Structure, Inference, and Optimization in Complex Networks

Tzu-Chi Yen, Dissertation Defense, June 29th, 2023

Graph 1O1

Input, raw

Output, processed Data

DATA | QUESTION | FRAMEWORK | REPRESENTATION

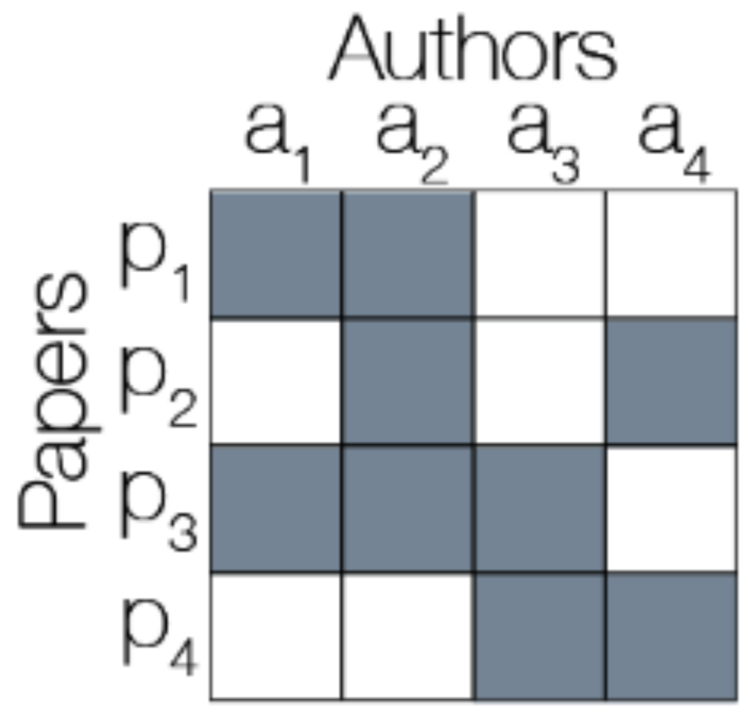
Paper : Authors

$p_1 : a_1, a_2$

$p_2 : a_2, a_4$

$p_3 : a_1, a_2, a_3$

$p_4 : a_3, a_4$



Graph 1O1

Input, raw

$V = \{a_1, a_2, \ldots, a_4\}$

Output, processed Data

DATA

QUESTION

FRAMEWORK

REPRESENTATION

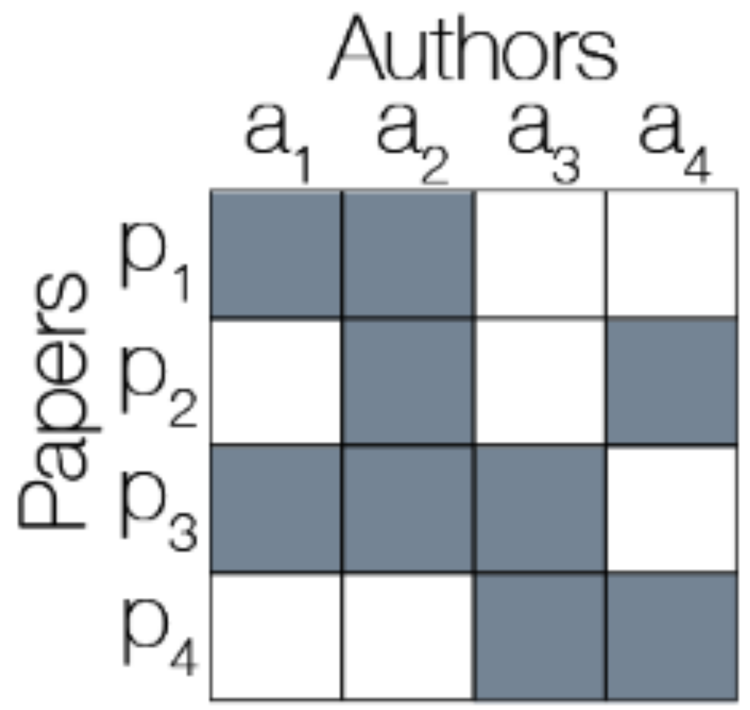
Paper : Authors

$p_1 : a_1, a_2$

$p_2 : a_2, a_4$

$p_3 : a_1, a_2, a_3$

$p_4 : a_3, a_4$



Graph 1O1

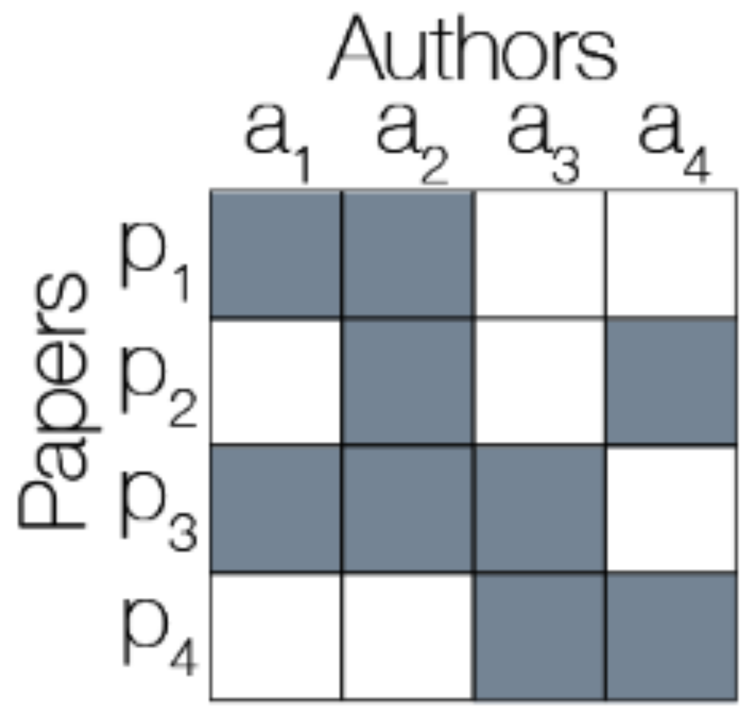
Input, raw

$V = \{a_1, a_2, \ldots, a_4\}$

Output, processed Data

DATA

Paper : Authors
$p_1 : a_1, a_2$
$p_2 : a_2, a_4$
$p_3 : a_1, a_2, a_3$
$p_4 : a_3, a_4$



QUESTION

Has this pair of authors written together? →

FRAMEWORK →

REPRESENTATION

Graph 1O1

Input, raw

$V = \{a_1, a_2, \dots, a_4\}$

Output, processed Data

DATA | QUESTION | FRAMEWORK | REPRESENTATION

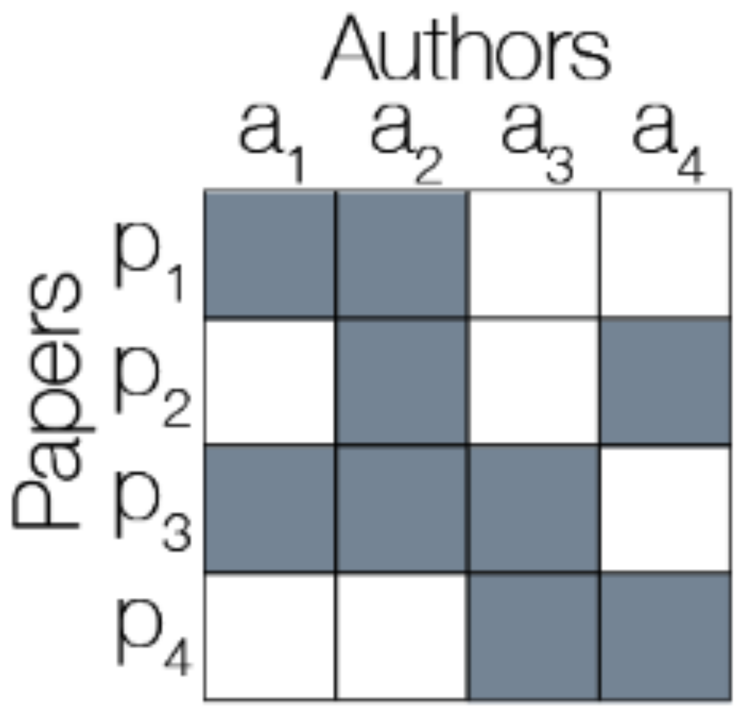
Paper : Authors

$p_1 : a_1, a_2$

$p_2 : a_2, a_4$

$p_3 : a_1, a_2, a_3$

$p_4 : a_3, a_4$



Has this pair of authors written together?

Pick any 2 nodes as an edge

Graph 1O1

Input, raw

$V = \{a_1, a_2, \ldots, a_4\}$

Output, processed Data

DATA

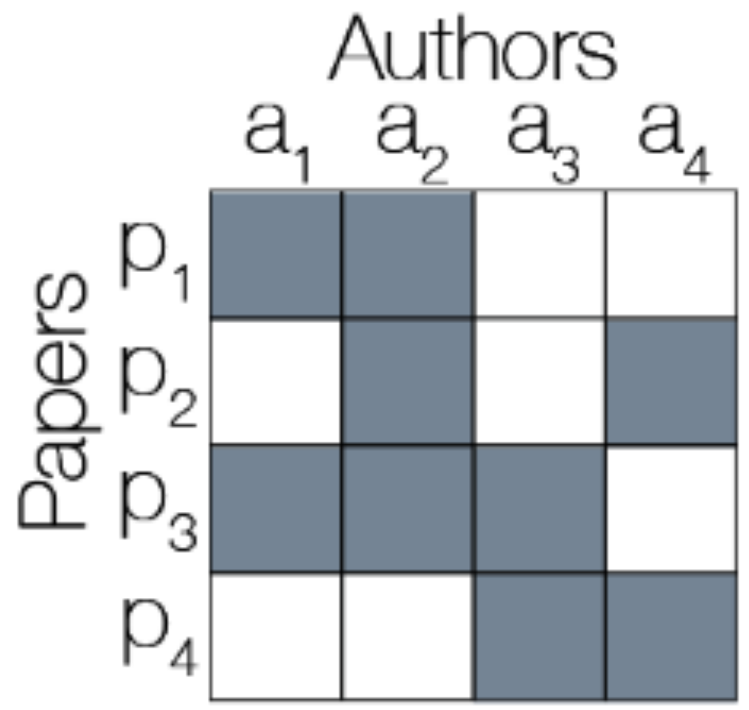
Paper : Authors

$p_1 : a_1, a_2$

$p_2 : a_2, a_4$

$p_3 : a_1, a_2, a_3$

$p_4 : a_3, a_4$



QUESTION

Has this pair of authors written together?

Pick any 2 nodes as an edge

FRAMEWORK

Graph

REPRESENTATION

Graph 1O1

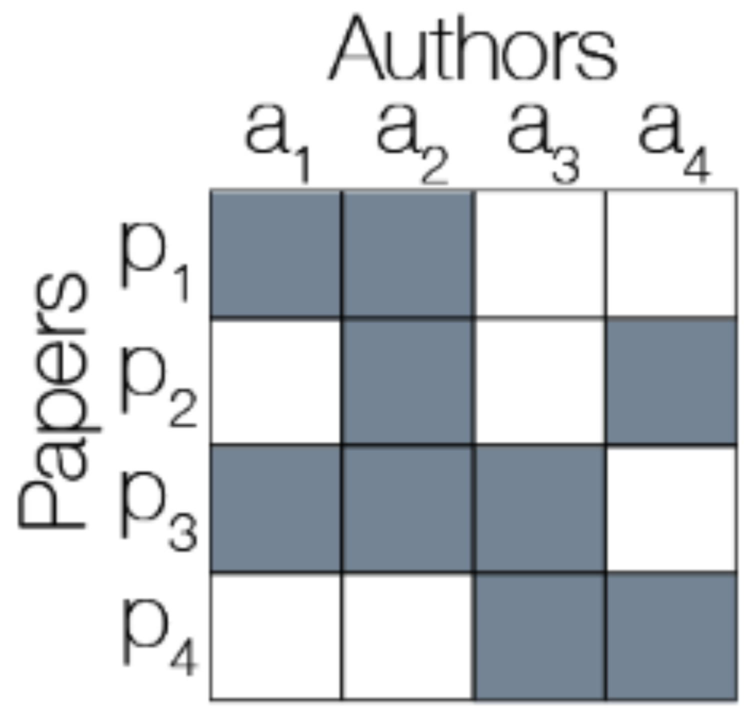
Input, raw

$V = \{a_1, a_2, \ldots, a_4\}$

Output, processed Data

DATA

Paper : Authors
$p_1 : a_1, a_2$
$p_2 : a_2, a_4$
$p_3 : a_1, a_2, a_3$
$p_4 : a_3, a_4$



QUESTION

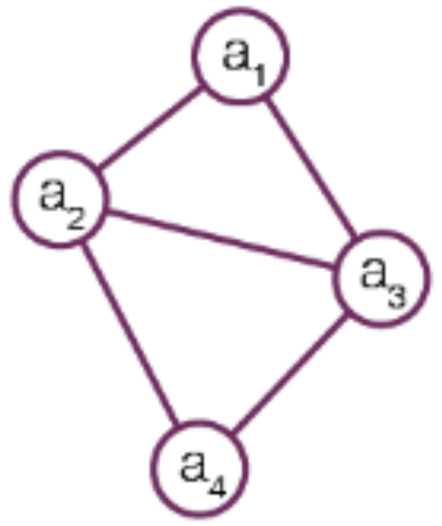
Has this pair of authors written together?

Pick any 2 nodes as an edge

FRAMEWORK

Graph

REPRESENTATION



Graph 1O1

Input, raw

$V = \{a_1, a_2, \ldots, a_4\}$

Output, processed Data

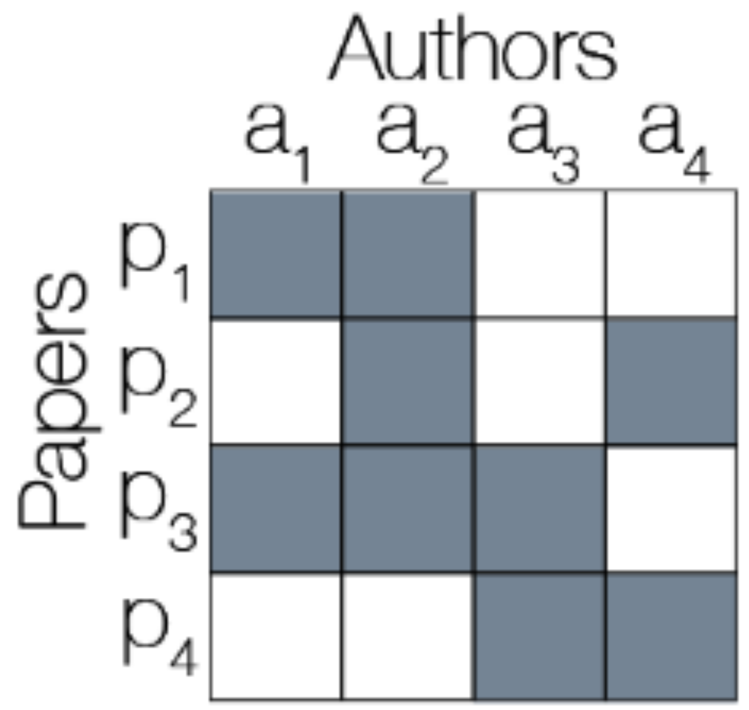
DATA

QUESTION

FRAMEWORK

REPRESENTATION

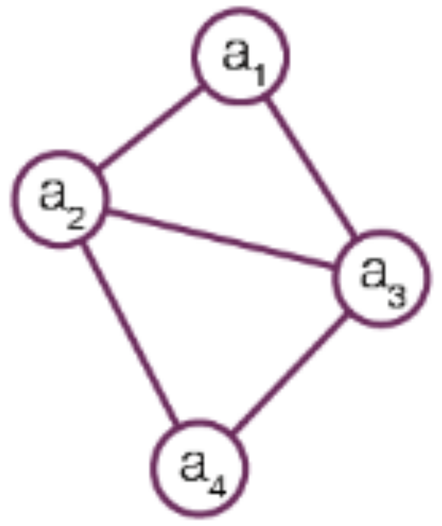
Paper : Authors
$p_1 : a_1, a_2$
$p_2 : a_2, a_4$
$p_3 : a_1, a_2, a_3$
$p_4 : a_3, a_4$



Has this pair of authors written together?

Pick any 2 nodes as an edge

→ *Graph* →



$$G = (V, \{e\})$$

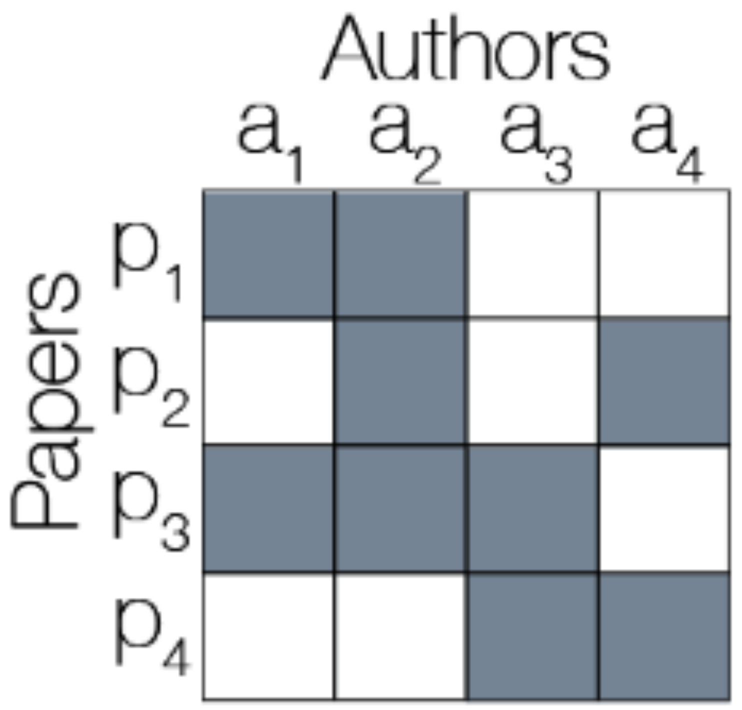
$$\{e\} = \{(1,\ 2),\ (2,\ 4),\ (1,\ 2),\ (1,\ 3),\ (2,\ 3),\ (3,\ 4)\}$$

