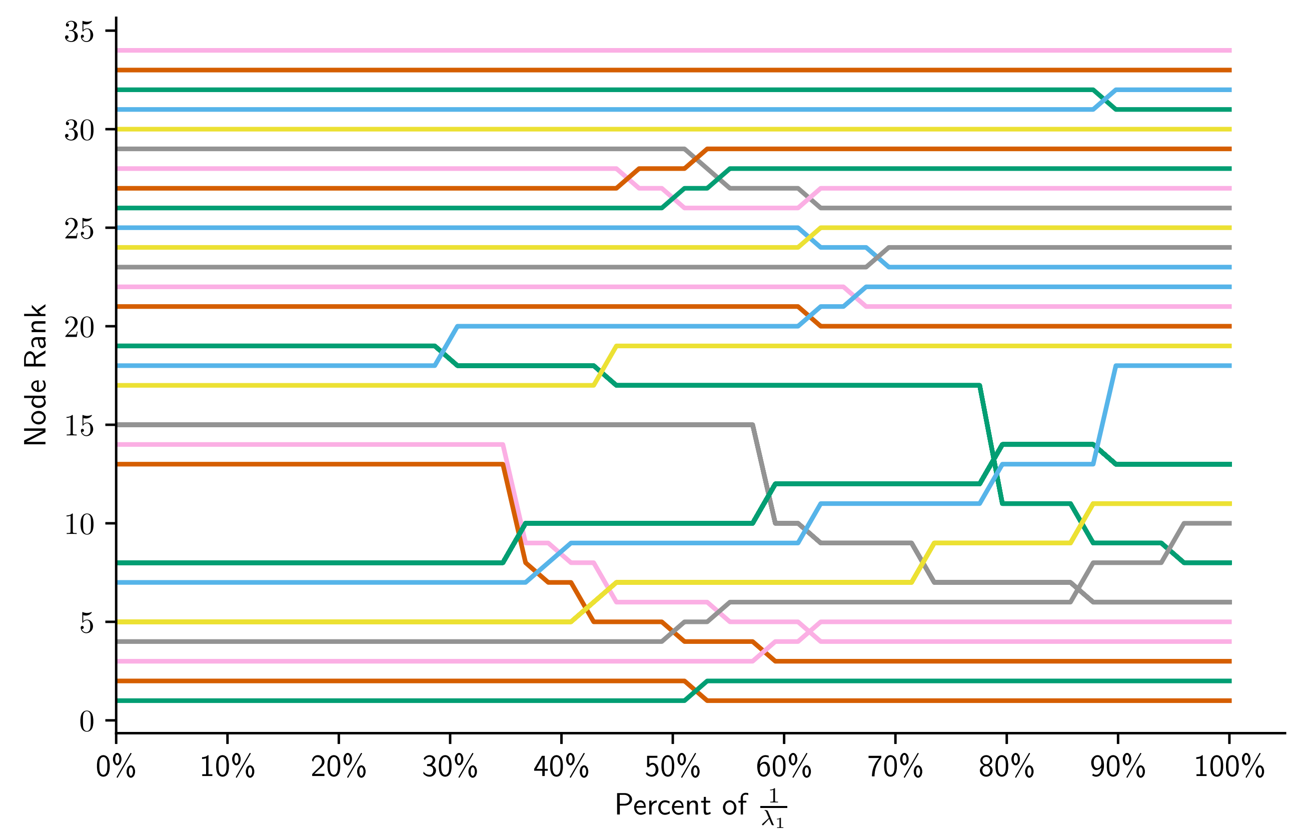


Walk-classes in Zachary's Karate Club

Total communicability, merged



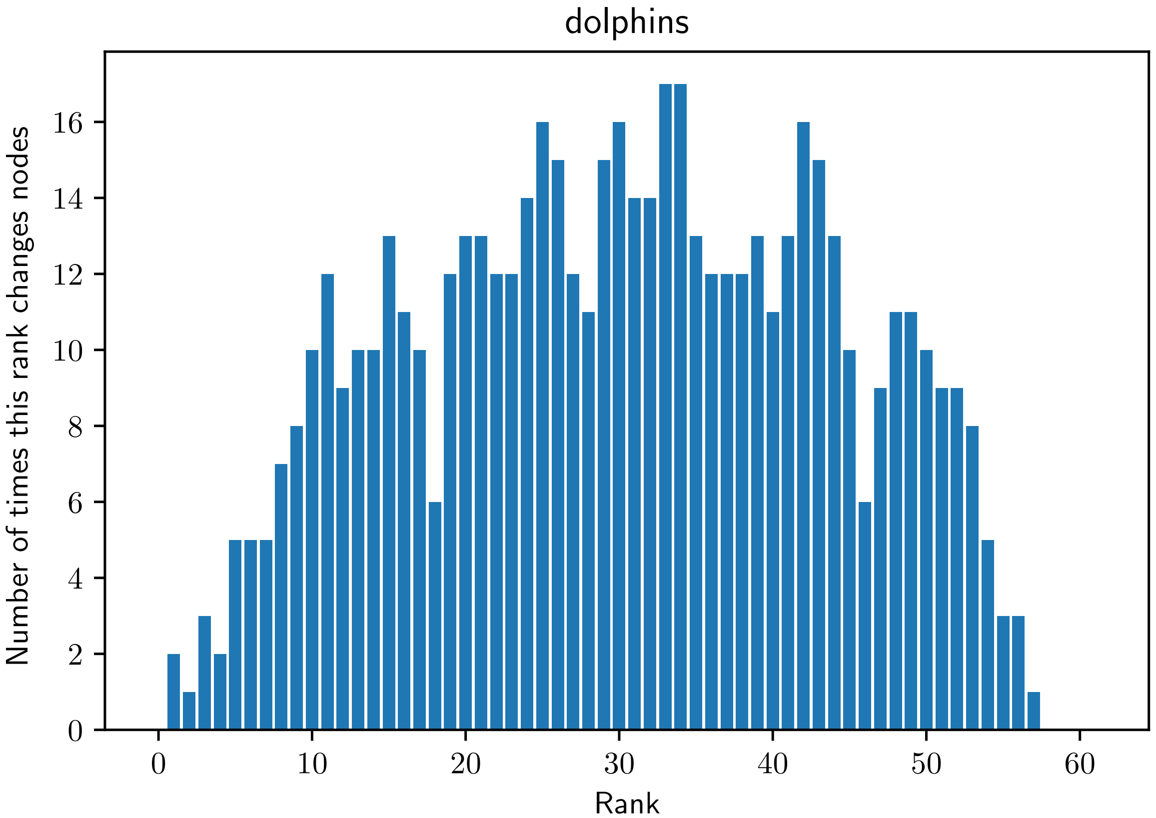
Katz, walk-classes merged