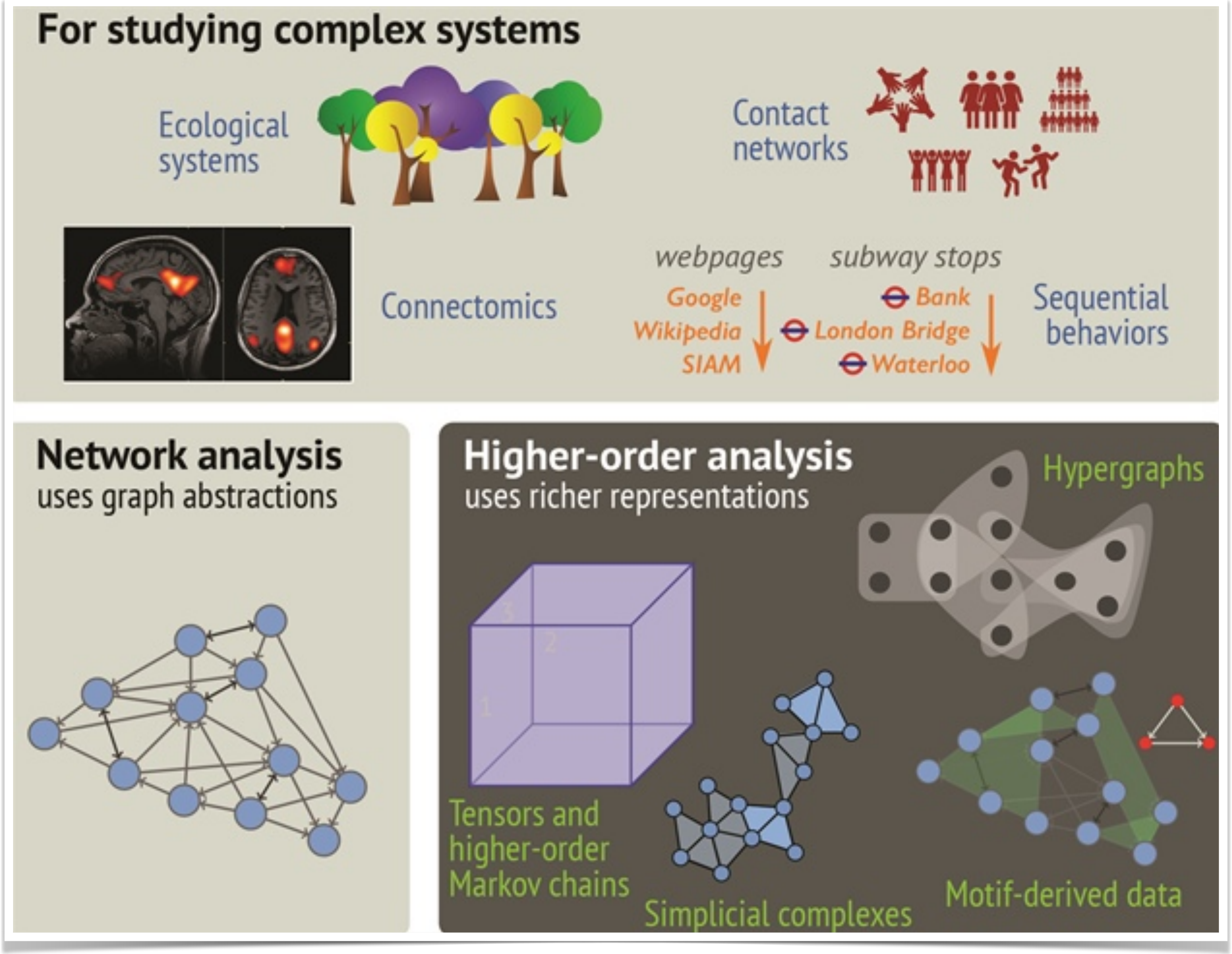
Graph 1O1

Input, raw

DATA

Paper : Authors
$p_1 : a_1, a_2$
$p_2 : a_2, a_4$
$p_3 : a_1, a_2, a_3$
$p_4 : a_3, a_4$



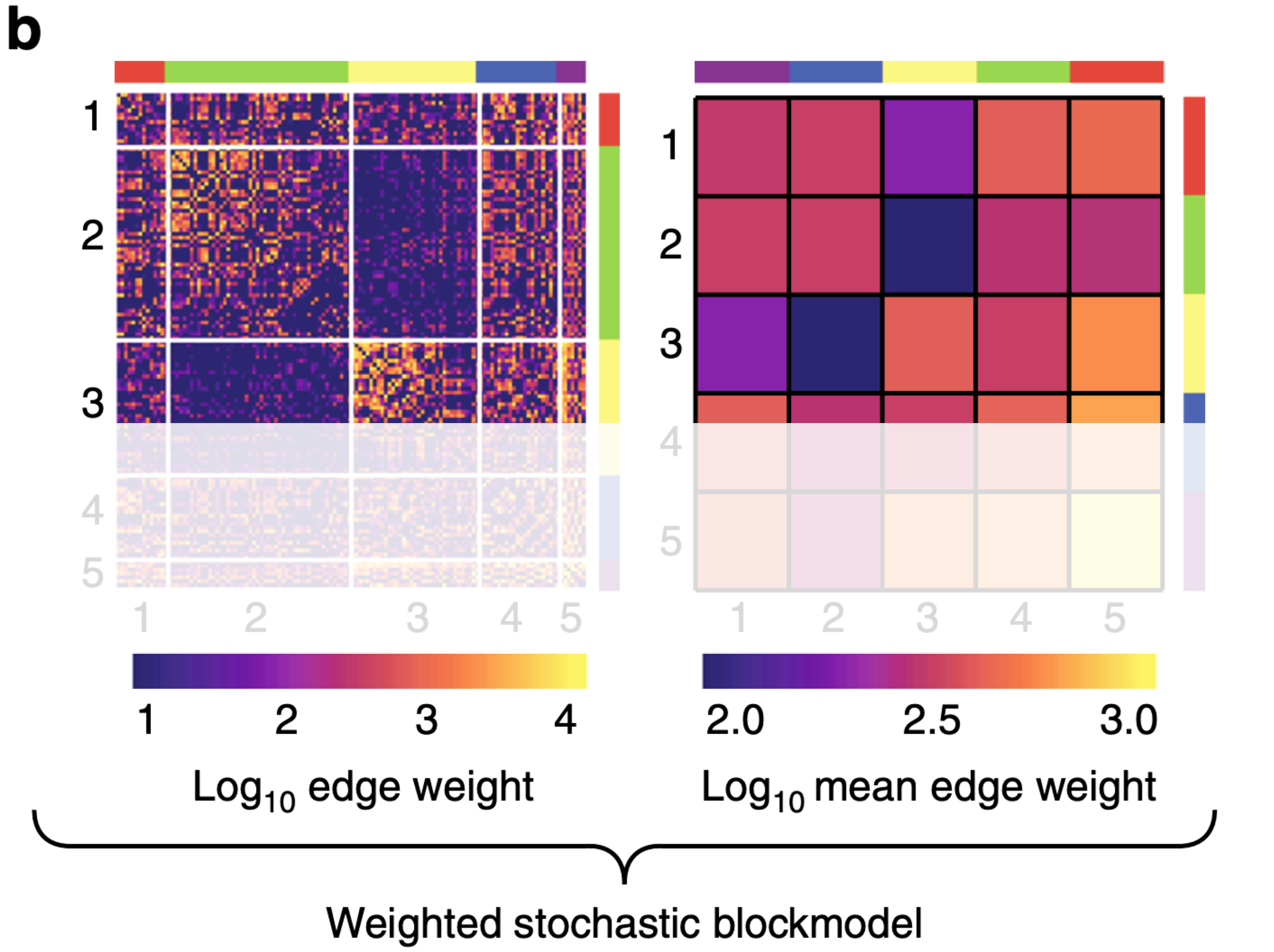


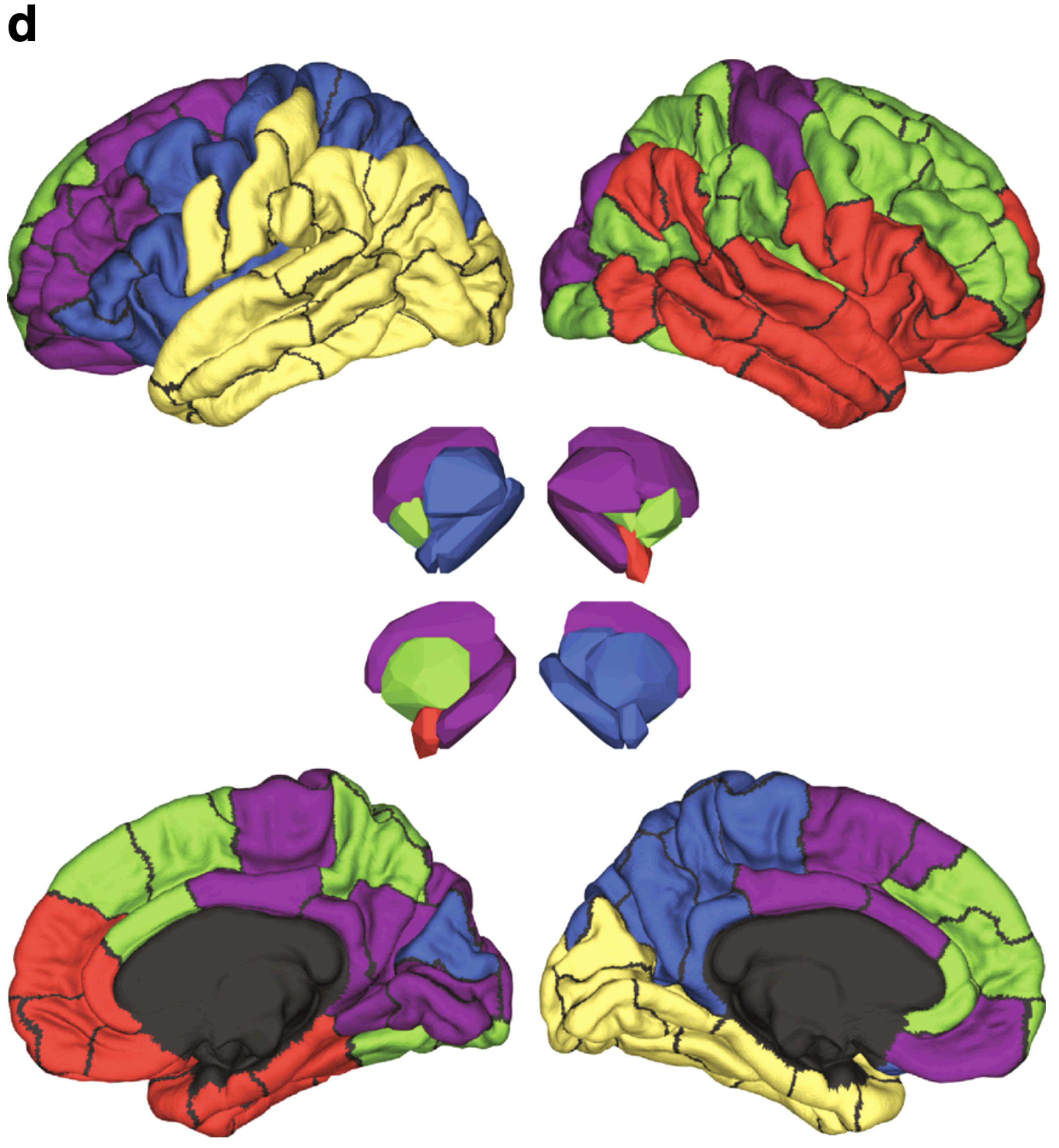
Output, processed Data

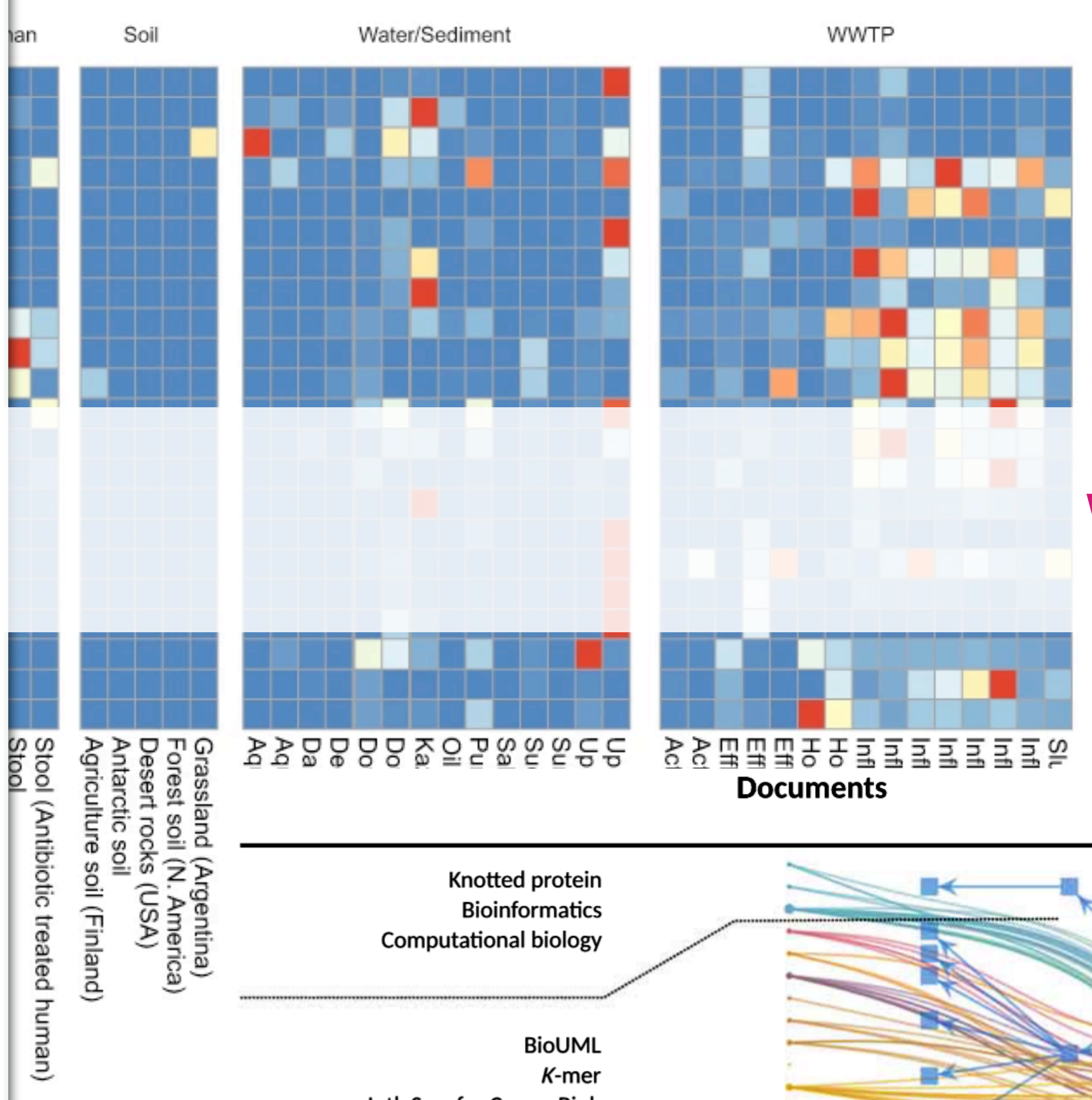
$$G = (V, \{e\})$$

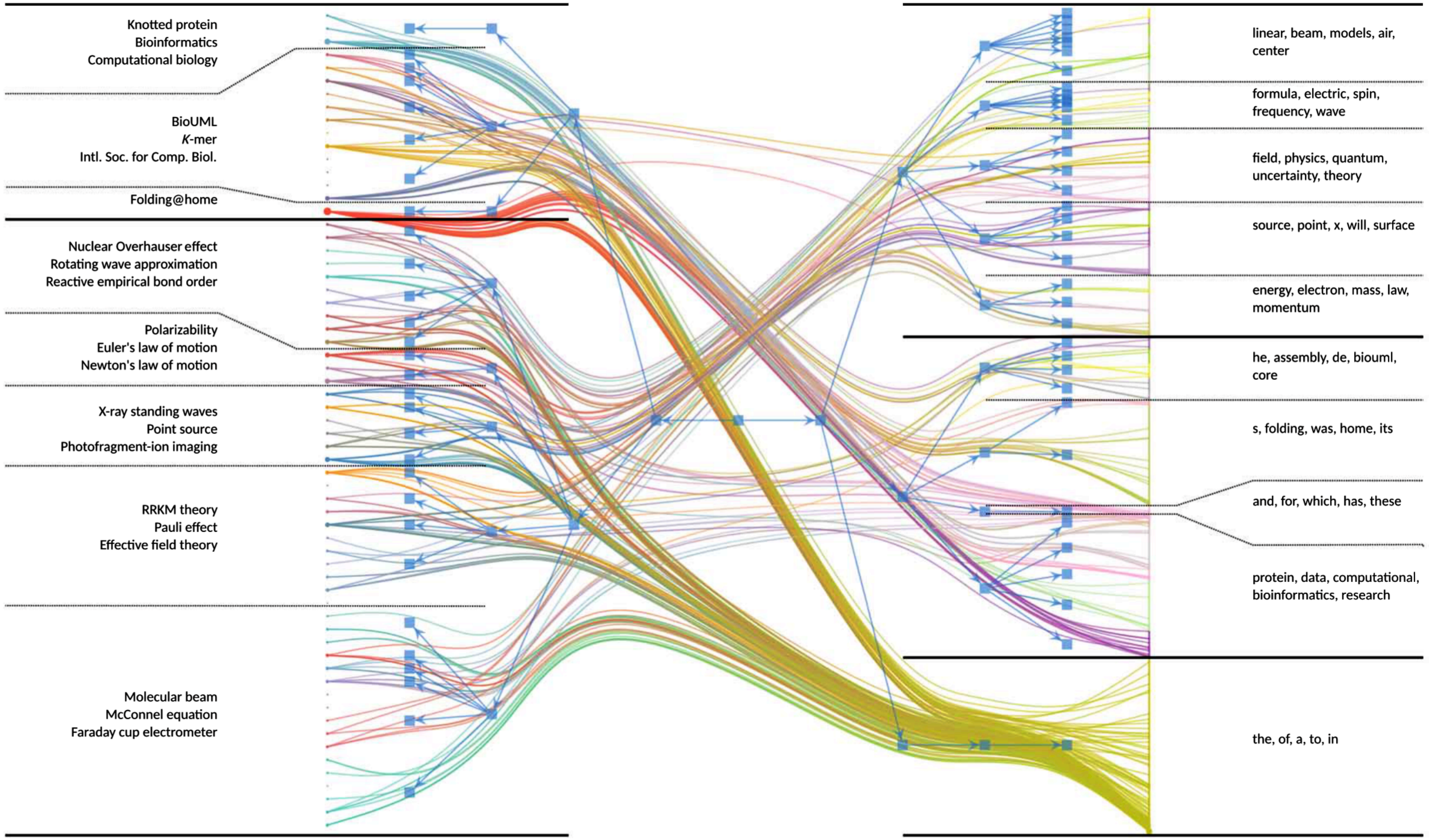
$$\{e\} = \{(1, 2), (2, 4), (1, 2), (1, 3), (2, 3), (3, 4)\}$$

What structure can we infer from a network?

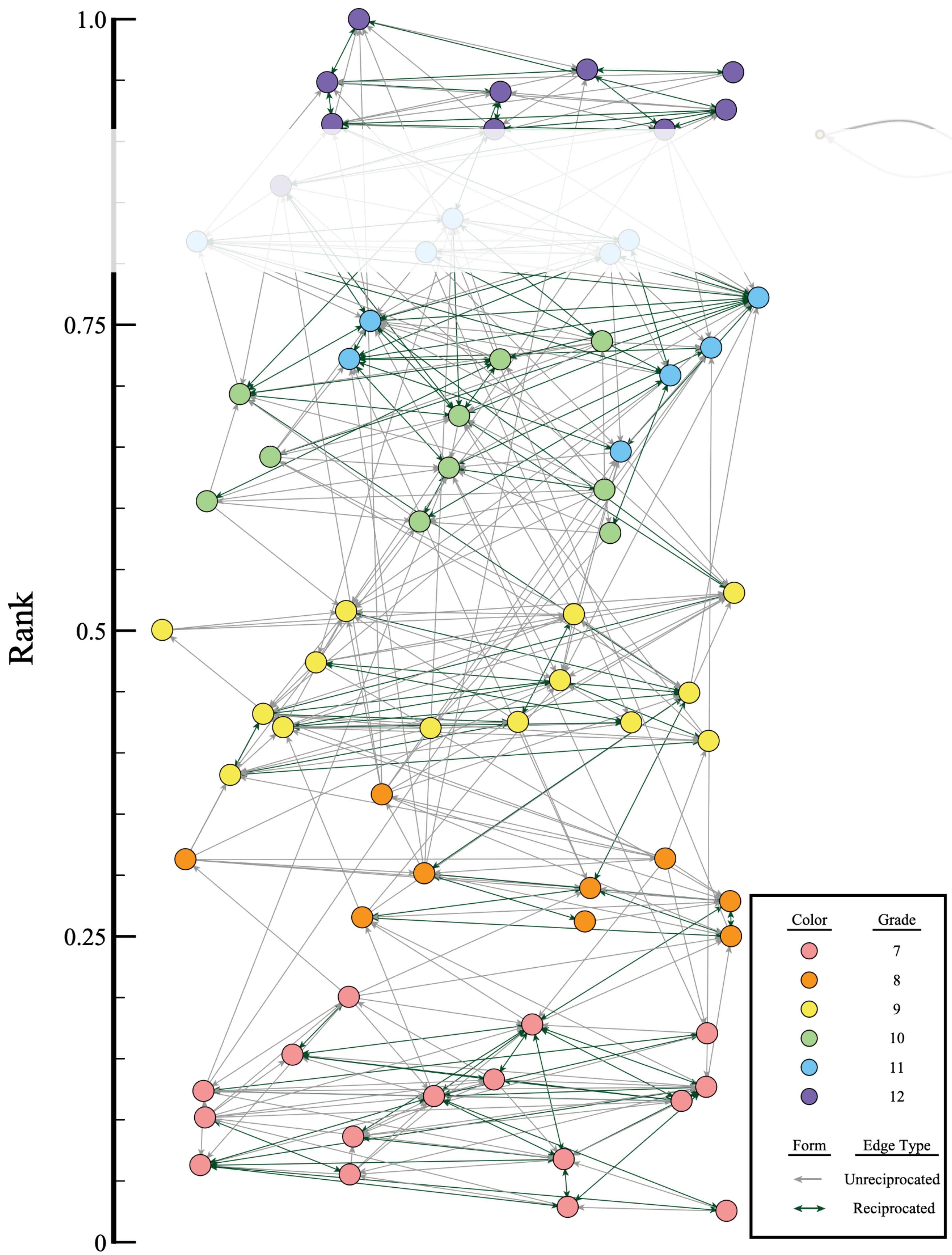


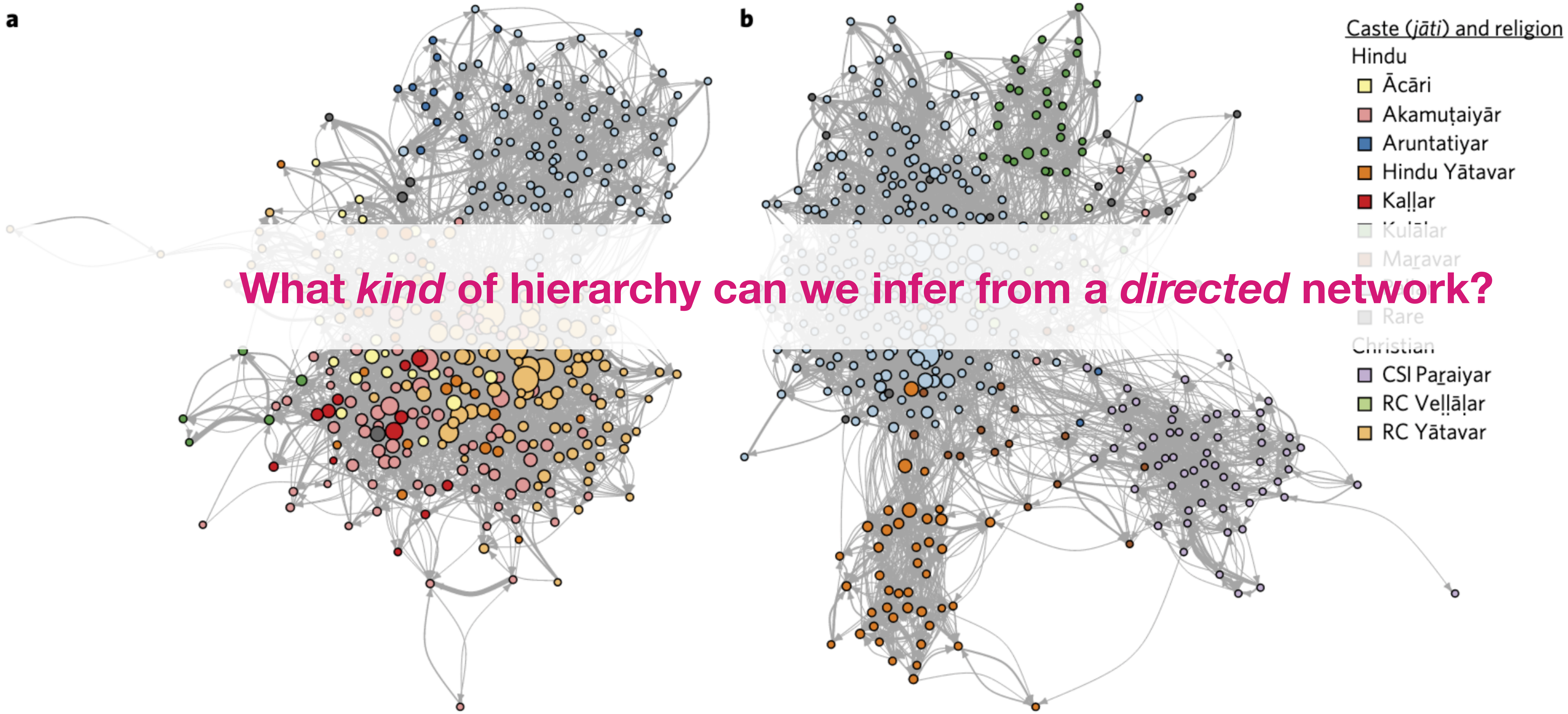


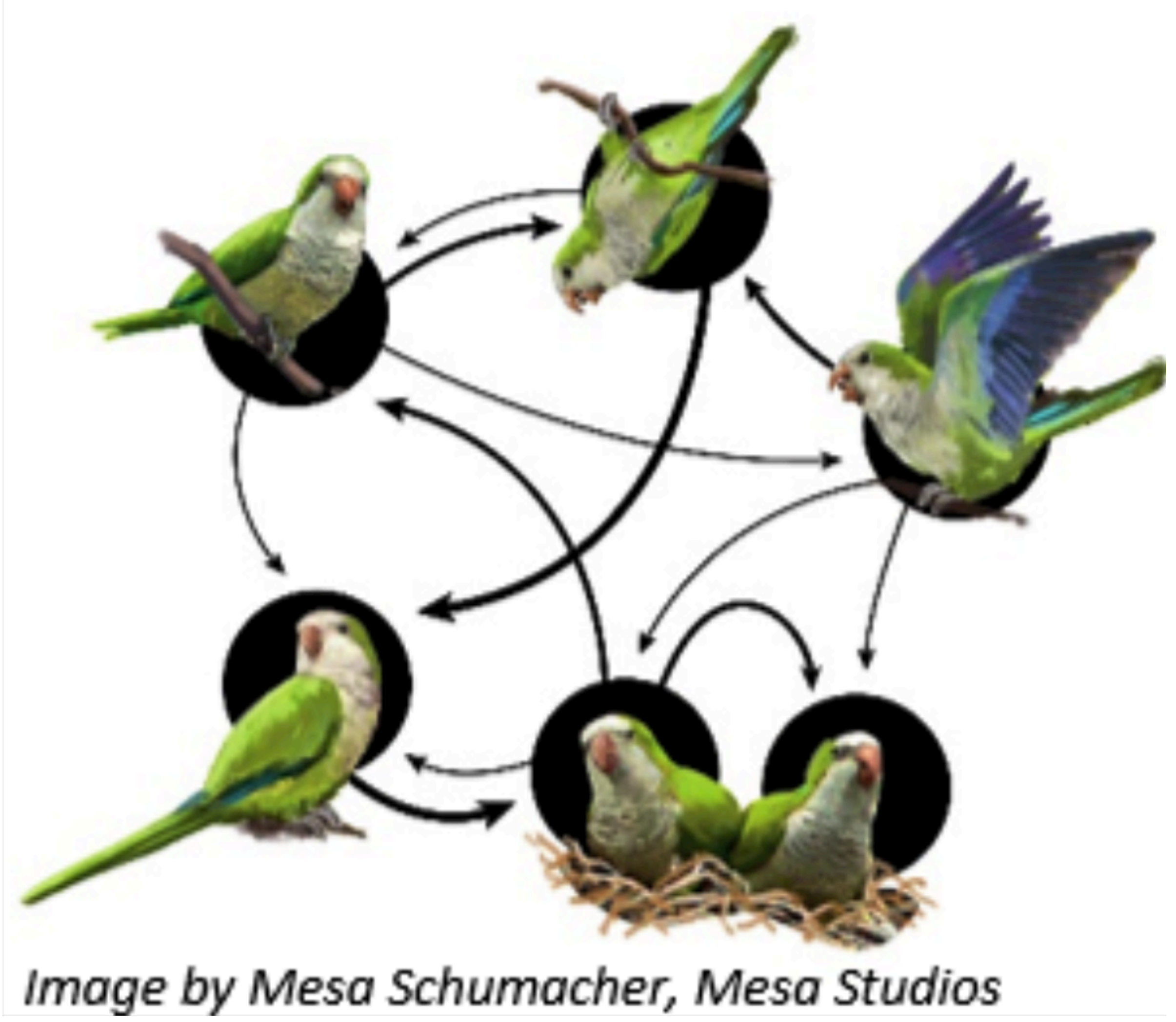


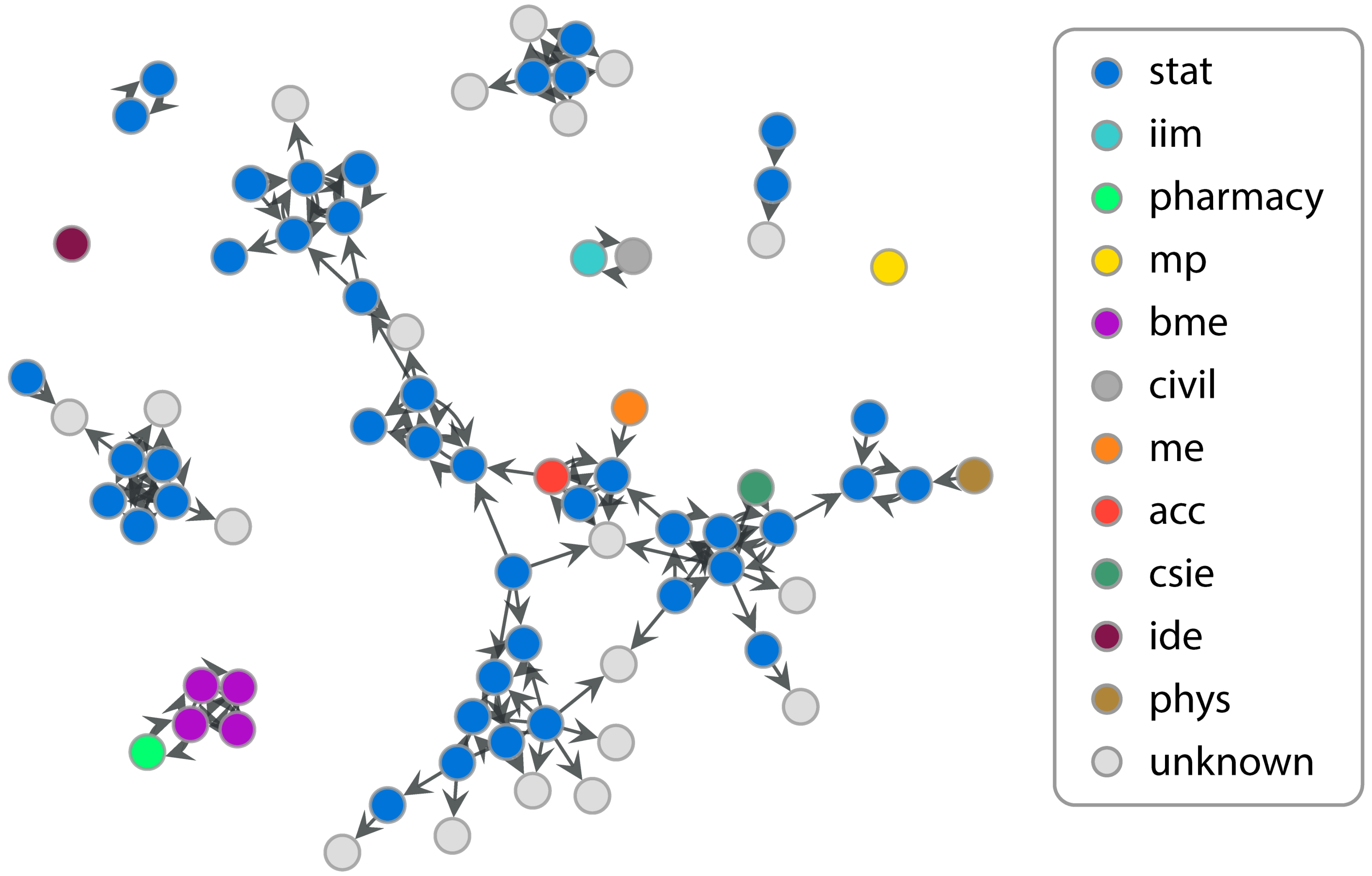


What *kind* of hierarchy can we infer from a *directed* network?









Outline

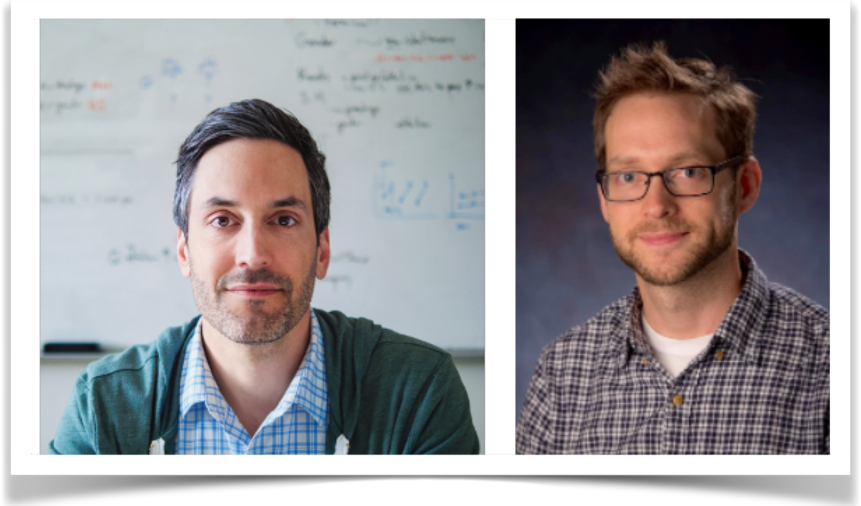
1. Context & motivation (chapter 1)
2. Aspiration of prestige in institutional peer selection (chapter 5)
3. Community detection in bipartite networks (chapter 2)
4. Regularized methods for efficient ranking (chapter 3)
5. A detour in higher-order structures (chapter 4)
6. Conclusions & future research (chapter 6)

Outline

1. Context & motivation (chapter 1)
2. Aspiration of prestige in institutional peer selection (chapter 5)
3. Community detection in bipartite networks (chapter 2)
4. Regularized methods for efficient ranking (chapter 3)
5. A detour in higher-order structures (chapter 4)
6. Conclusions & future research (chapter 6)

Outline

1. Context & motivation (chapter 1)
2. Aspiration of prestige in institutional peer selection (chapter 5)
3. Community detection in bipartite netv
4. Regularized methods for efficient rank
5. A detour in higher-order structures (cha
6. Conclusions & future research (chapter 6



Motivation (Structure vs. Psychology)

Social hierarchy. directional bias in interactions, observed interactions tend to point from higher-ranked individuals to lower-ranked ones

Motivation (Structure vs. Psychology)

Social hierarchy. directional bias in interactions, observed interactions tend to point from higher-ranked individuals to lower-ranked ones

U.S. academic system. strong self-reinforcing hierarchy of status and signaling.

Motivation (Structure vs. Psychology)

Social hierarchy. directional bias in interactions, observed interactions tend to point from higher-ranked individuals to lower-ranked ones

U.S. academic system. strong self-reinforcing hierarchy of status and signaling.

Prestige drives epistemic inequality in the diffusion of scientific ideas

Allison C. Morgan, Dimitrios J. Economou, Samuel F. Way & Aaron Clauset

EPJ Data Science **7**, Article number: 40 (2018) | Cite this article

Motivation (Structure vs. Psychology)

Social hierarchy. directional bias in interactions, observed interactions tend to point from higher-ranked individuals to lower-ranked ones

U.S. academic system. strong self-reinforcing hierarchy of status and signaling.

Labor advantages drive the greater productivity of faculty at elite universities

SAM ZHANG, K. HUNTER WAPMAN, DANIEL B. LARREMORE, AND AARON CLAUSET Authors Info & Affiliations

SCIENCE ADVANCES • 18 Nov 2022 • Vol 8, Issue 46 • DOI: 10.1126/sciadv.abq7056

Prestige drives epistemic inequality in the diffusion of scientific ideas

Allison C. Morgan, Dimitrios J. Economou, Samuel F. Way & Aaron Clauset

EPJ Data Science **7**, Article number: 40 (2018) | Cite this article

Motivation (Structure vs. Psychology)

Social hierarchy. directional bias in interactions, observed interactions tend to point from higher-ranked individuals to lower-ranked ones

U.S. academic system. strong self-reinforcing hierarchy of status and signaling.

Labor advantages drive the greater productivity of faculty at elite universities

SAM ZHANG, K. HUNTER WAPMAN, DANIEL B. LARREMORE, AND AARON CLAUSET Authors Info & Affiliations

SCIENCE ADVANCES • 18 Nov 2022 • Vol 8, Issue 46 • DOI: 10.1126/sciadv.abq7056

Prestige drives epistemic inequality in the diffusion of scientific ideas

Allison C. Morgan, Dimitrios J. Economou, Samuel F. Way & Aaron Clauset

EPJ Data Science **7**, Article number: 40 (2018) | Cite this article

Quantifying hierarchy and dynamics in US faculty hiring and retention

K. Hunter Wapman, Sam Zhang, Aaron Clauset & Daniel B. Larremore

Nature **610**, 120–127 (2022) | Cite this article

Motivation (Structure vs. Psychology)

Social hierarchy. directional bias in interactions, observed interactions tend to point from higher-ranked individuals to lower-ranked ones

U.S. academic system. strong self-reinforcing hierarchy of status and signaling.

Aspiration behavior. individuals routinely aim "higher" and judge themselves to be peers of those of higher status.

Labor advantages drive the greater productivity of faculty at elite universities

SAM ZHANG, K. HUNTER WAPMAN, DANIEL B. LARREMORE, AND AARON CLAUSET Authors Info & Affiliations

SCIENCE ADVANCES • 18 Nov 2022 • Vol 8, Issue 46 • DOI: 10.1126/sciadv.abq7056

Prestige drives epistemic inequality in the diffusion of scientific ideas

Allison C. Morgan, Dimitrios J. Economou, Samuel F. Way & Aaron Clauset

EPJ Data Science **7**, Article number: 40 (2018) | Cite this article

Quantifying hierarchy and dynamics in US faculty hiring and retention

K. Hunter Wapman, Sam Zhang, Aaron Clauset & Daniel B. Larremore

Nature **610**, 120–127 (2022) | Cite this article

Motivation (Structure vs. Psychology)

Social hierarchy. directional bias in interactions, observed interactions tend to point from higher-ranked individuals to lower-ranked ones

U.S. academic system. strong self-reinforcing hierarchy of status and signaling.

Aspiration behavior. individuals routinely aim “higher” and judge themselves to be peers of those of higher status.