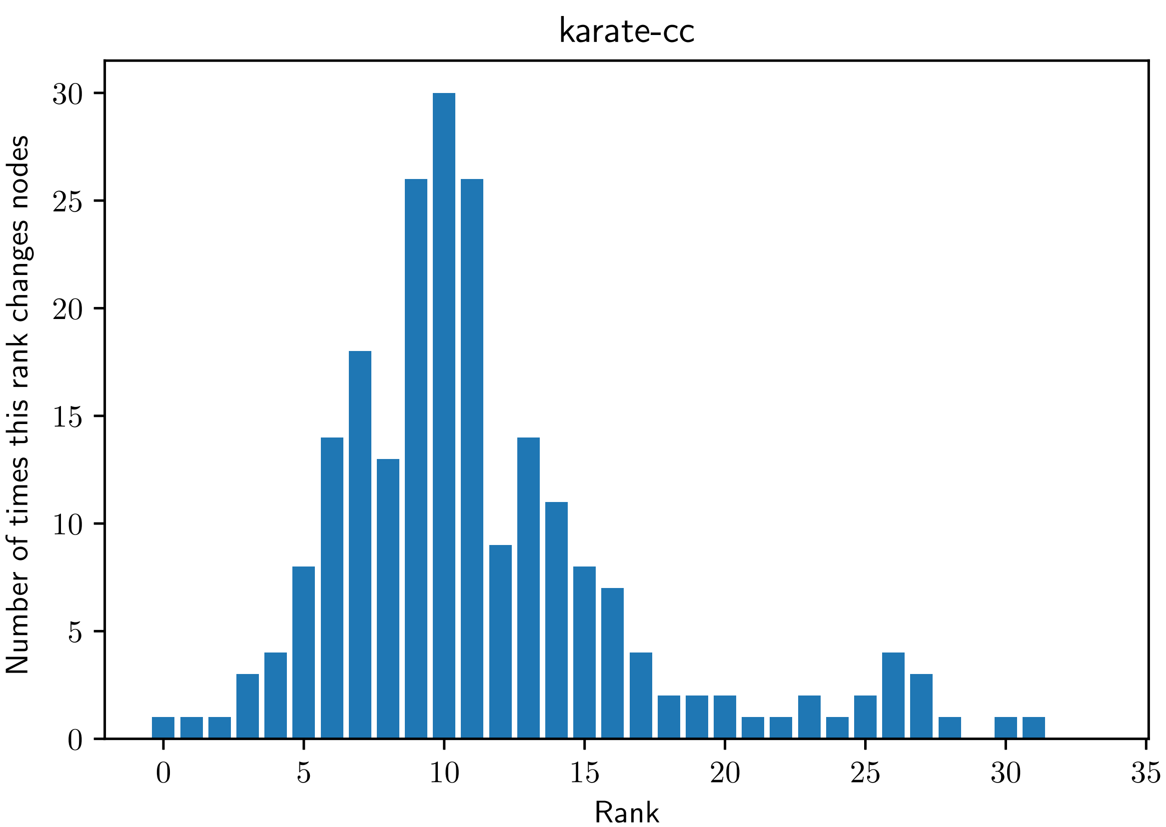


Same node curves are merged in both plots because walk-classes are regardless of *f*

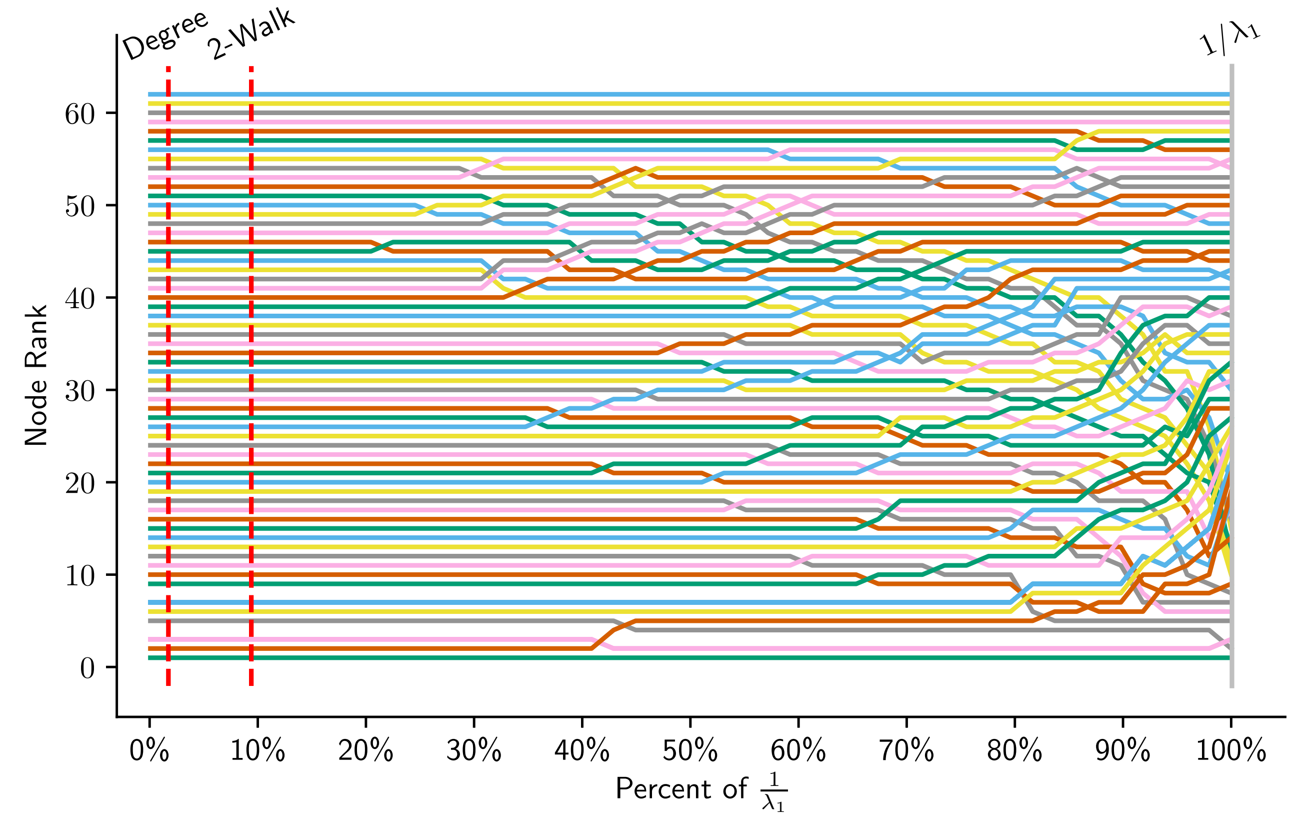
Collisions are abundant; Iso-centrality is not