Labor advantages drive the greater productivity of faculty at elite universities

SAM ZHANG, K. HUNTER WAPMAN, DANIEL B. LARREMORE, AND AARON CLAUSET Authors Info & Affiliations

SCIENCE ADVANCES • 18 Nov 2022 • Vol 8, Issue 46 • DOI: 10.1126/sciadv.abq7056

Prestige drives epistemic inequality in the diffusion of scientific ideas

Allison C. Morgan, Dimitrios J. Economou, Samuel F. Way & Aaron Clauset

EPJ Data Science 7, Article number: 40 (2018) | Cite this article

Quantifying hierarchy and dynamics in US faculty hiring and retention

K. Hunter Wapman, Sam Zhang, Aaron Clauset & Daniel B. Larremore

Nature 610, 120–127 (2022) | Cite this article

SCIENCE ADVANCES | RESEARCH ARTICLE

SOCIAL SCIENCES

Aspirational pursuit of mates in online dating markets

Elizabeth E. Bruch[1,2]* and M. E. J. Newman[2,3]

Motivation (Structure vs. Psychology)

Social hierarchy. directional bias in interactions, observed interactions tend to point from higher-ranked individuals to lower-ranked ones

U.S. academic system. strong self-reinforcing hierarchy of status and signaling.

Aspiration behavior. individuals routinely aim "higher" and judge themselves to be peers of those of higher status.

Labor advantages drive the greater productivity of faculty at elite universities

SAM ZHANG, K. HUNTER WAPMAN, DANIEL B. LARREMORE, AND AARON CLAUSET Authors Info & Affiliations

SCIENCE ADVANCES • 18 Nov 2022 • Vol 8, Issue 46 • DOI: 10.1126/sciadv.abq7056

Prestige drives epistemic inequality in the diffusion of scientific ideas

Allison C. Morgan, Dimitrios J. Economou, Samuel F. Way & Aaron Clauset

EPJ Data Science **7**, Article number: 40 (2018) | Cite this article

Quantifying hierarchy and dynamics in US faculty hiring and retention

K. Hunter Wapman, Sam Zhang, Aaron Clauset & Daniel B. Larremore

Nature **610**, 120–127 (2022) | Cite this article

SCIENCE ADVANCES | RESEARCH ARTICLE

SOCIAL SCIENCES

Aspirational pursuit of mates in online dating markets

Elizabeth E. Bruch[1,2]* and M. E. J. Newman[2,3]

Friendship networks and social status

Published online by Cambridge University Press: **15 April 2013**

BRIAN BALL and M.E.J. NEWMAN

Peer institutions (dataset to study the *psychology* behind structure)

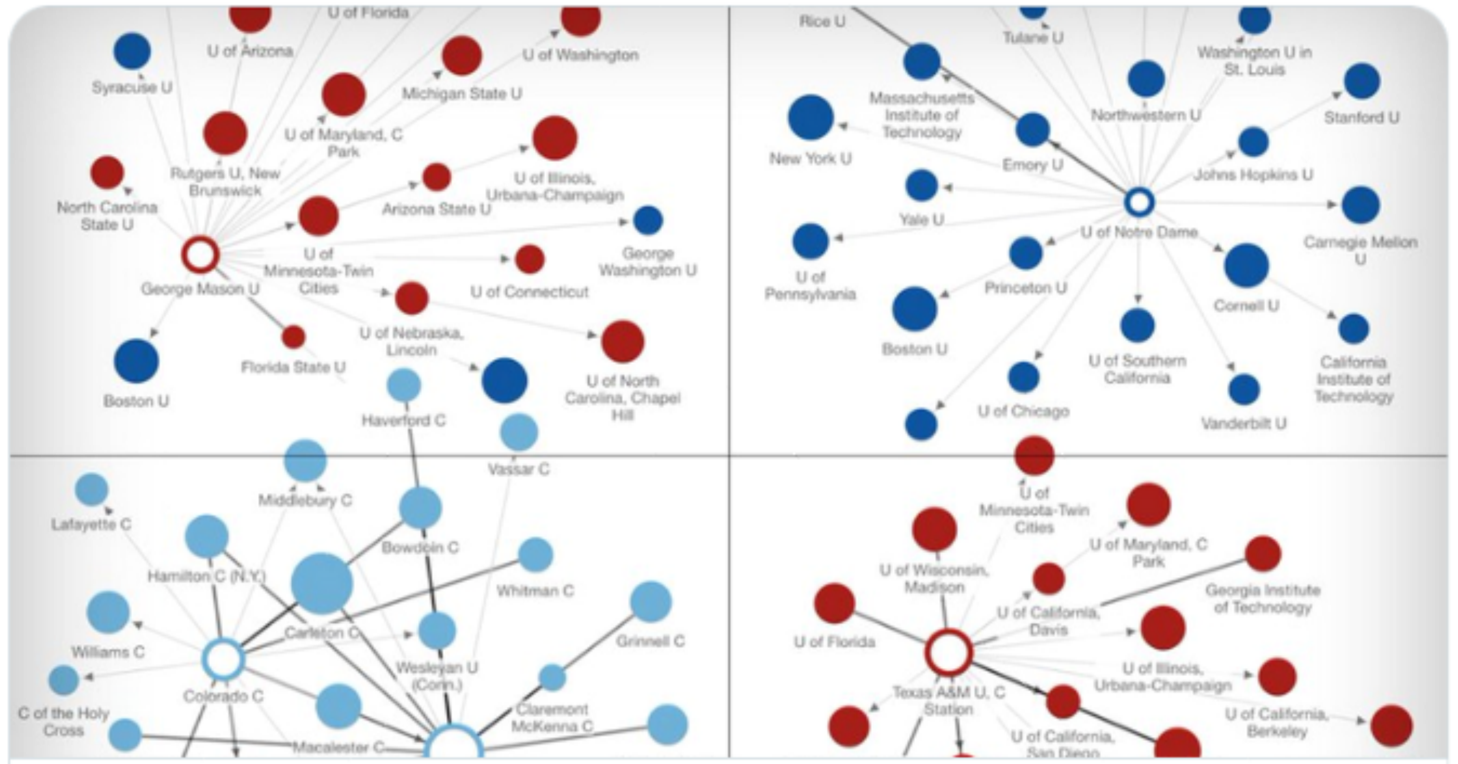
Neil Lewis, Jr., PhD
@NeilLewisJr

Harvard: Our peers are Princeton, Stanford, and Yale

Princeton: We have no peers, we're in a league of our own.

Cornell: All of the other Ivies are peers, except Dartmouth.

This dataset is interesting to explore:



chronicle.com
Who Does Your College Think Its Peers Are?
The Chronicle compiled the peer institutions for nearly 1,500 institutions from the 2020-21 year.

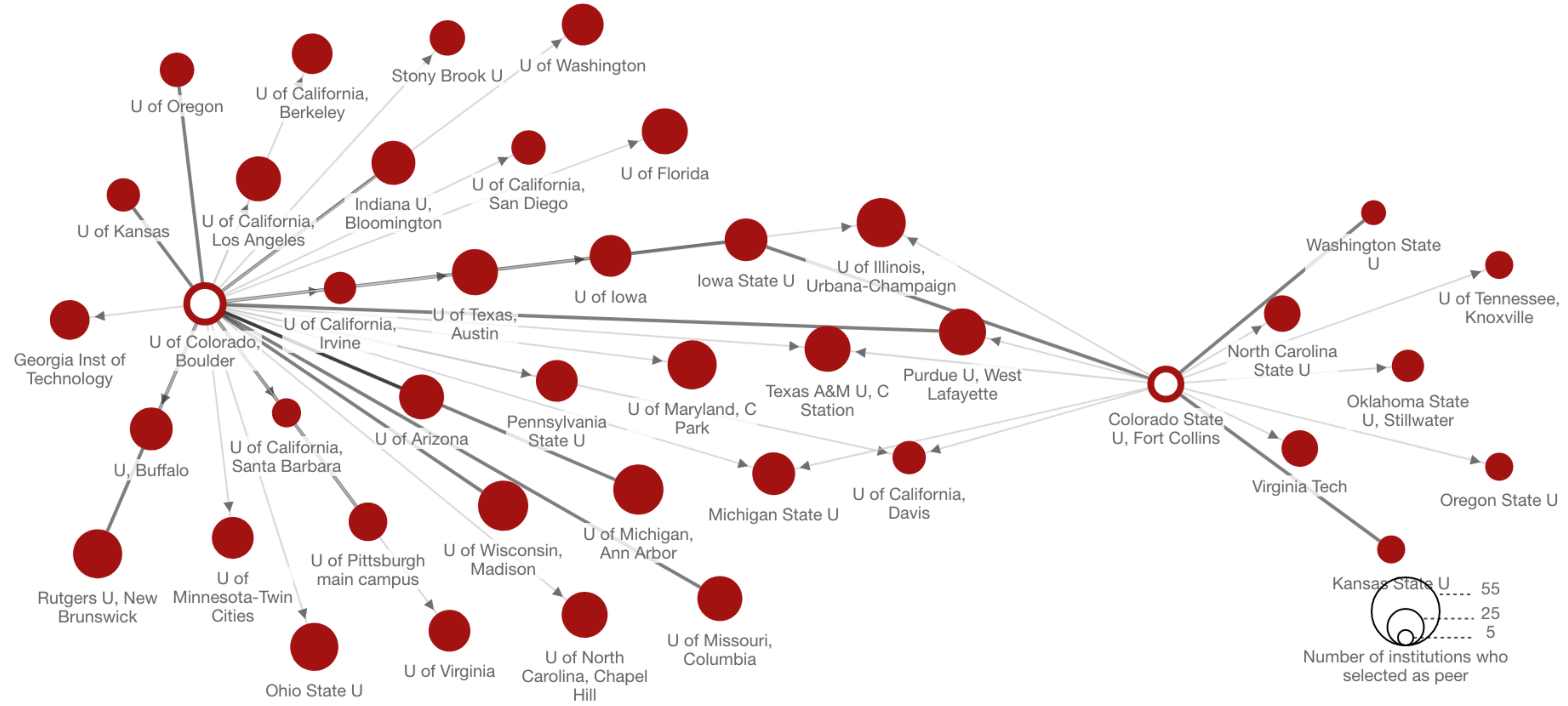
2:04 PM · Mar 27, 2022 · Twitter Web App

Peer group | Who this college chose | Who chose this college | Institution | Single institution | Compare institutions

Search for an institution X

Peer | Mutual Peer | Public Doctoral

Click an institution's circle below to select it for comparison.



University of Colorado at Boulder x
Boulder, CO | Public Doctoral

33 colleges selected as peers by this college. **25** colleges selected this college as a peer. **11** peer colleges that also chose this college as a peer.

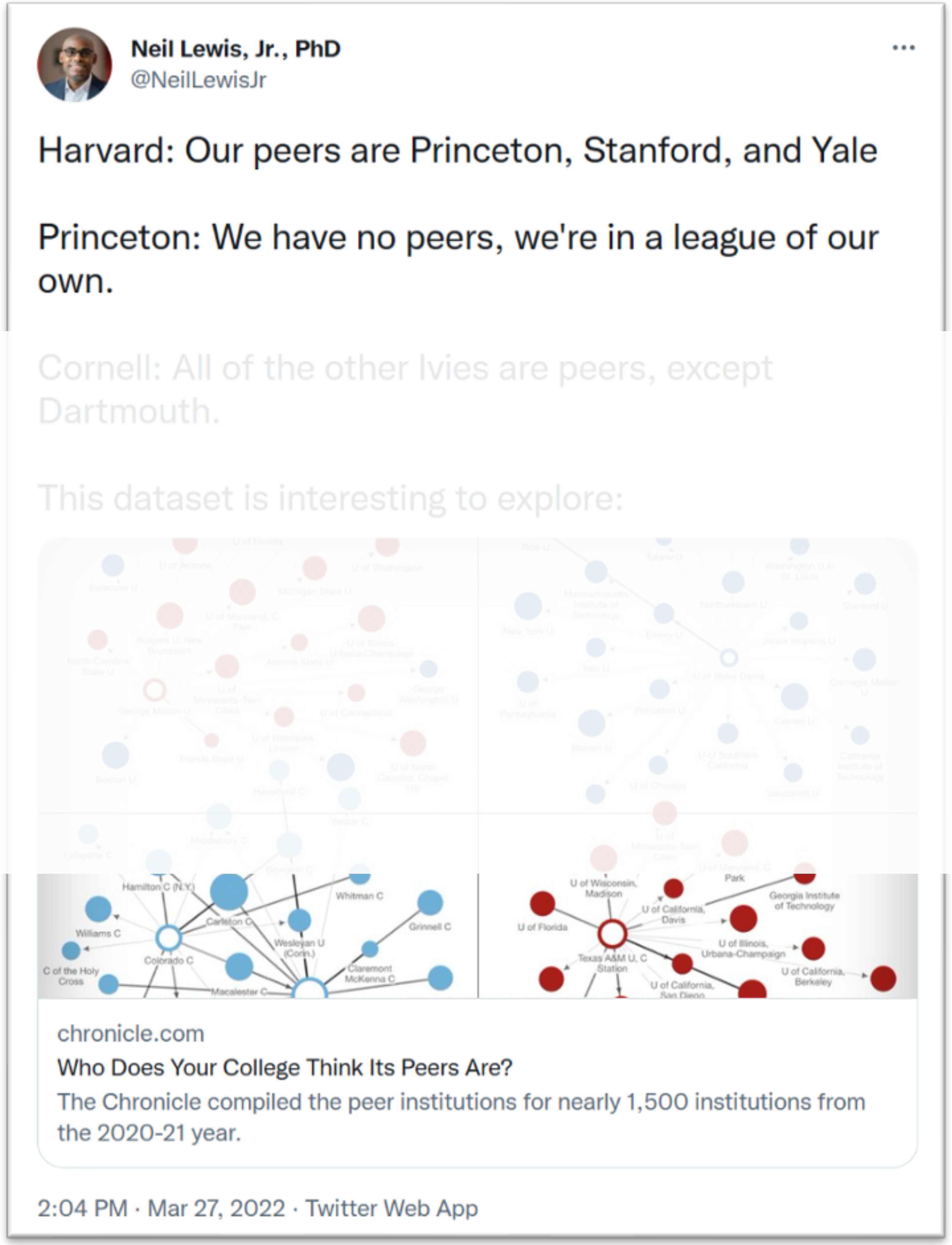
	University	Peer institutions
Admissions Rate	80%	47%
Graduation Rate	74%	83%
Enrollment	31,369	37,400

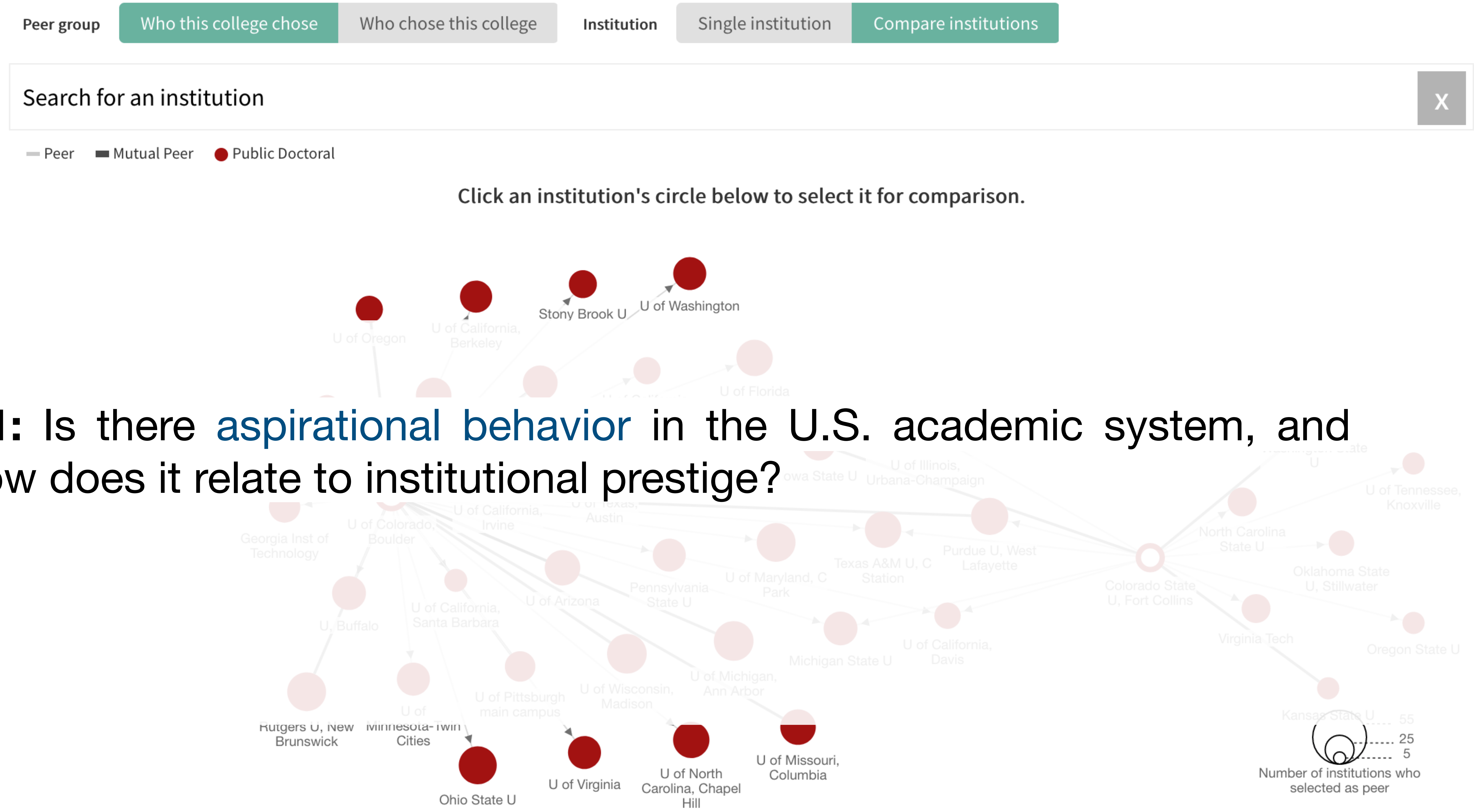
Colorado State University at Fort Collins x
Fort Collins, CO | Public Doctoral

13 colleges selected as peers by this college. **19** colleges selected this college as a peer. **3** peer colleges that also chose this college as a peer.