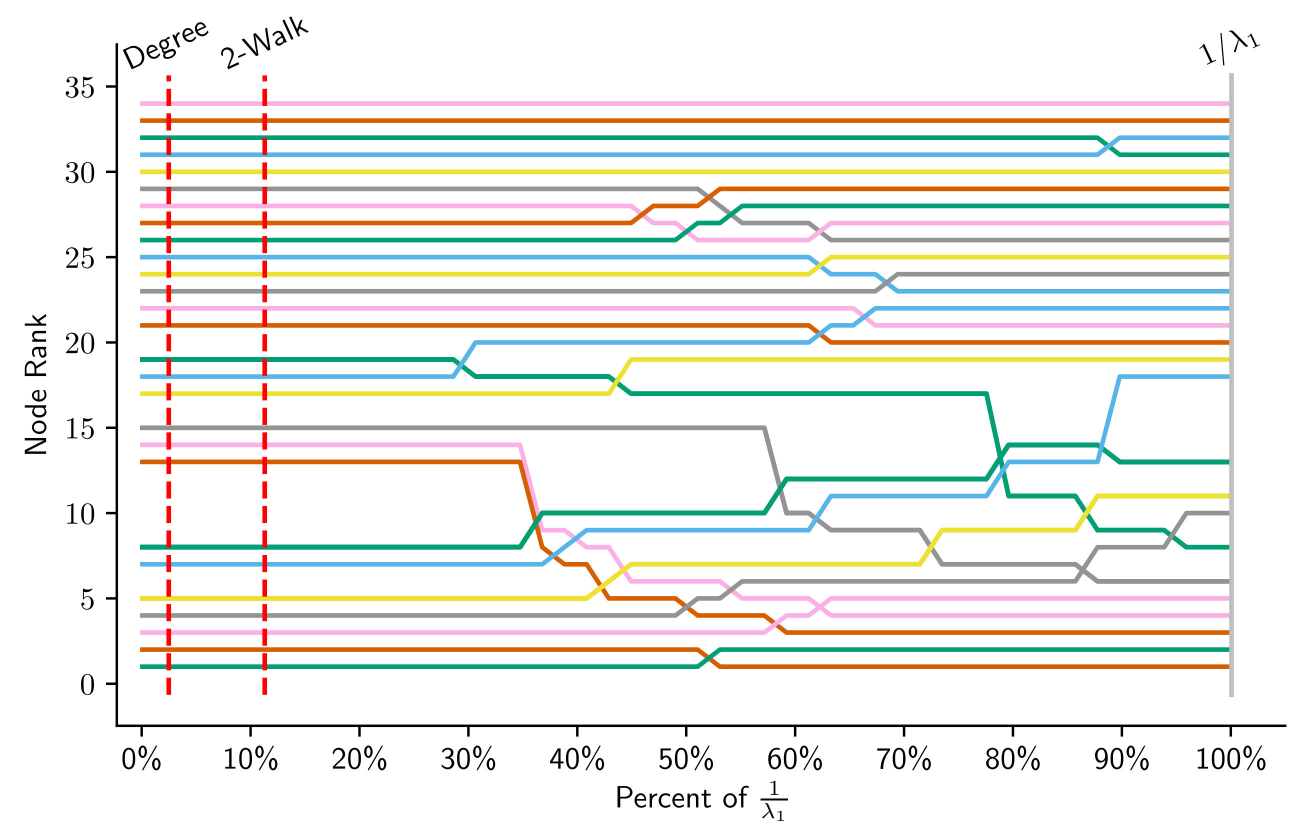
Collisions





Walk-classes merged





Walk-classes and iso-centrality

Definition:

We say nodes satisfying these properties are **isocentral**.

Theorem: [Horton, K., Sullivan] The following are equivalent:

1. Nodes *i,j* are in the same walk-class
2. Nodes *i,j* have >= *m* Katz collisions
3. Nodes *i,j* have identical **walk centrality:**

$C_f(j,\beta) = C_f(i,\beta)$ for all *f* and β

4. Nodes *i,j* have the "same" eigenvector centrality

Walk-class determines centrality behavior

Theorem. *Let* $\boldsymbol{M} \in \mathbb{R}^{n \times n}$ *be diagonalizable with* $|\lambda_1| > \cdots > |\lambda_m|$*, and eigen-decomposition* $\boldsymbol{M} = \sum_{k=1}^{m} \lambda_k \boldsymbol{V}_k \boldsymbol{U}_k^T$*. Then nodes* i, j *are in the same walk-class if and only if* $[\boldsymbol{V}_k \boldsymbol{U}_k^T \boldsymbol{s}(i)]_i = [\boldsymbol{V}_k \boldsymbol{U}_k^T \boldsymbol{s}(j)]_j$ *holds for* $k = 1, \cdots, m$*.*

Walk-class determines centrality behavior

Theorem. *Let* $\boldsymbol{M} \in \mathbb{R}^{n \times n}$ *be diagonalizable with* $|\lambda_1| > \cdots > |\lambda_m|$, *and eigen-decomposition* $\boldsymbol{M} = \sum_{k=1}^{m} \lambda_k \boldsymbol{V}_k \boldsymbol{U}_k^T$. *Then nodes* i, j *are in the same walk-class if and only if* $[\boldsymbol{V}_k \boldsymbol{U}_k^T \boldsymbol{s}(i)]_i = [\boldsymbol{V}_k \boldsymbol{U}_k^T \boldsymbol{s}(j)]_j$ *holds for* $k = 1, \cdots, m$.

Theorem:

Nodes in a walk-class have "same" **eigenvector centrality**:

i,j in the same walk-class if and only if

$[\boldsymbol{V}_k \boldsymbol{U}_k^T \mathbf{s}(i)]_i = [\boldsymbol{V}_k \boldsymbol{U}_k^T \mathbf{s}(j)]_j$ for each $k = 1, \ldots, m$

Walk-class determines centrality behavior

Depending on ***s*** and ***M***, this means *i, j* have same…

- Perron-Frobenius — eigenvector score
- Fiedler vector — score magnitude
- spectral clustering —positions symmetric in embedding

Theorem:

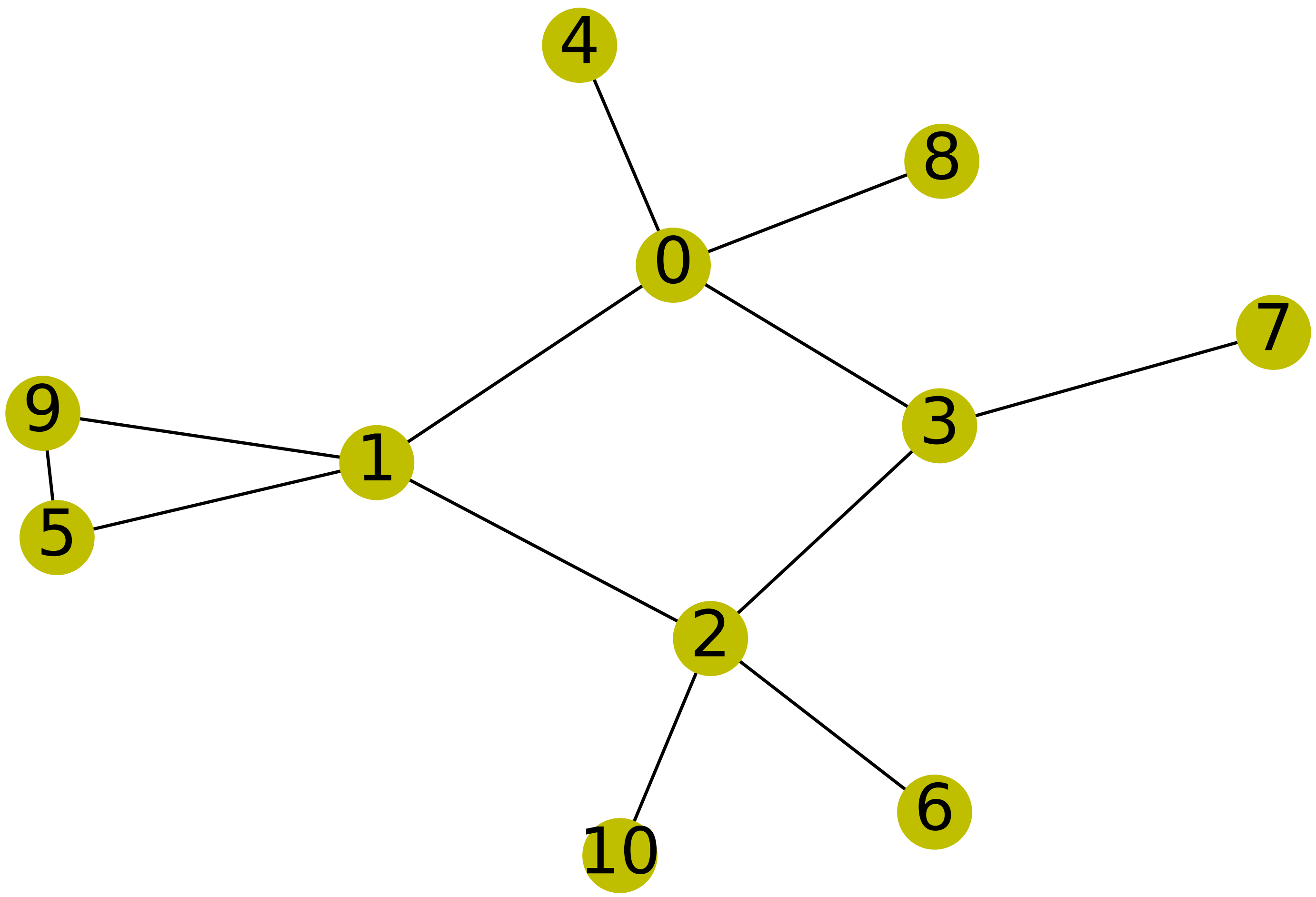
Nodes in a walk-class have "same" **eigenvector centrality**:

i,j in the same walk-class if and only if

$[\boldsymbol{V}_k \boldsymbol{U}_k^T \mathbf{s}(i)]_i = [\boldsymbol{V}_k \boldsymbol{U}_k^T \mathbf{s}(j)]_j$ for each $k = 1, \ldots, m$

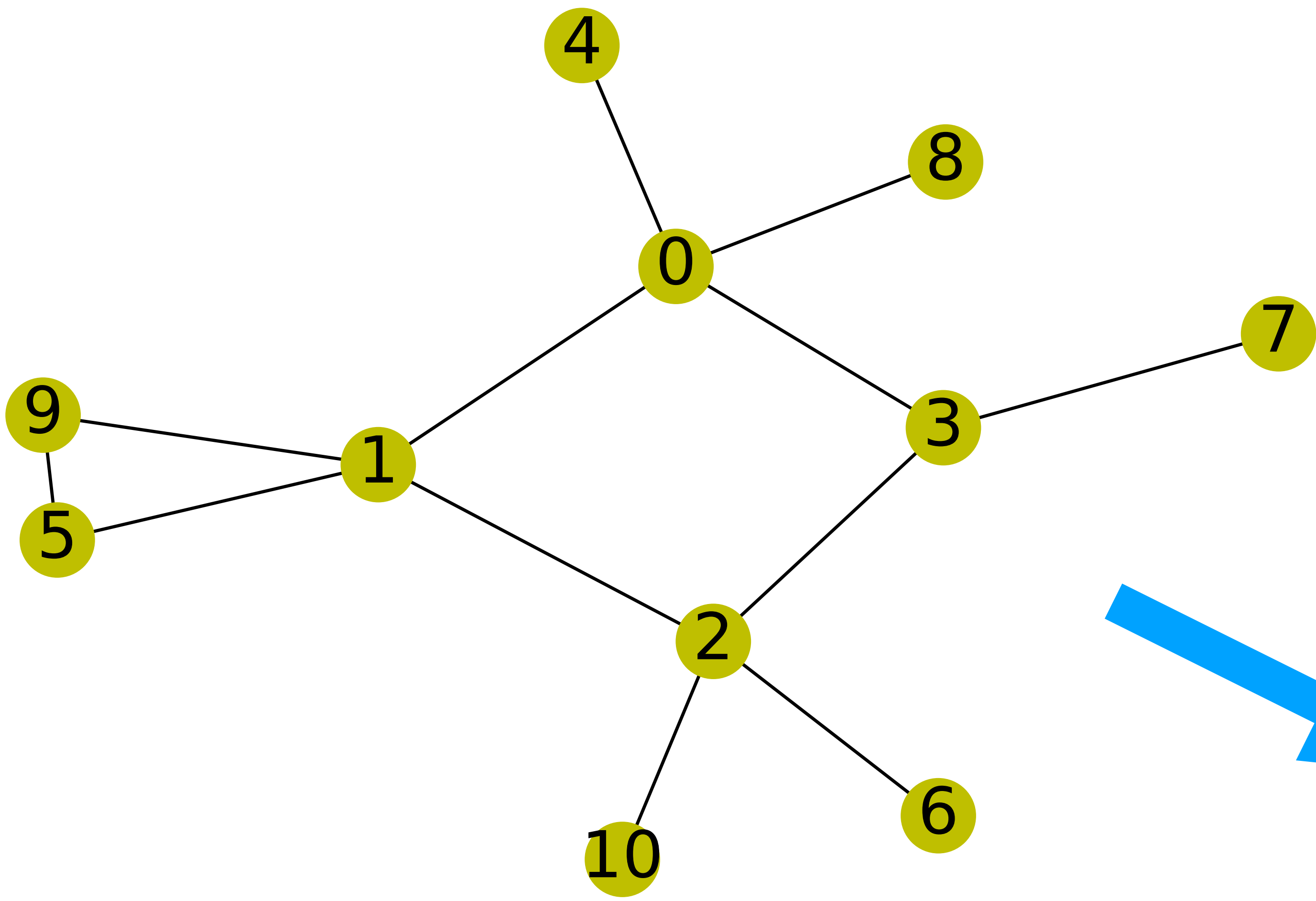
Walk-class determines centrality behavior

Example Nodes in a walk-class <-> same spectral clustering



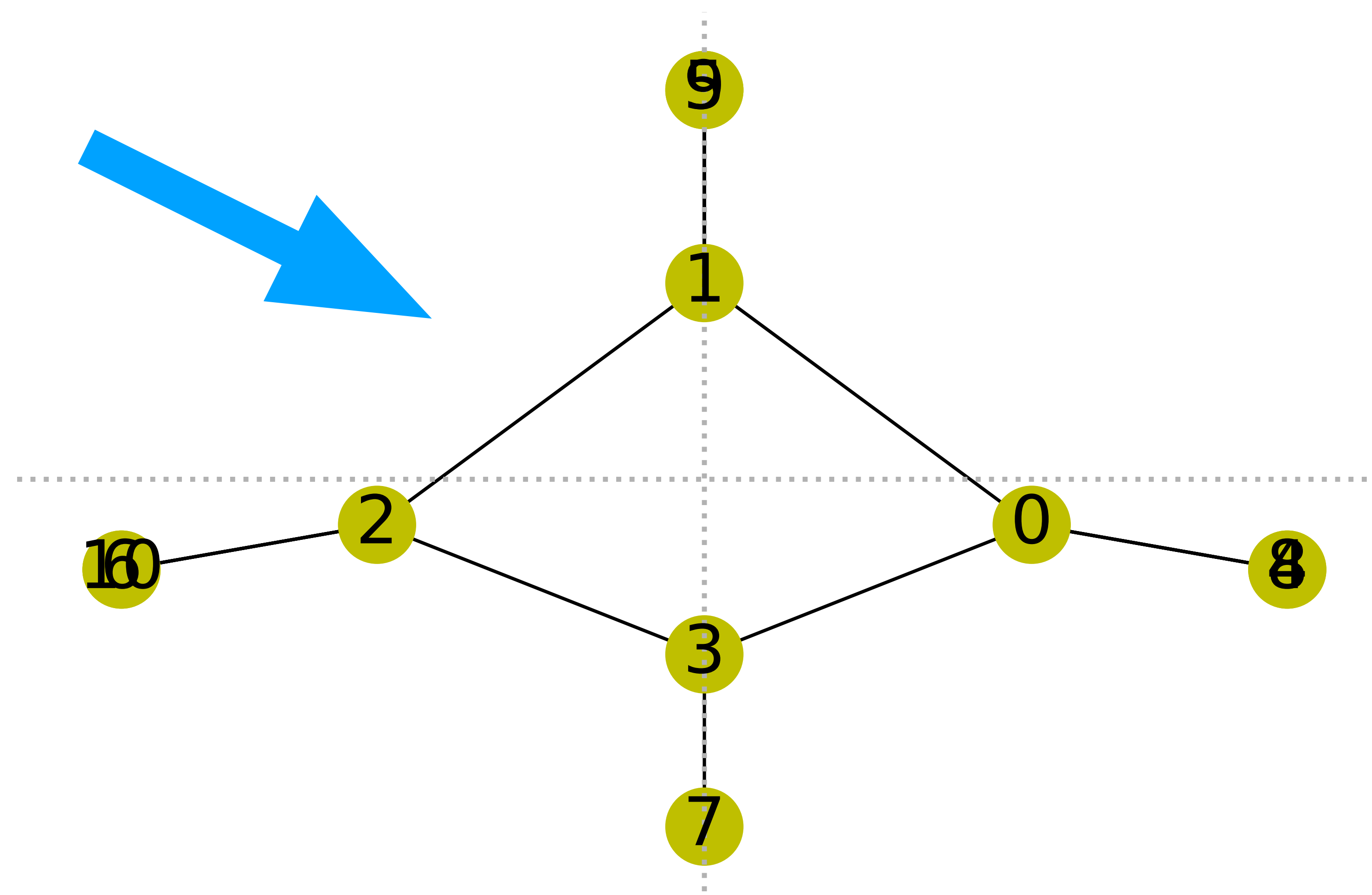
Walk-class determines centrality behavior