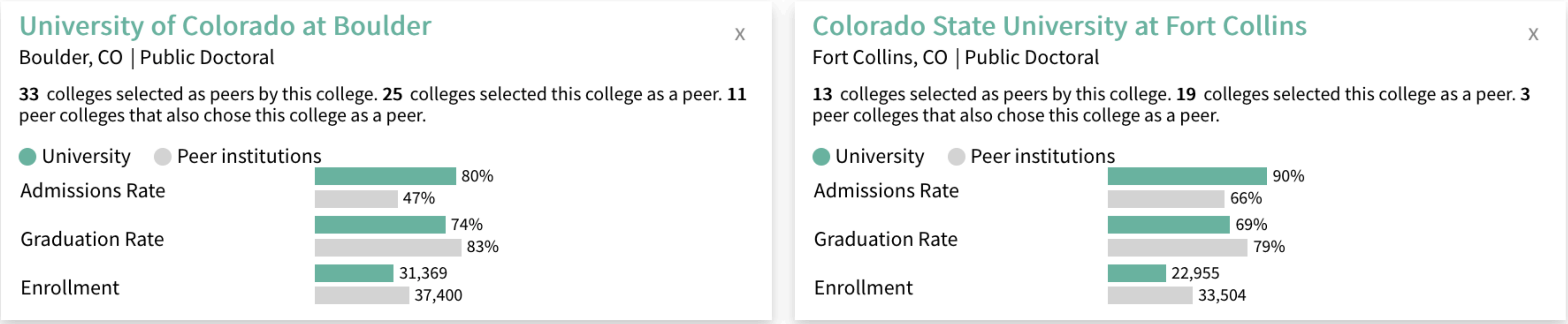
	University	Peer institutions
Admissions Rate	90%	66%
Graduation Rate	69%	79%
Enrollment	22,955	33,504

Peer institutions (dataset to study the *psychology* behind structure)

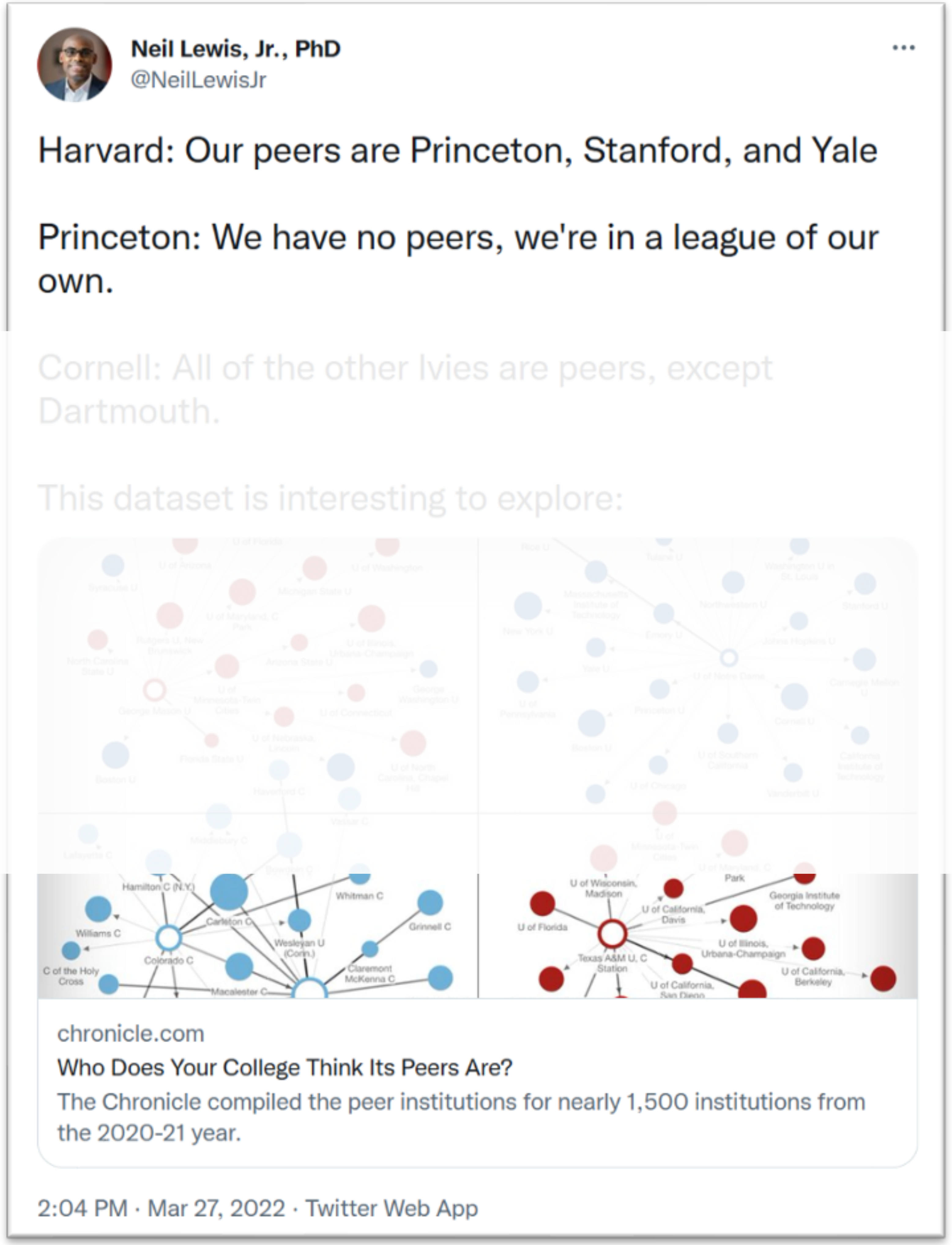




R1: Is there aspirational behavior in the U.S. academic system, and how does it relate to institutional prestige?



Peer institutions (dataset to study the *psychology* behind structure)


Neil Lewis, Jr., PhD
@NeilLewisJr

Harvard: Our peers are Princeton, Stanford, and Yale

Princeton: We have no peers, we're in a league of our own.

Cornell: All of the other Ivies are peers, except Dartmouth.

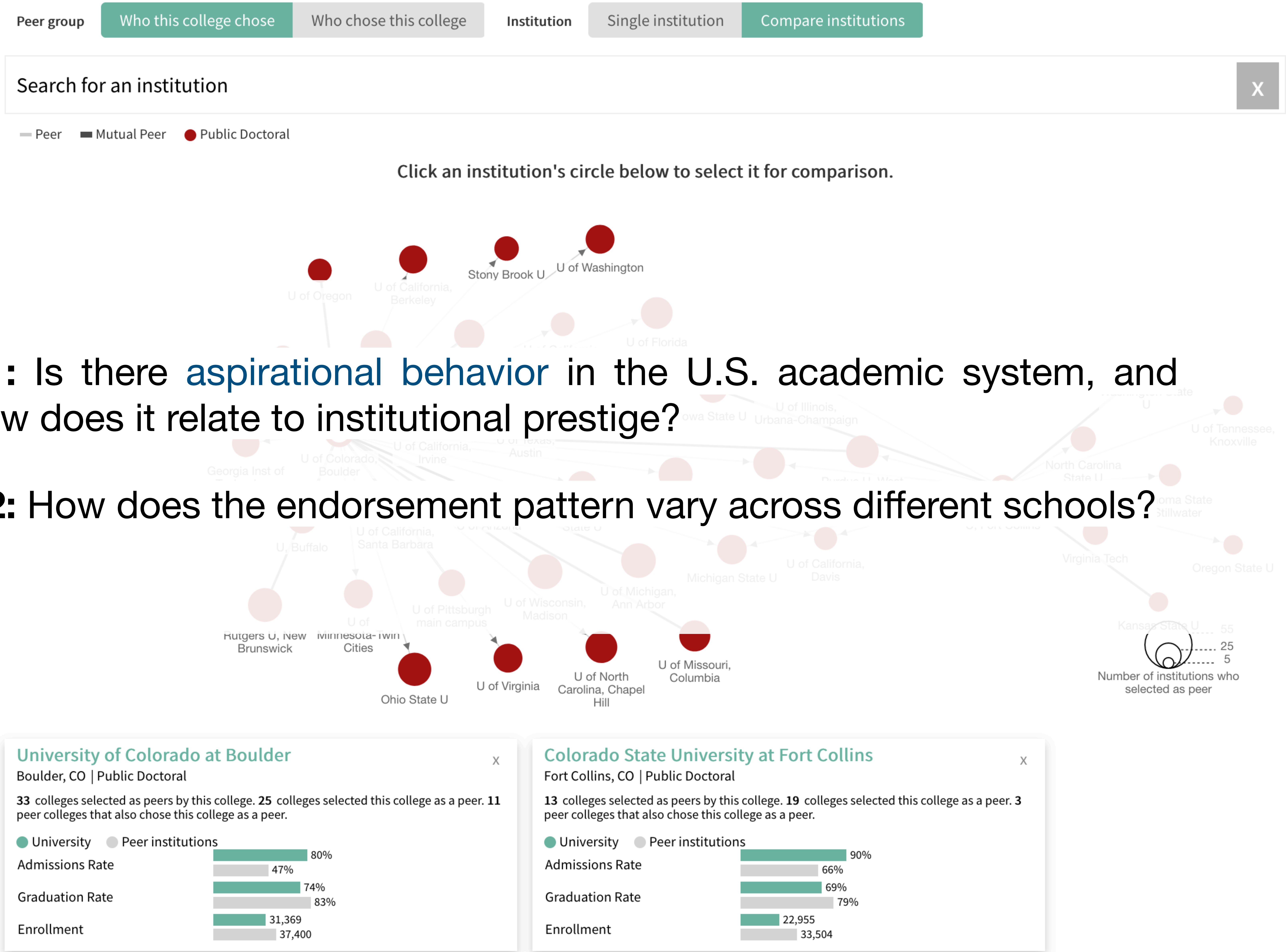
This dataset is interesting to explore:

chronicle.com
Who Does Your College Think Its Peers Are?
The Chronicle compiled the peer institutions for nearly 1,500 institutions from the 2020-21 year.

2:04 PM · Mar 27, 2022 · Twitter Web App

R1: Is there aspirational behavior in the U.S. academic system, and how does it relate to institutional prestige?

R2: How does the endorsement pattern vary across different schools?



Method (data collection)

Code	Type of Institutions	N
15 (R1)	Doctoral Universities: Highest Research Activity	125
16 (R2)	Doctoral Universities: Higher Research Activity	114
17 (R3)	Doctoral Universities: Moderate Research Activity	30
18 (M:Large)	Master's Colleges & Universities: Larger Programs	18
19 (M:Med)	Master's Colleges & Universities: Medium Programs	2
20 (M:Small)	Master's Colleges & Universities: Small Programs	1
21 (B:A&S)	Baccalaureate Colleges: Arts & Sciences Focus	6
26 (S:Health)	Special Focus Four-Year: Other Health Professions Schools	1
27 (S:Engineer)	Special Focus Four-Year: Engineering Schools	1
		total = 298

Aspiration: Peer institutions survey (IPEDS, 2020–21). Of the 1179 institutions and 18817 nominations collected, we matched the names with the faculty hiring dataset & restricted to Ph.D.-granting ones, resulting in 298 institutions.

Institutional prestige: Faculty hiring dataset (existing, 2011–12). Each institution is annotated by a data-driven estimate of its "prestige" within the faculty hiring system. Proxy for an institution's ability to place its graduates as faculty at other institutions.

Method (data collection)

Code	Type of Institutions	N
15 (R1)	Doctoral Universities: Highest Research Activity	125
16 (R2)	Doctoral Universities: Higher Research Activity	114
17 (R3)	Doctoral Universities: Moderate Research Activity	30
18 (M:Large)	Master's Colleges & Universities: Larger Programs	18
19 (M:Med)	Master's Colleges & Universities: Medium Programs	2
20 (M:Small)	Master's Colleges & Universities: Small Programs	1
21 (B:A&S)	Baccalaureate Colleges: Arts & Sciences Focus	6
26 (S:Health)	Special Focus Four-Year: Other Health Professions Schools	1
27 (S:Engineer)	Special Focus Four-Year: Engineering Schools	1

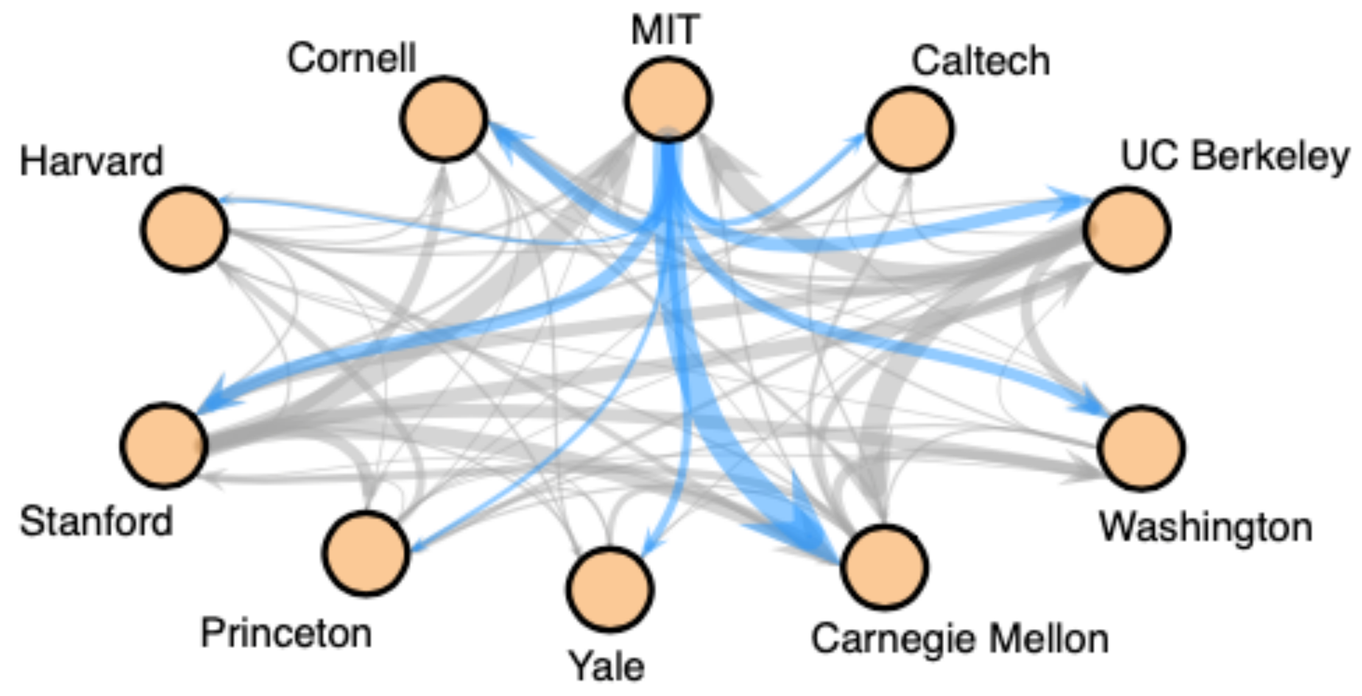
total = 298

Aspiration: Peer institutions survey (IPEDS, 2020–21). Of the 1179 institutions and 18817 nominations collected, we matched the names with the faculty hiring dataset & restricted to Ph.D.-granting ones, resulting in 298 institutions.

Institutional prestige: Faculty hiring dataset (existing, 2011–12). Each institution is annotated by a data-driven estimate of its "prestige" within the faculty hiring system. Proxy for an institution's ability to place its graduates as faculty at other institutions.

Method (quantifying prestige/desirability via SpringRank)

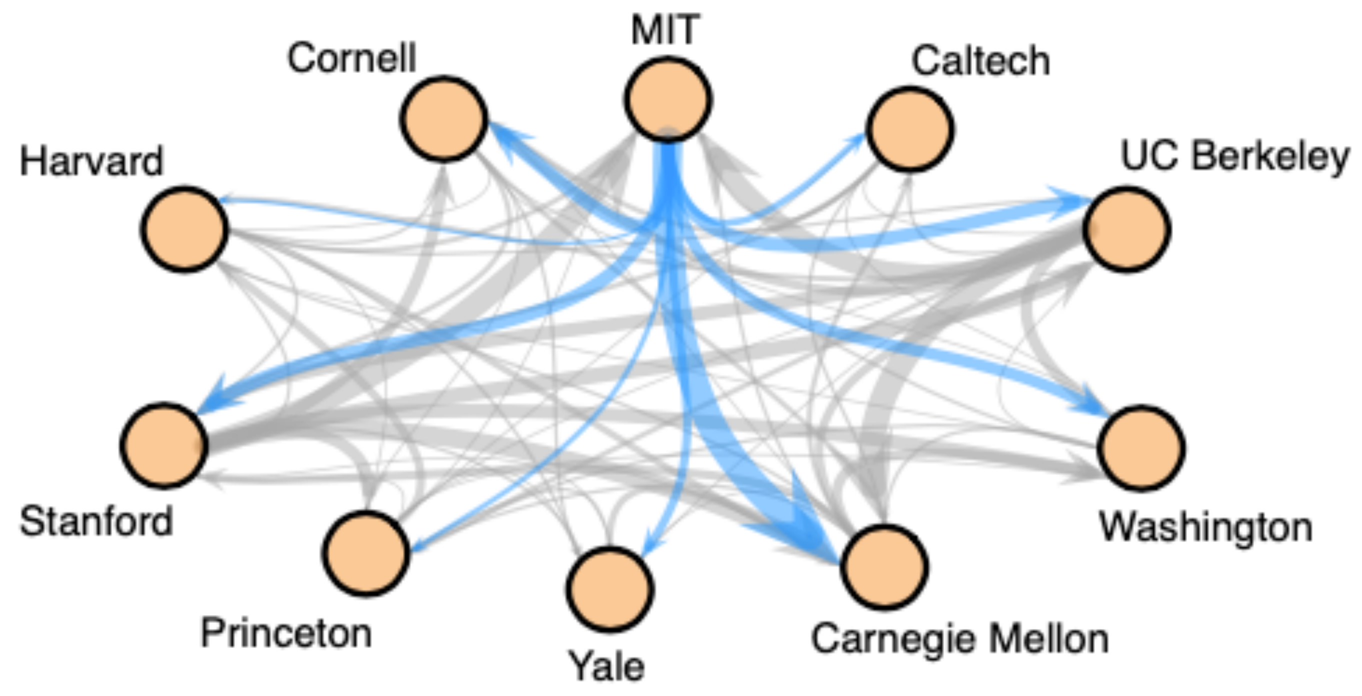
Each directed edge $u \to v$ implies
PhD from u becomes faculty at v



Computer science faculty hiring network; http://tuvalu.santafe.edu/~aaronc/facultyhiring/. See also: *Sci. Adv.* 1(1), e1400005, 2015.