Example Nodes in a walk-class <-> same spectral clustering



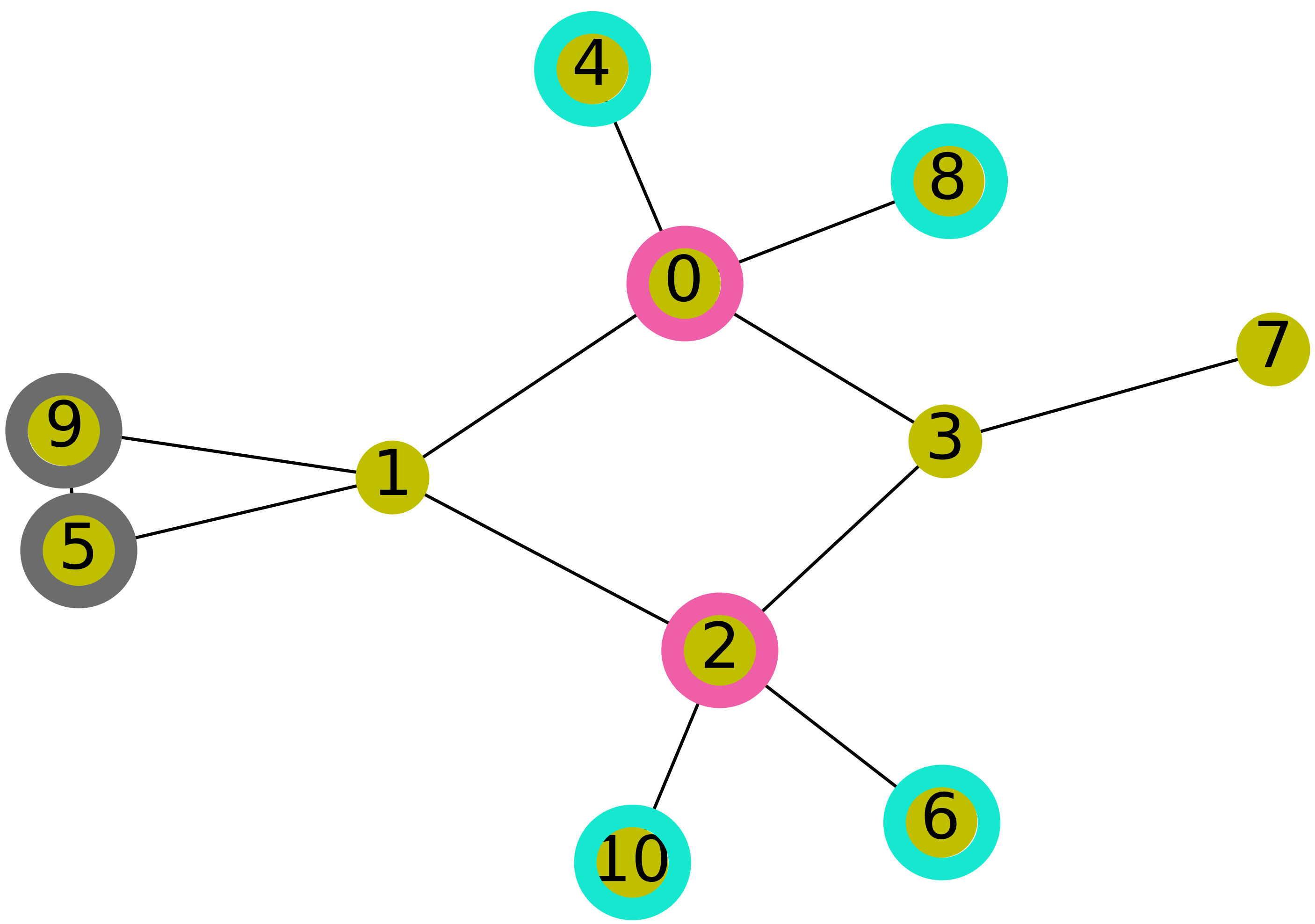
Embed nodes in $\mathbb{R}^d$

according to $\mathrm{e}_i^T \boldsymbol{U}_k$

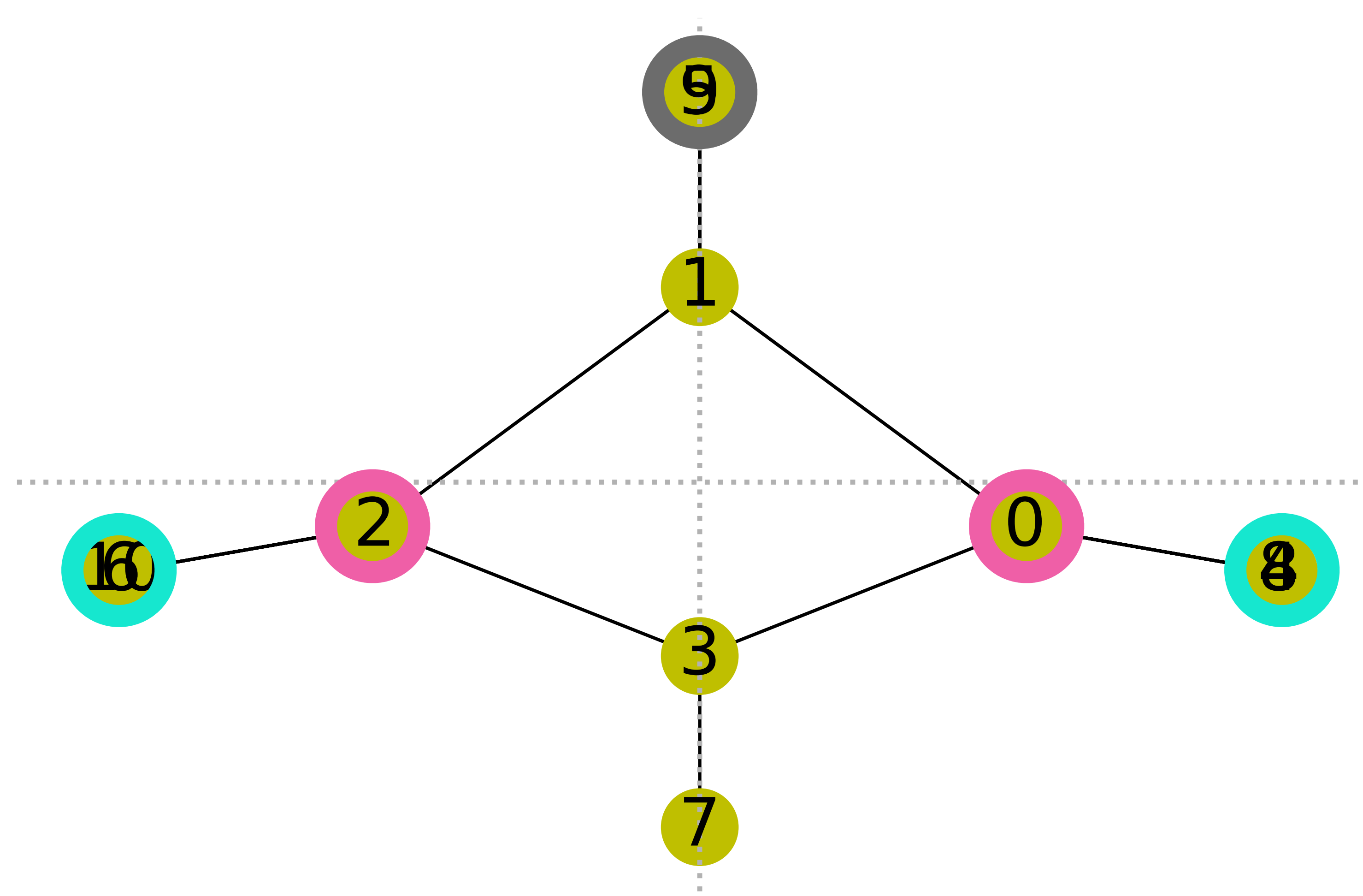


Walk-class determines centrality behavior

Example Nodes in a walk-class <-> same spectral clustering

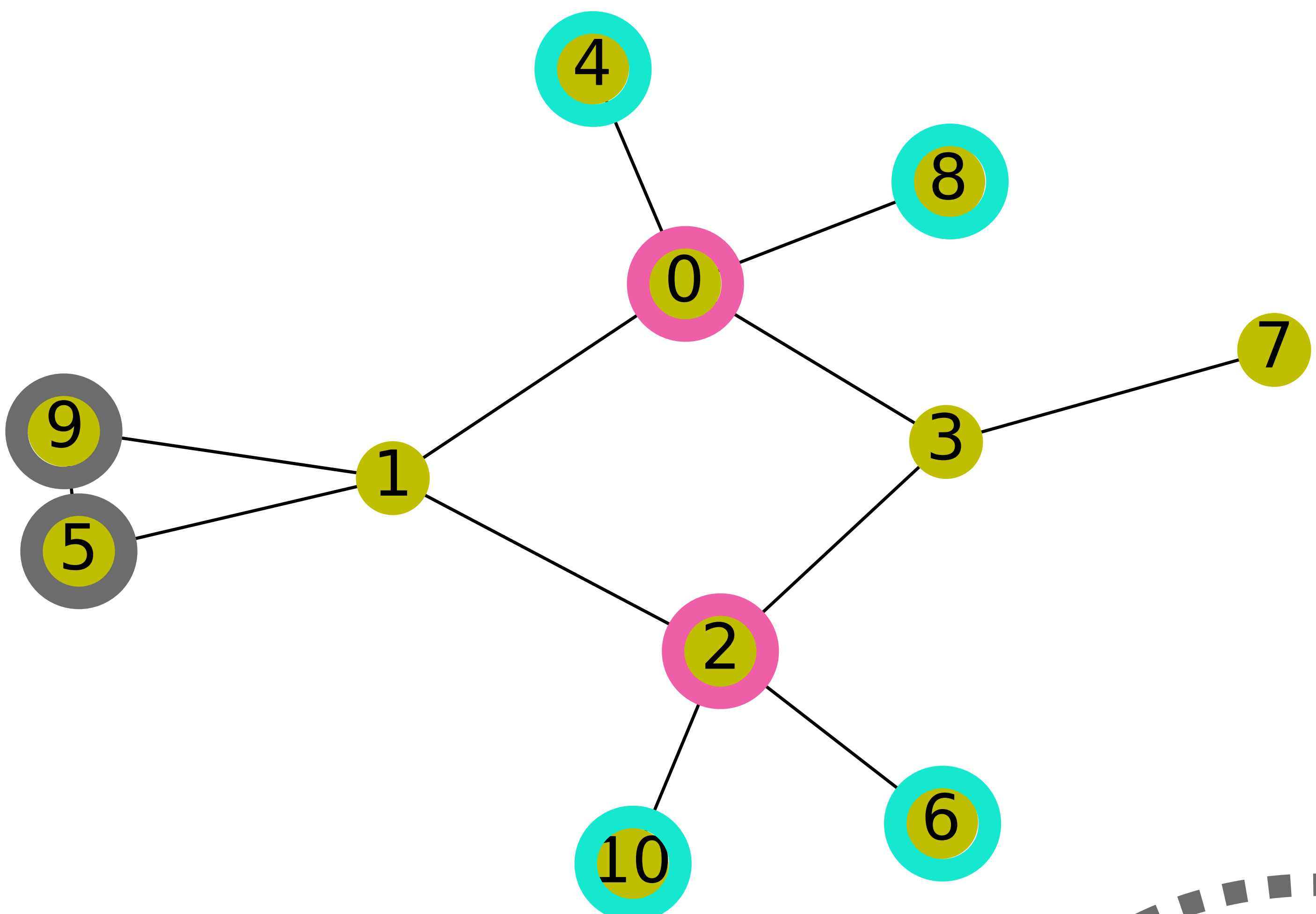


High-light walk classes

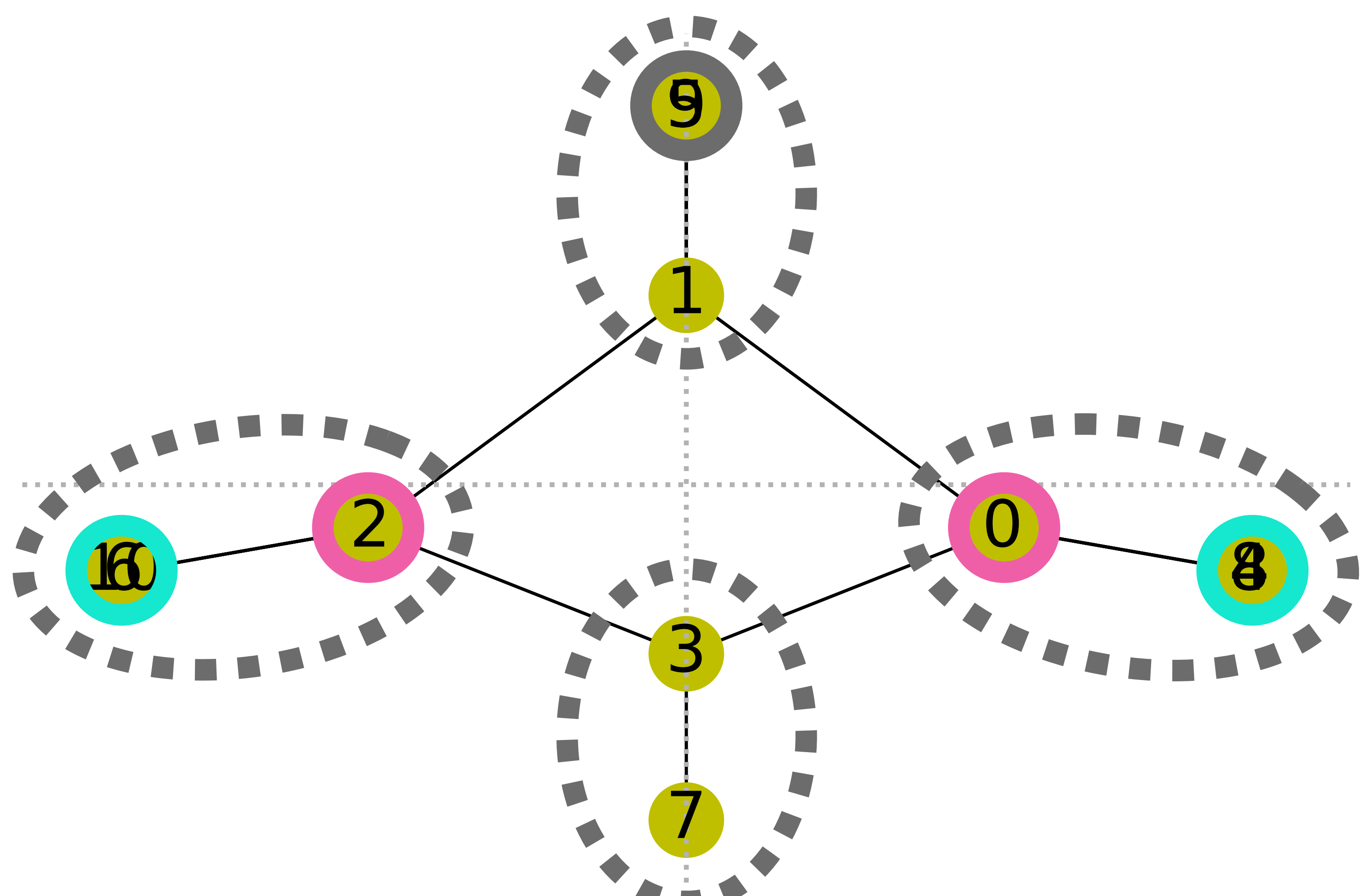


Walk-class determines centrality behavior

Example Nodes in a walk-class <-> same spectral clustering

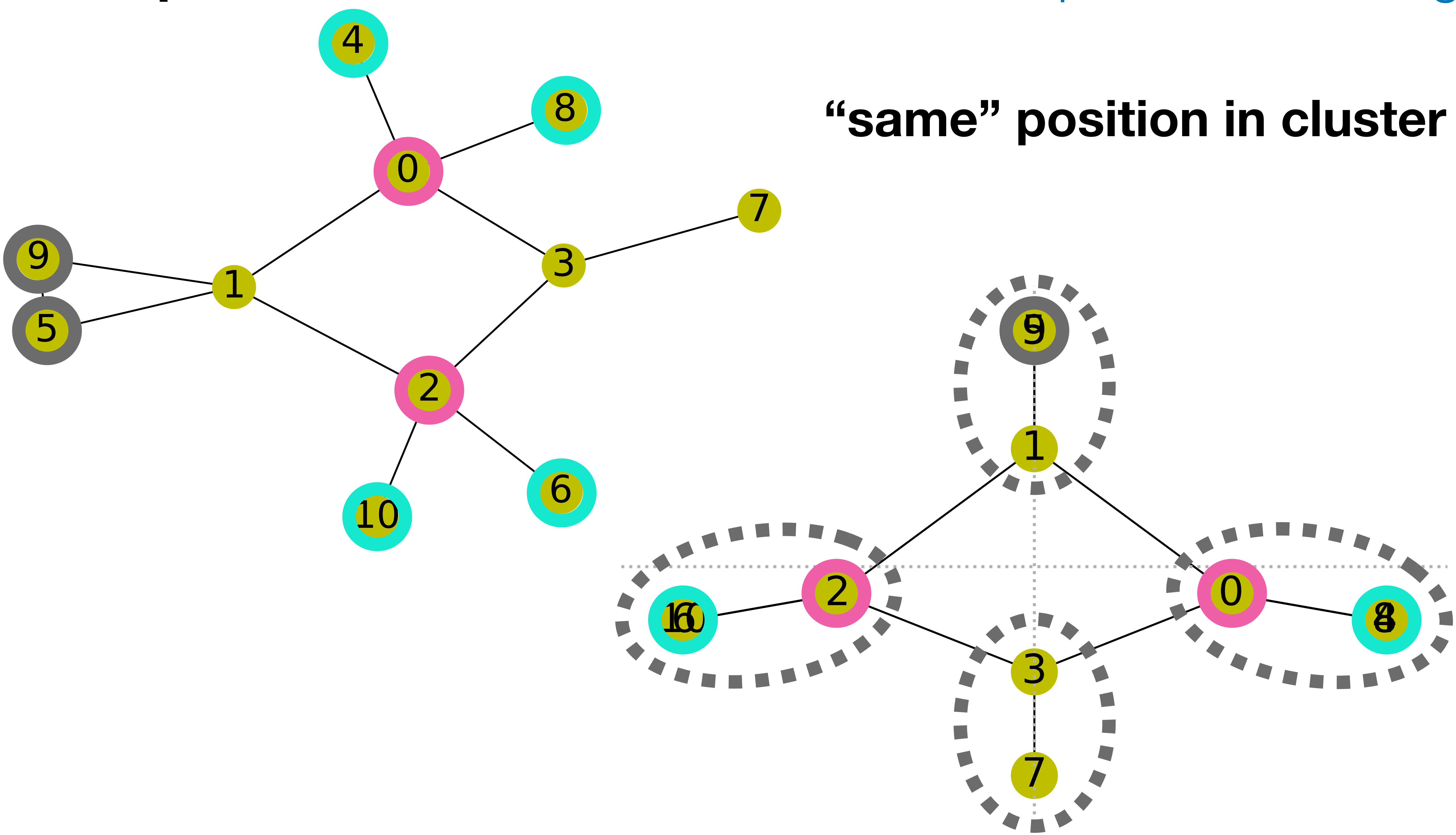


cluster by $\mathrm{e}_i^T \boldsymbol{U}_k$