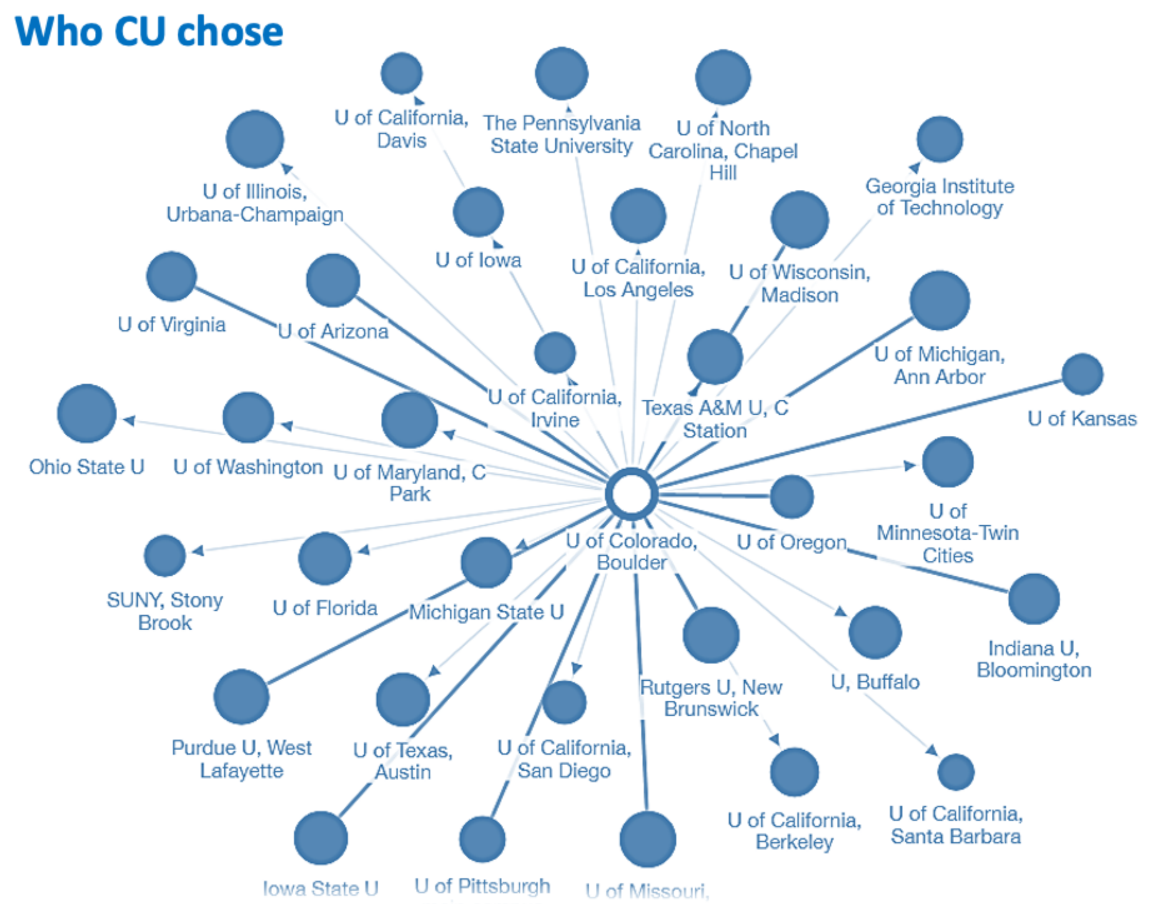
Method (quantifying prestige/desirability via SpringRank)

Each directed edge $u \to v$ implies PhD from u becomes faculty at v



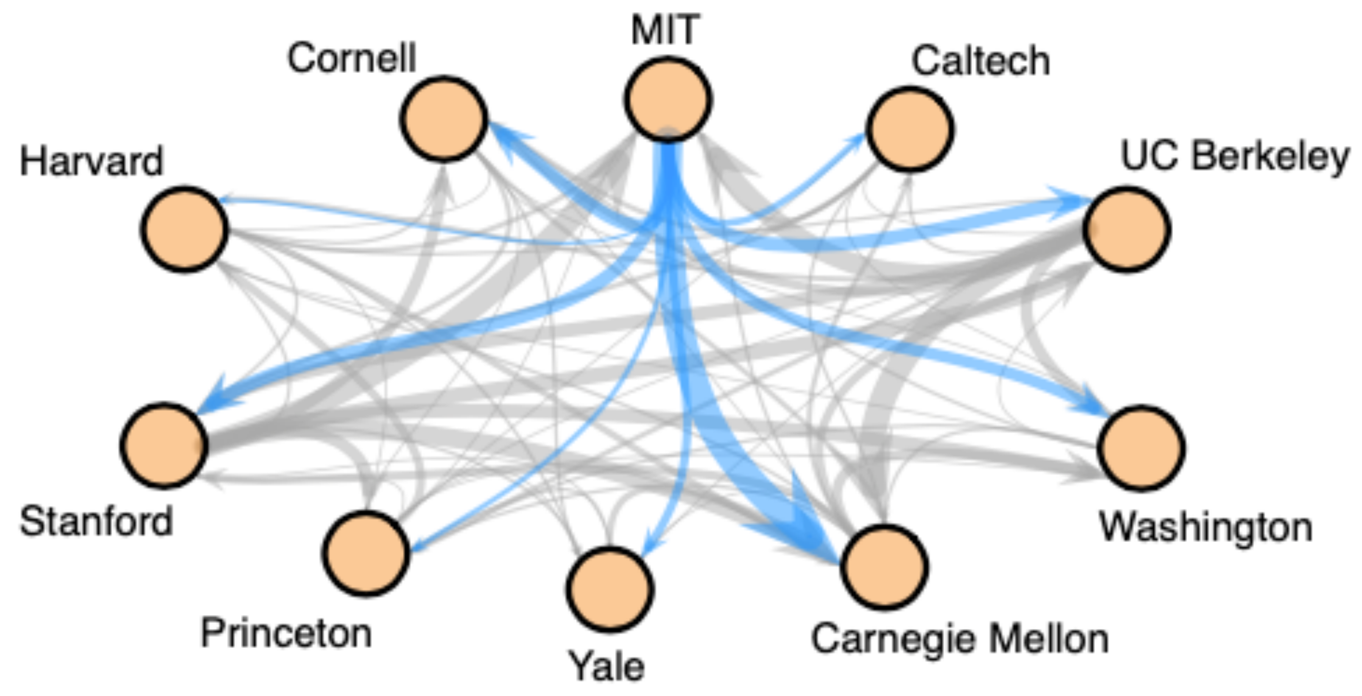
Computer science faculty hiring network; http://tuvalu.santafe.edu/~aaronc/facultyhiring/. See also: *Sci. Adv.* 1(1), e1400005, 2015.

Each directed edge $u \to v$ implies institution u selects institution v



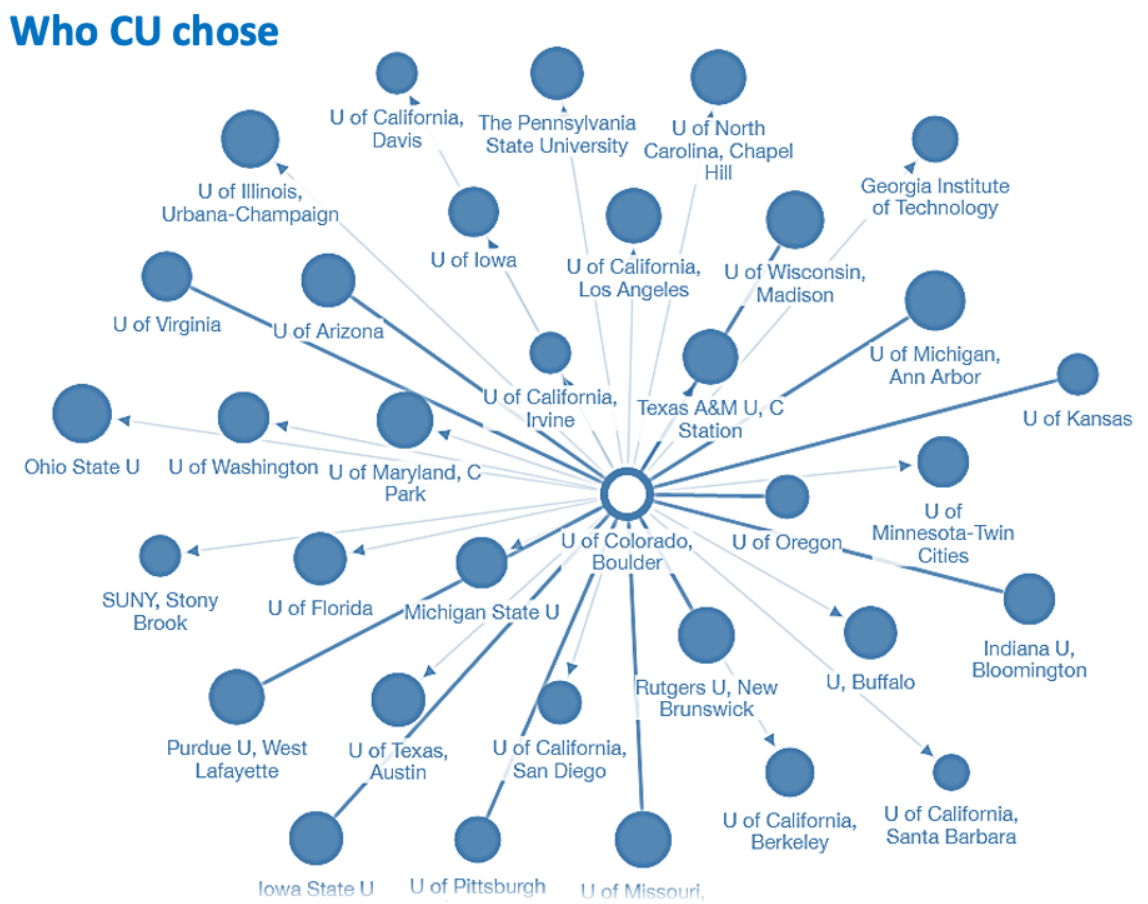
Method (quantifying prestige/desirability via SpringRank)

Each directed edge $u \to v$ implies PhD from u becomes faculty at v



Computer science faculty hiring network; http://tuvalu.santafe.edu/~aaronc/facultyhiring/. See also: *Sci. Adv.* 1(1), e1400005, 2015.

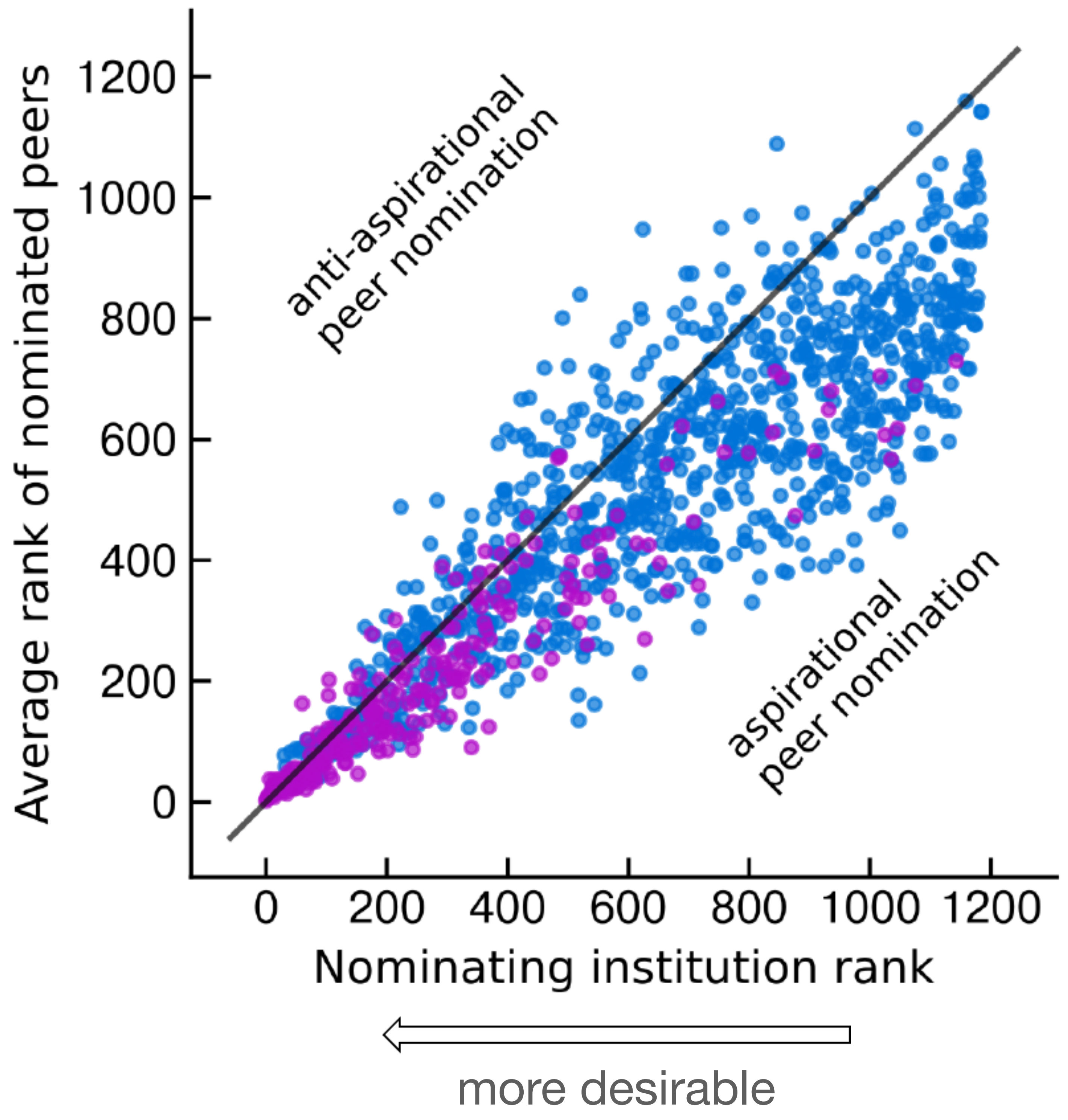
Each directed edge $u \to v$ implies institution u selects institution v

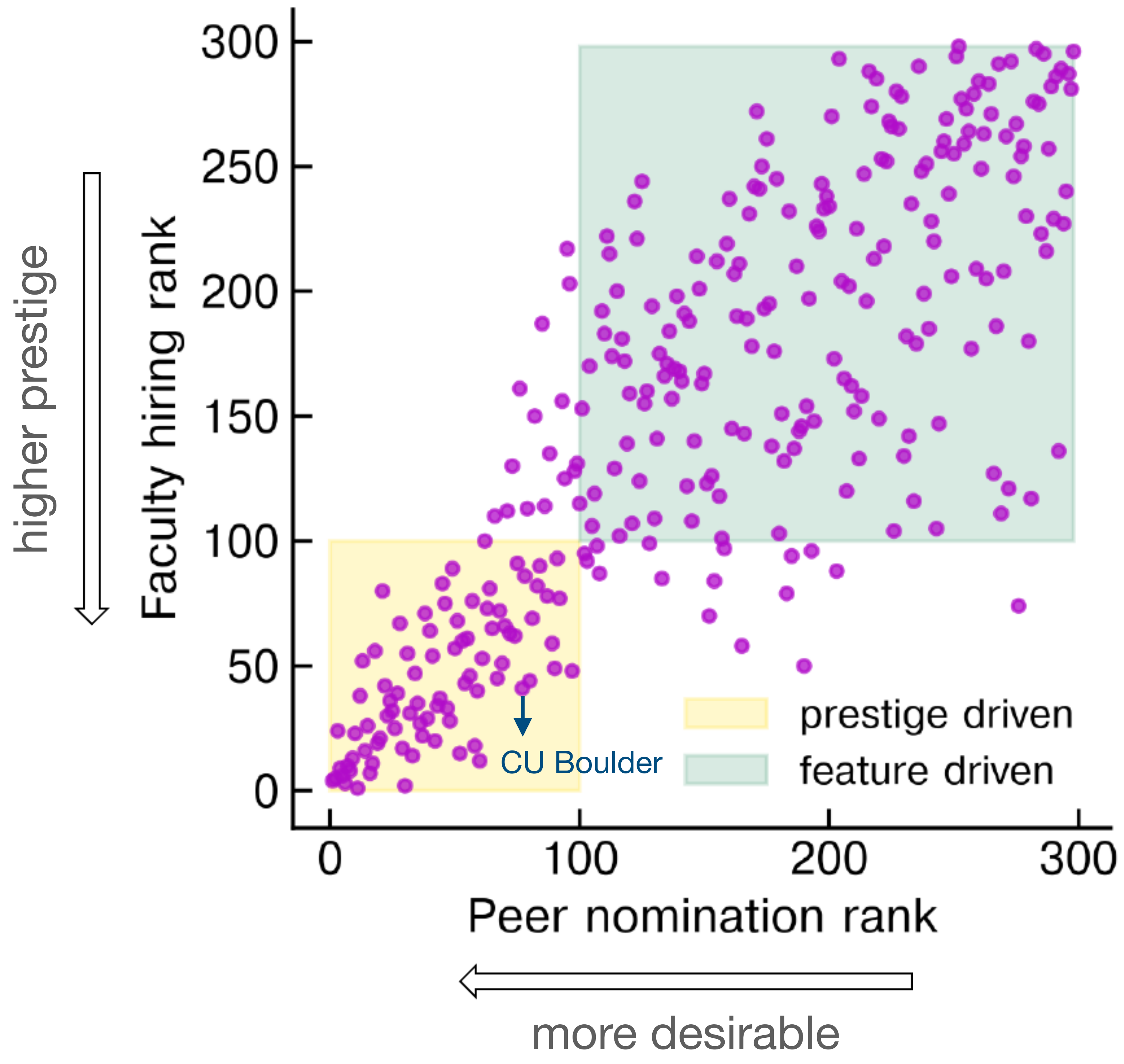


SpringRank: Let $\mathbf{A} = \{A_{ij}\}$ be the adjacency matrix of a directed multigraph. The SpringRank Model finds the ranking $\mathbf{s}^*$ of the nodes by solving:

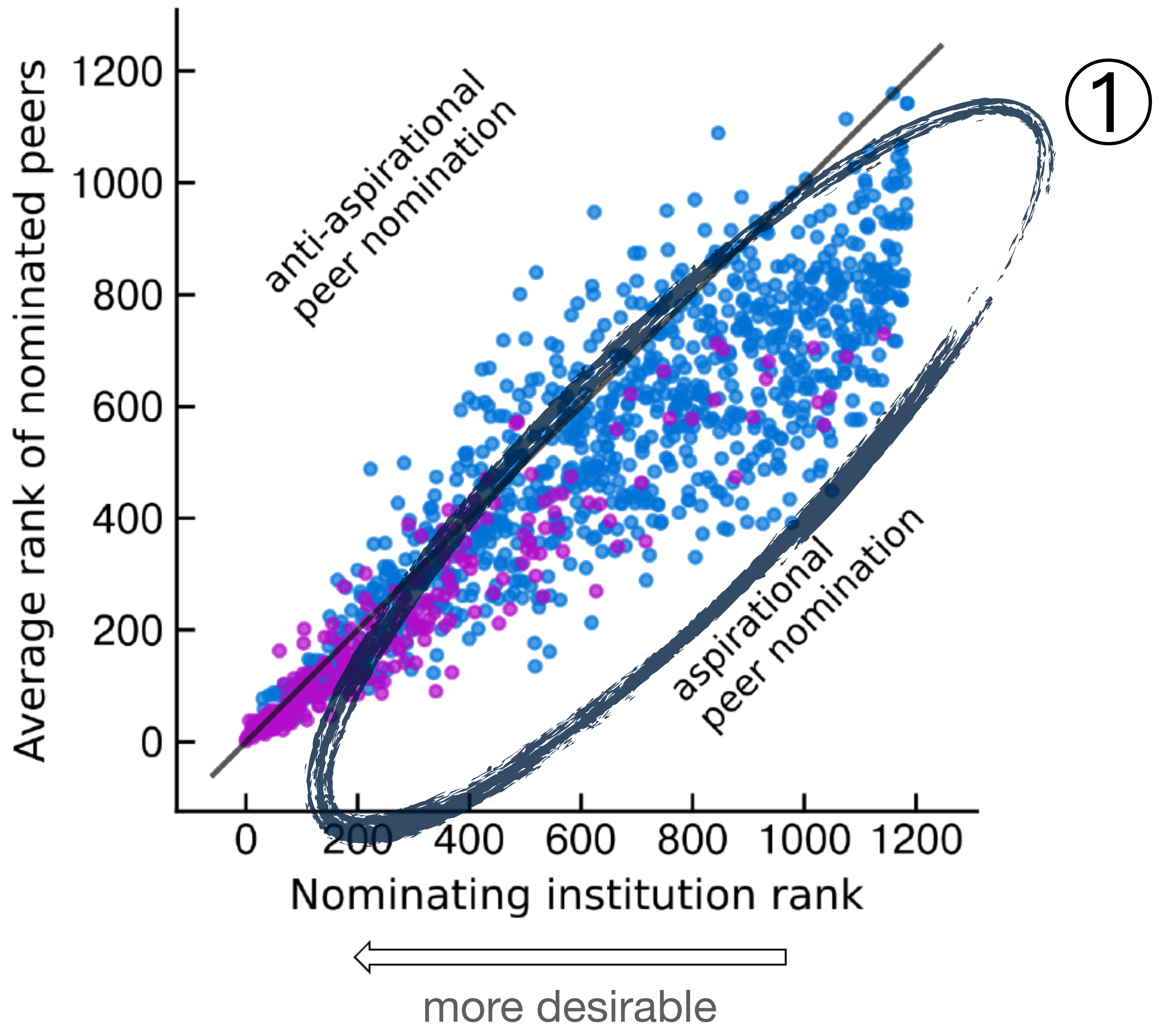
$$\underset{s \in \mathbb{R}^N}{\text{minimize}} \frac{1}{2} \sum_{ij} A_{ij}(s_i - s_j - 1)^2 \ .$$