Walk-class determines centrality behavior

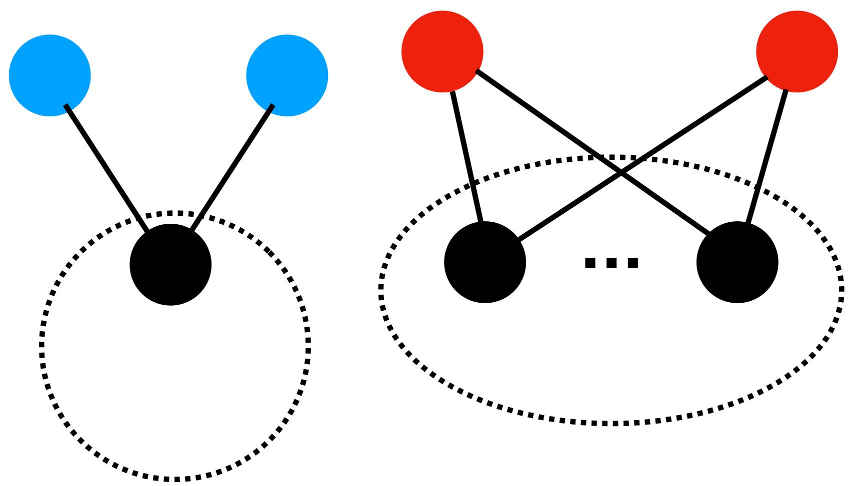
Example Nodes in a walk-class <-> same spectral clustering



Walk-class -> centrality: take-aways

- Iso-centrality can occur in real networks
- But it's rare
- Usual cause: small symmetries, e.g.

two nodes that have same neighborhood



- Iso-central nodes have identical walk-centrality, walk structure, and "same" eigen-centrality

Thank you!

Concluding thoughts:

- The more *j* (partially) majorizes *i*,

the larger interval near 0 where *j* ranks above *i*

- # collisions is finite, unless nodes are iso-central
- Iso-central nodes rare; same walk- and eigen-centrality

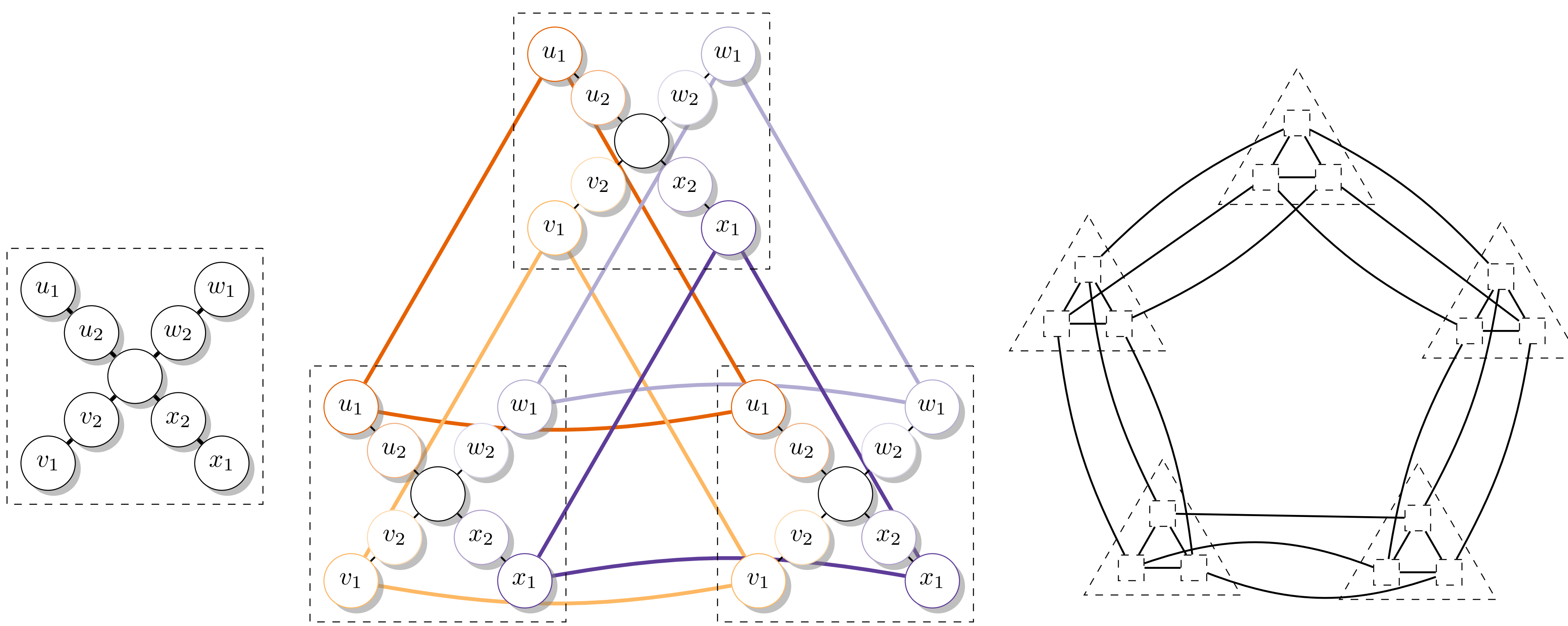
Future work: Better understand collision bounds, effect on ranking

Bonus slides, if time

Walk-classes in Spider Donuts

The spider donut graph family:

- This spider donut has exactly 3 walk-classes
- Constructed so they collide simultaneously
- Also constructed a 4-walk-class spider donut, all 4 class collide simultaneously
- Motivation: wanted to understand how many walk-classes can be involved in one collision at a single parameter value.



$S(4,2)$ $SC(4,2,[3])$ $ST(4,2,[5,3])$