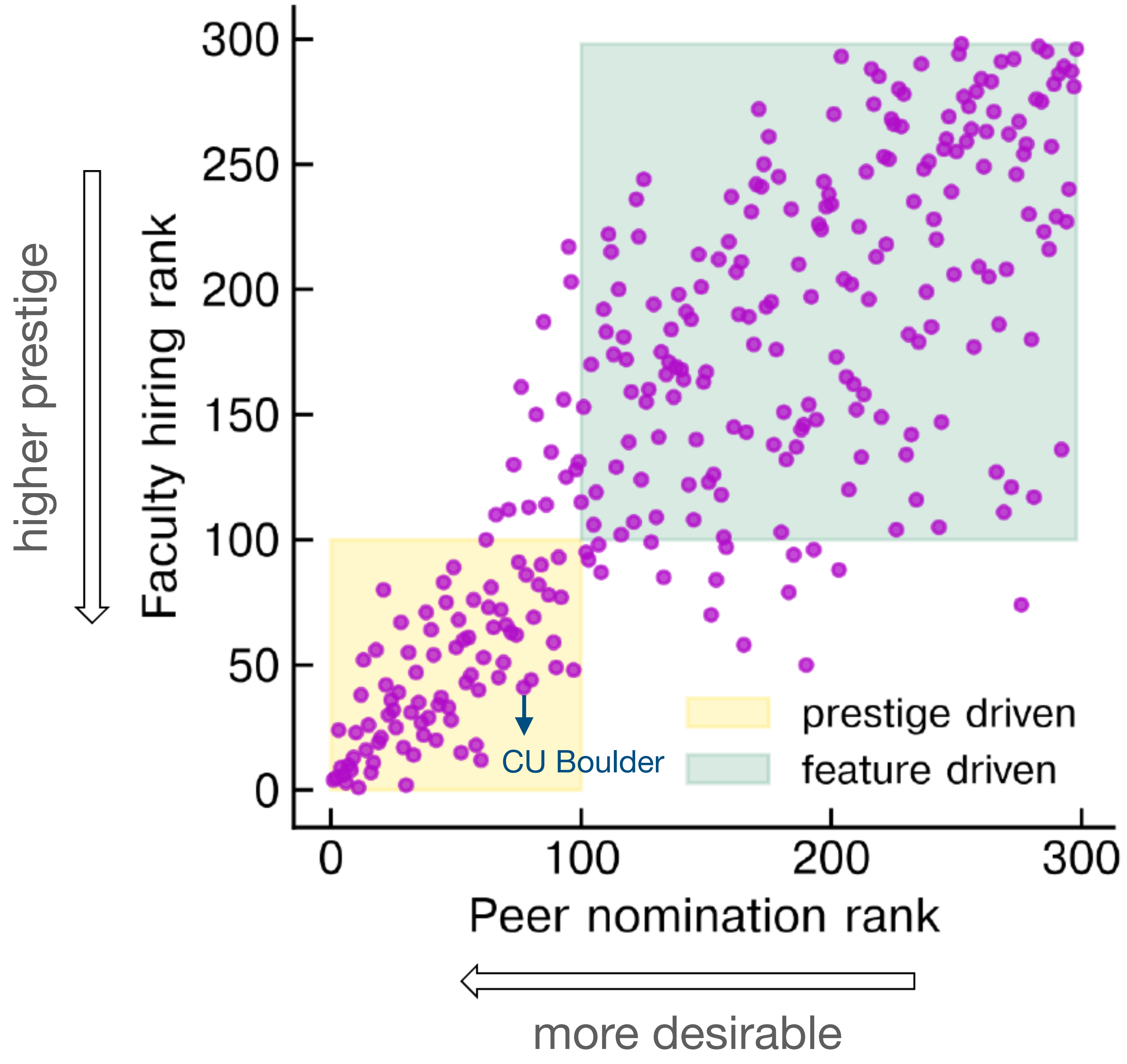
Institutions name higher prestige institutions as their peers



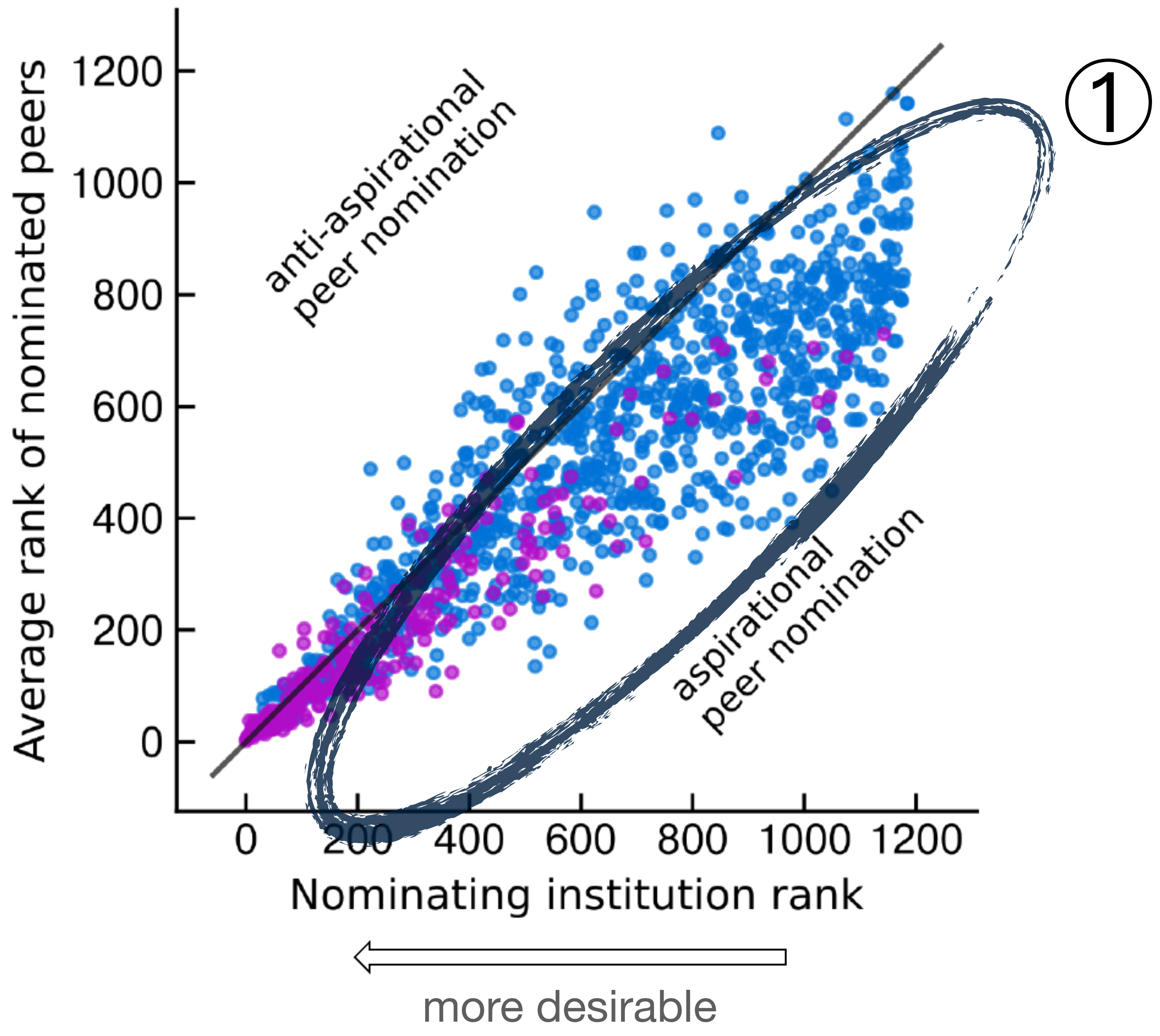


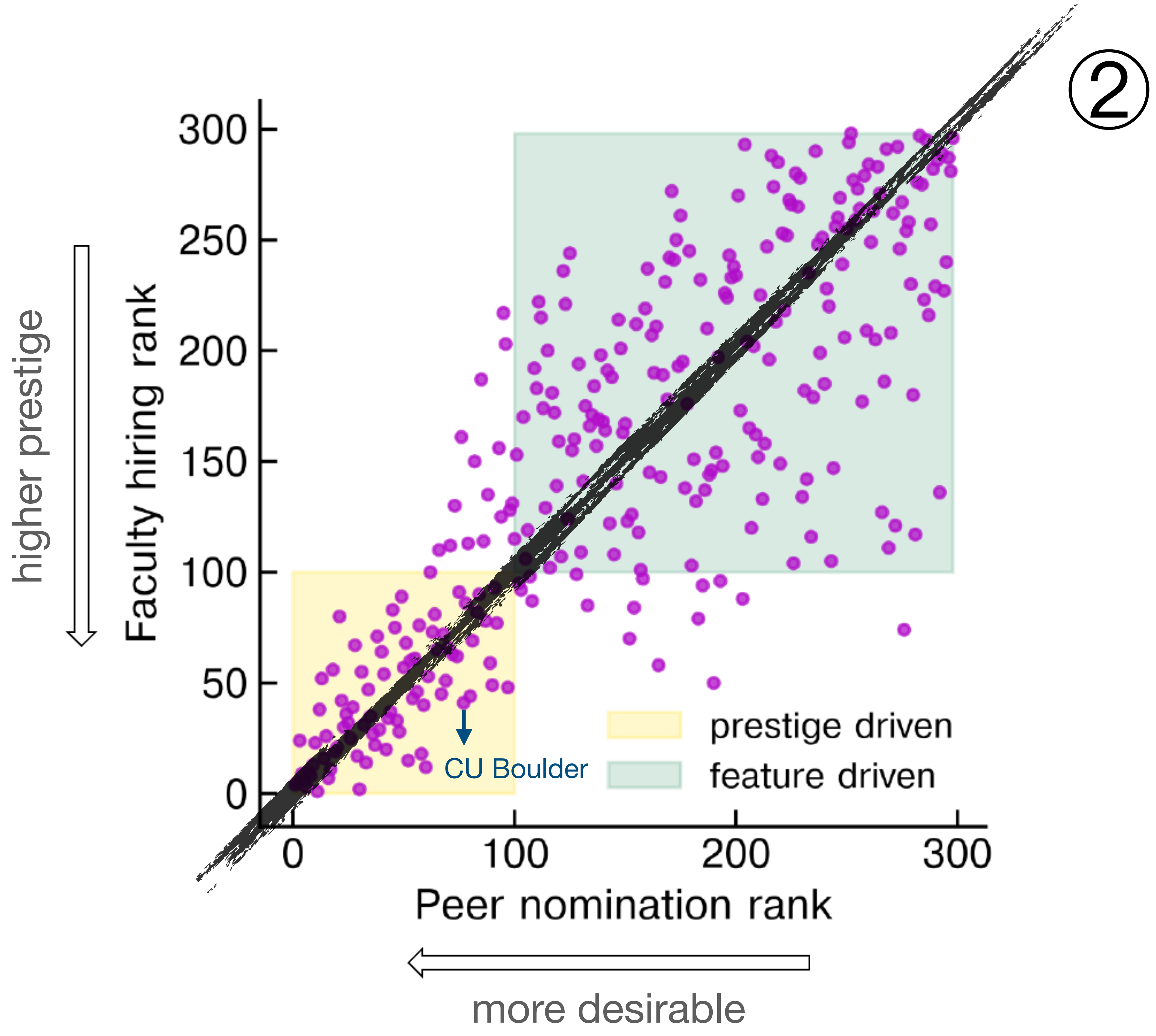
Institutions name higher prestige institutions as their peers



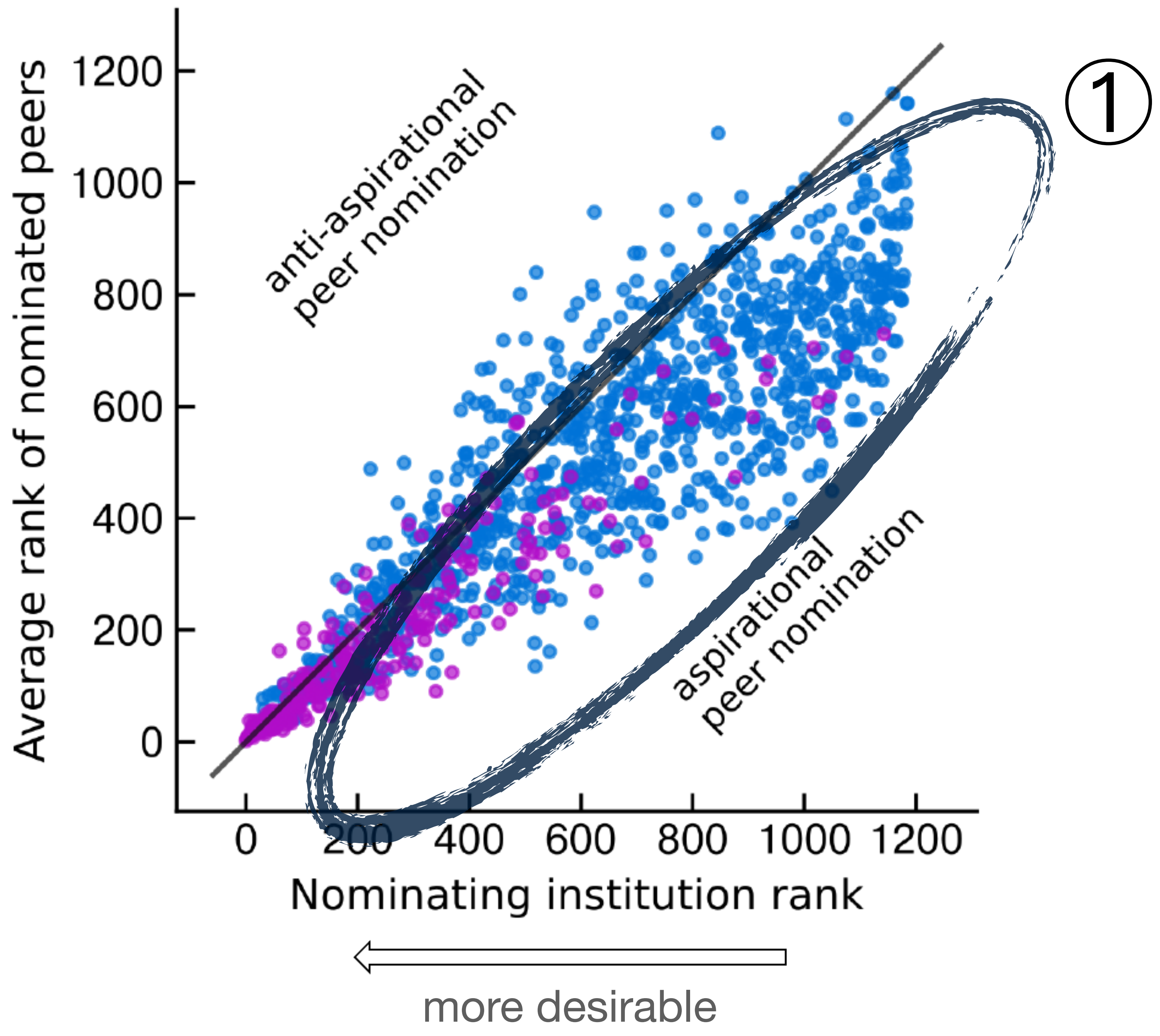


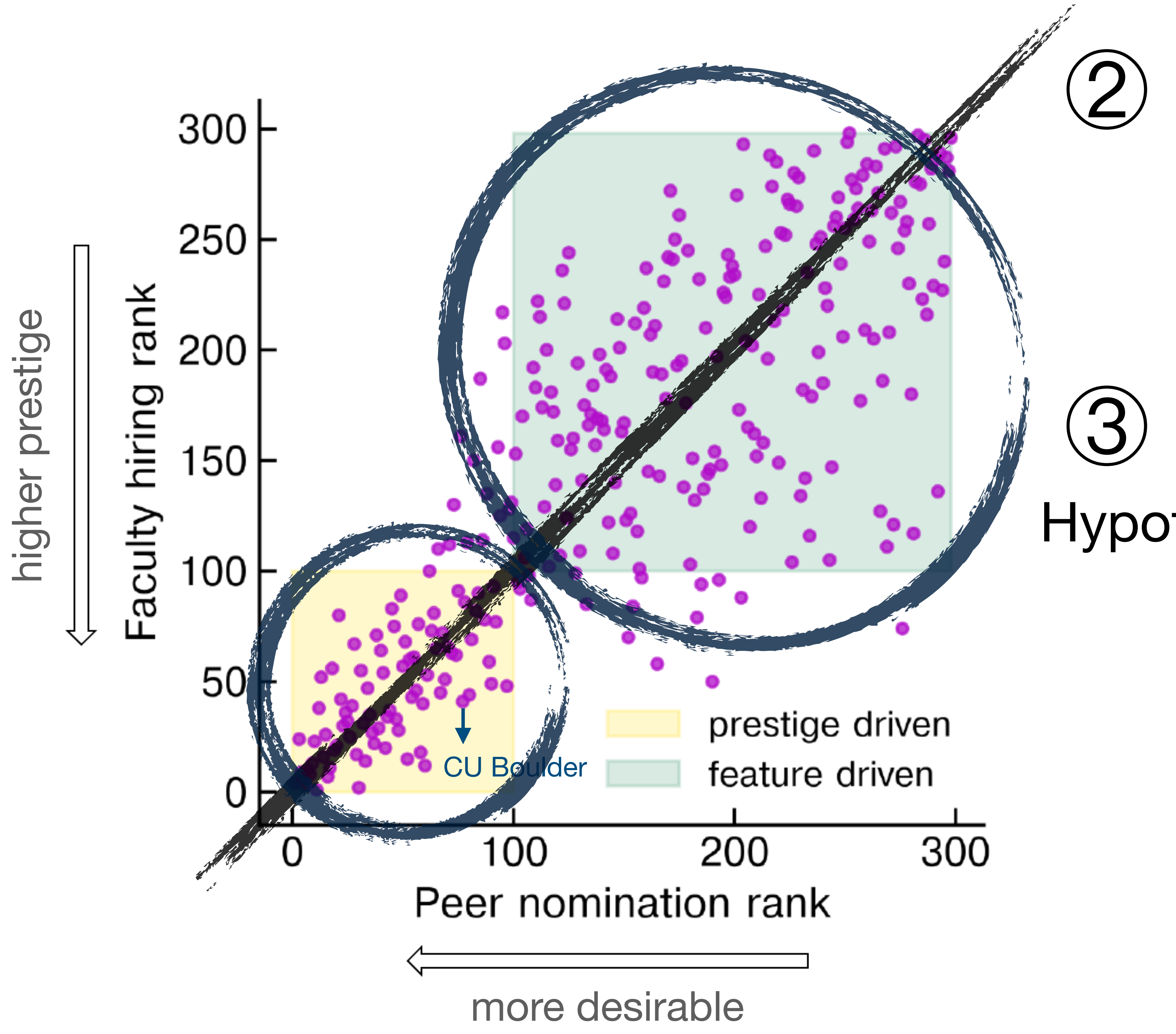
Institutions name higher prestige institutions as their peers





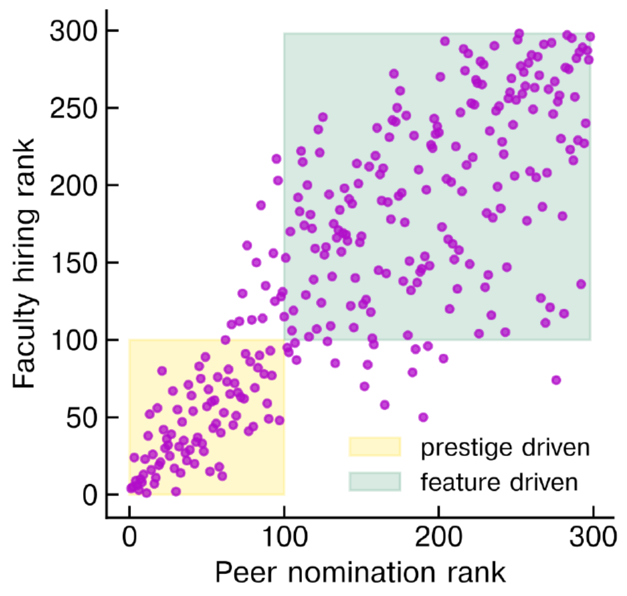
Institutions name higher prestige institutions as their peers





Hypothesized strata 🤔

Method (logistic regression with LASSO and Laplacian regularization*)



$$\underset{\boldsymbol{\beta}\in\mathbb{R}^p}{\text{minimize}}\, \ell(\boldsymbol{y}, \boldsymbol{\beta}^\mathsf{T}\boldsymbol{x}) + \underbrace{\lambda\|\boldsymbol{\beta}\|_1}_{\text{LASSO } (\ell 1 \text{ regularization})}$$

(logistic) loss

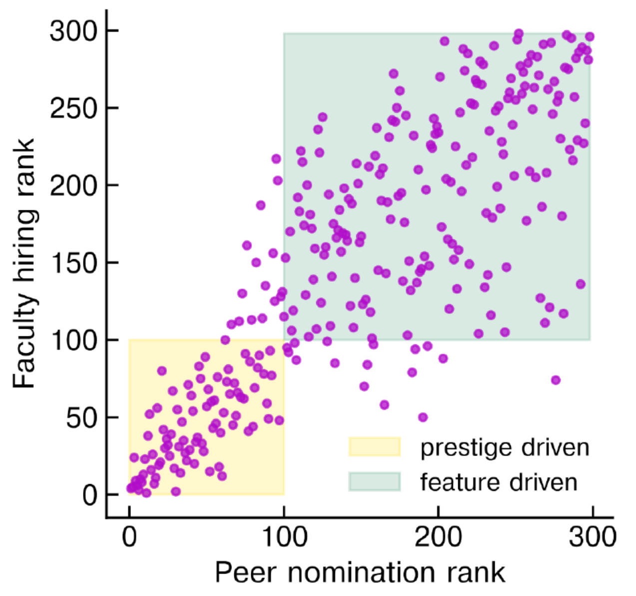
actual data

predicted data

Institutional features (nodal features)	Type	Total features	Total derived features (one for each edge)
Historically black?	binary	2	4
Sector (public or private)	categorical	2	4
Carnegie Classification	categorical	9	78
Two-letter state abbreviation	categorical	50	2494
Avg enrollment of full-time students	numeric	1	2
Graduation rates	numeric	1	2
Admission rates	numeric	1	2
Prestige rank	numeric	1	2

*Solution method: alternating direction method of multipliers (ADMM), see Tuck, Barrett, Boyd in J. Machine Learning Research 22, 1-37 (2021).

Method (logistic regression with LASSO and Laplacian regularization*)



$$\underset{\boldsymbol{\beta}\in\mathbb{R}^p}{\text{minimize}}\,\ell(\boldsymbol{y},\underline{\boldsymbol{\beta}^\intercal \boldsymbol{x}}) + \underline{\lambda\|\boldsymbol{\beta}\|_1}$$

LASSO (ℓ1 regularization)

use when #features / #samples ~ 1

(logistic) loss

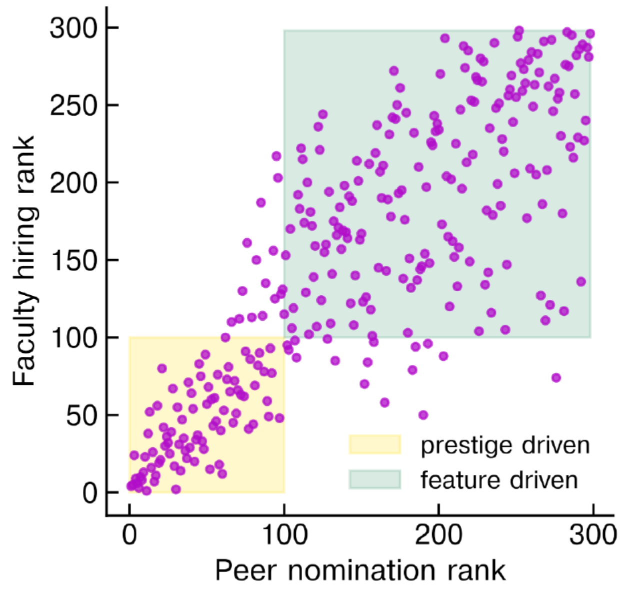
actual data

predicted data

Institutional features (nodal features)	Type	Total features	Total derived features (one for each edge)
Historically black?	binary	2	4
Sector (public or private)	categorical	2	4
Carnegie Classification	categorical	9	78
Two-letter state abbreviation	categorical	50	2494
Avg enrollment of full-time students	numeric	1	2
Graduation rates	numeric	1	2
Admission rates	numeric	1	2
Prestige rank	numeric	1	2

*Solution method: alternating direction method of multipliers (ADMM), see Tuck, Barrett, Boyd in J. Machine Learning Research 22, 1-37 (2021).

Method (logistic regression with LASSO and Laplacian regularization*)



$$\underset{\boldsymbol{\beta}\in\mathbb{R}^p}{\text{minimize}}\, \ell(\boldsymbol{y}, \underline{\boldsymbol{\beta}^\mathsf{T}\boldsymbol{x}}) + \underline{\lambda\|\boldsymbol{\beta}\|_1}$$

(logistic) loss

actual data

predicted data

LASSO (ℓ1 regularization)

use when #features / #samples ~ 1

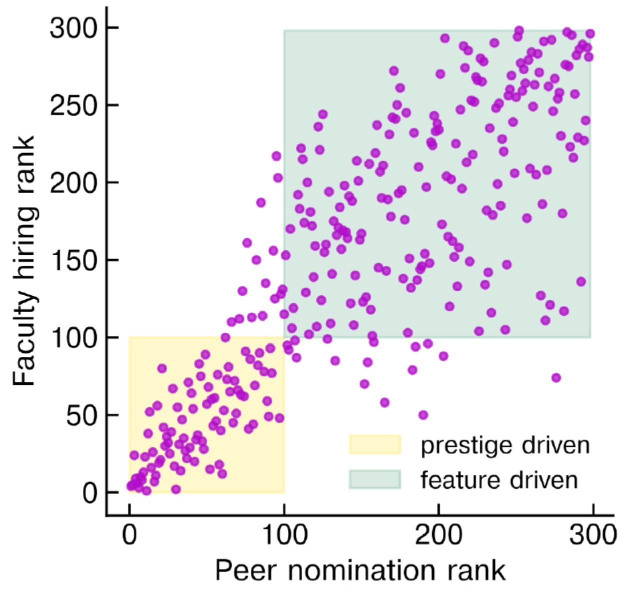
Institutional features (nodal features)	Type	Total features	Total derived features (one for each edge)
Historically black?	binary	2	4
Sector (public or private)	categorical	2	4
Carnegie Classification	categorical	9	78
Two-letter state abbreviation	categorical	50	2494
Avg enrollment of full-time students	numeric	1	2
Graduation rates	numeric	1	2
Admission rates	numeric	1	2
Prestige rank	numeric	1	2

$$\begin{aligned}\underset{\boldsymbol{\beta}_\text{high},\boldsymbol{\beta}_\text{low}\in\mathbb{R}^p}{\text{minimize}}\, & \ell(\boldsymbol{y}_\text{high}, \boldsymbol{\beta}_\text{high}^\mathsf{T}\boldsymbol{x}) + \lambda\|\boldsymbol{\beta}_\text{high}\|_1 \\ & + \ell(\boldsymbol{y}_\text{low}, \boldsymbol{\beta}_\text{low}^\mathsf{T}\boldsymbol{x}) + \lambda\|\boldsymbol{\beta}_\text{low}\|_1 \\ & + \lambda_L\|\boldsymbol{\beta}_\text{high} - \boldsymbol{\beta}_\text{low}\|_2^2\end{aligned}$$

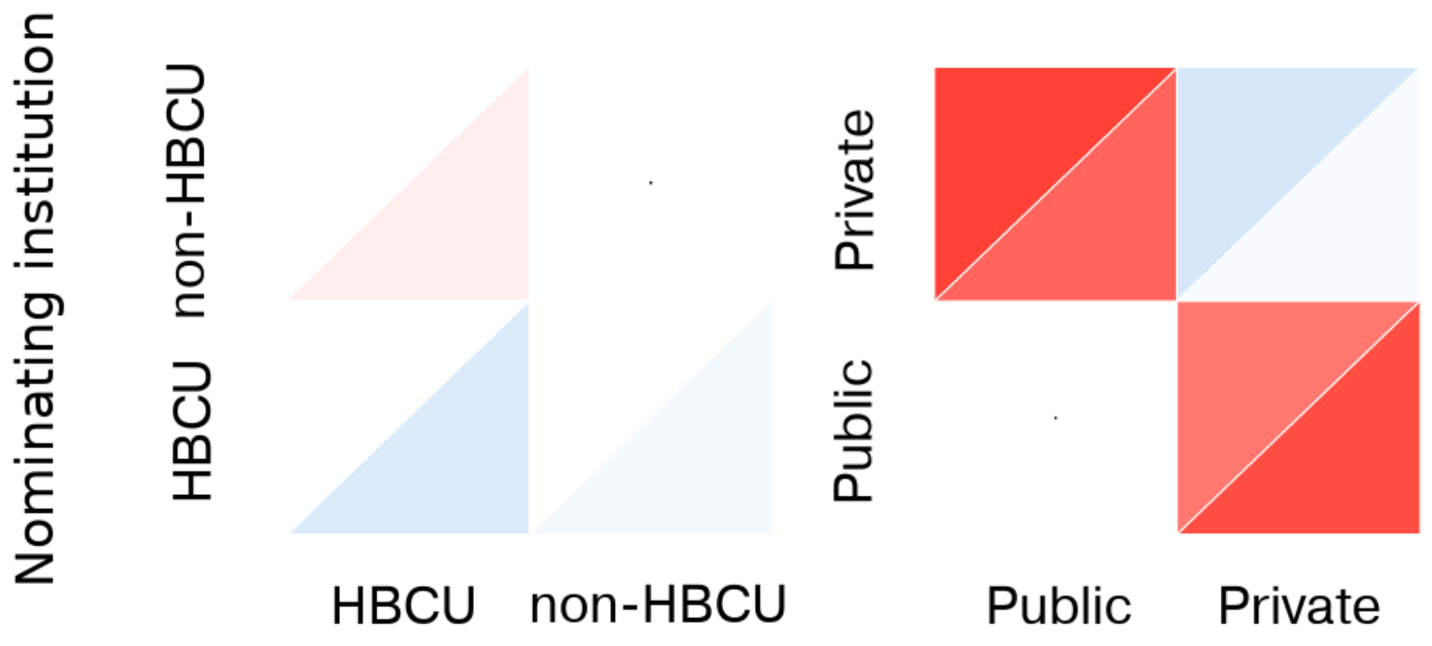
Idea: Parameters in adjacent strata should be similar. Plus, parameters are encouraged to be zero (not used).

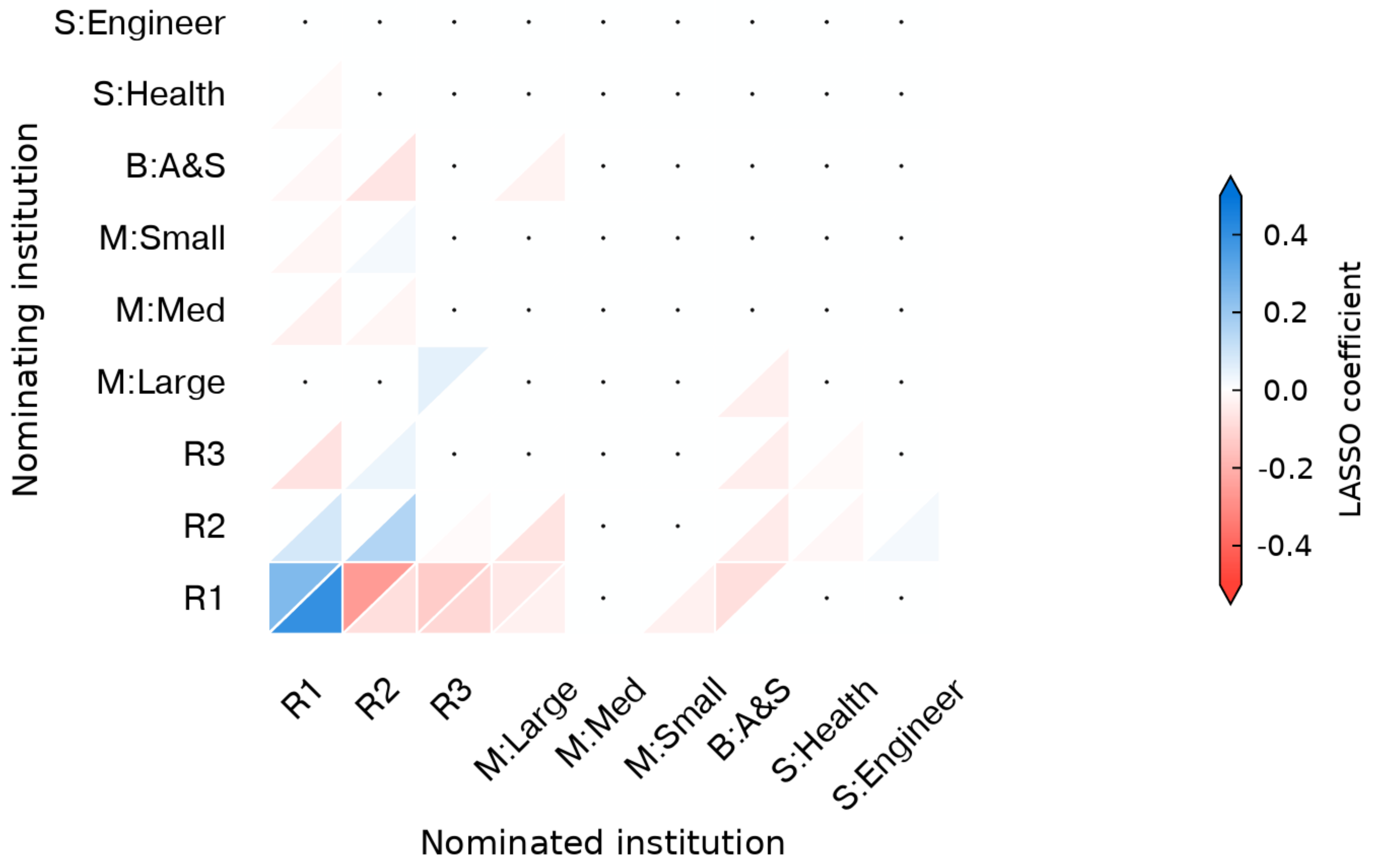
*Solution method: alternating direction method of multipliers (ADMM), see Tuck, Barrett, Boyd in J. Machine Learning Research 22, 1-37 (2021).

Rich peer decision patterns (categorical features)



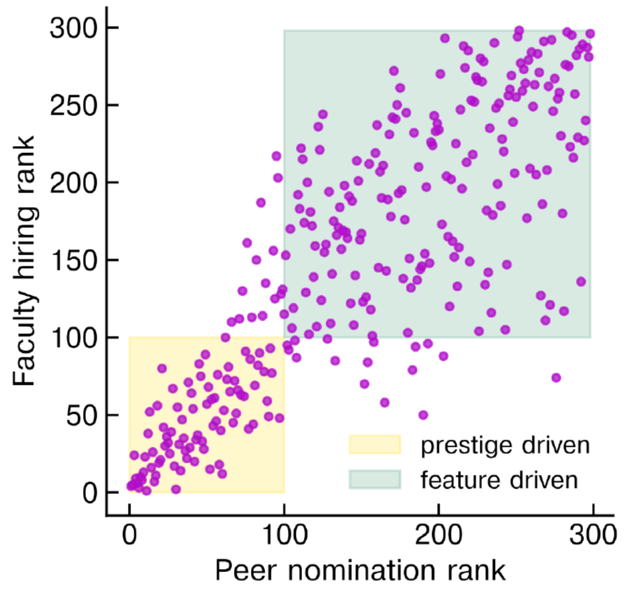
◸ : higher prestige
◺ : lower prestige



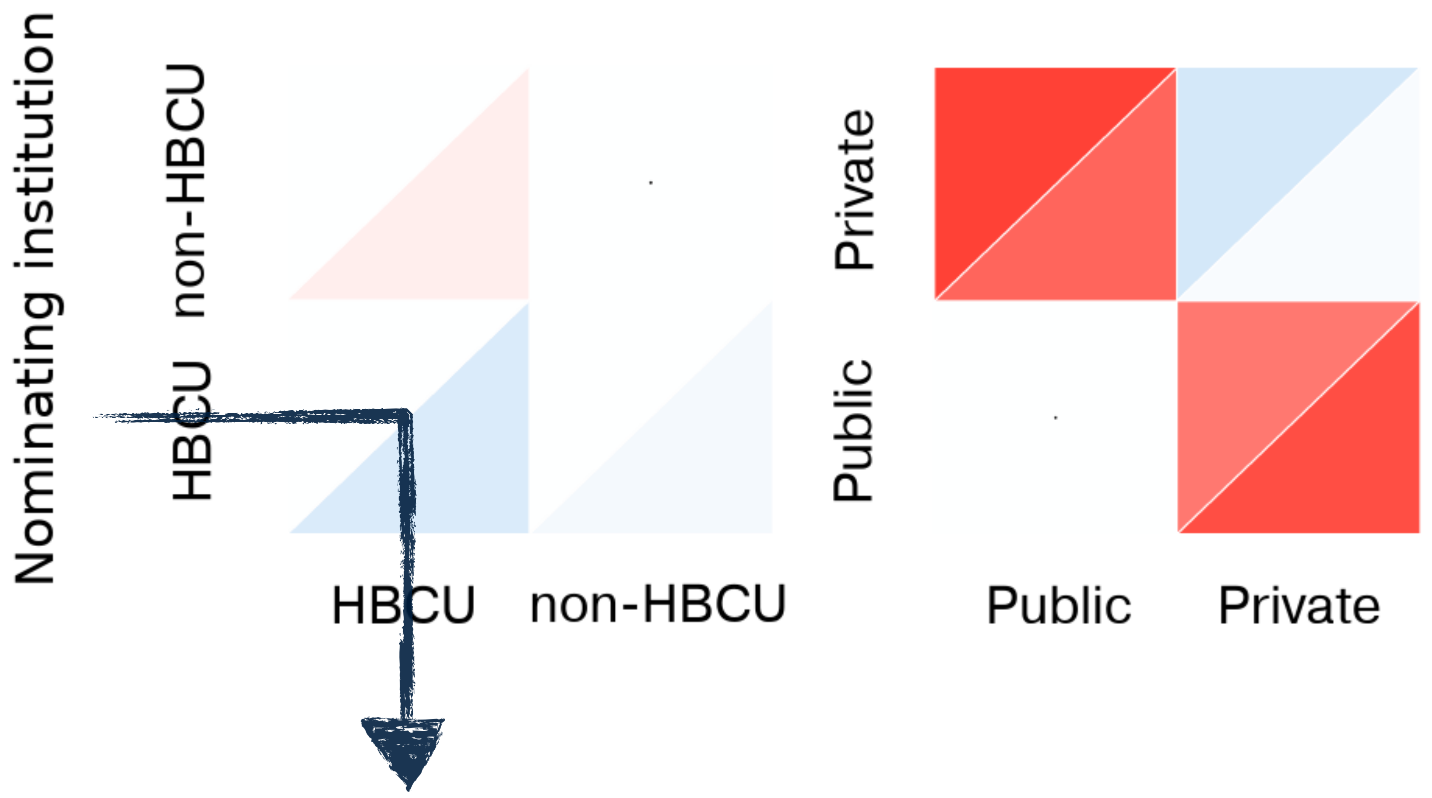


*Model validated via the average negative log likelihood (ANLL): stratified model (0.863); separate model (0.864); common model (0.912)

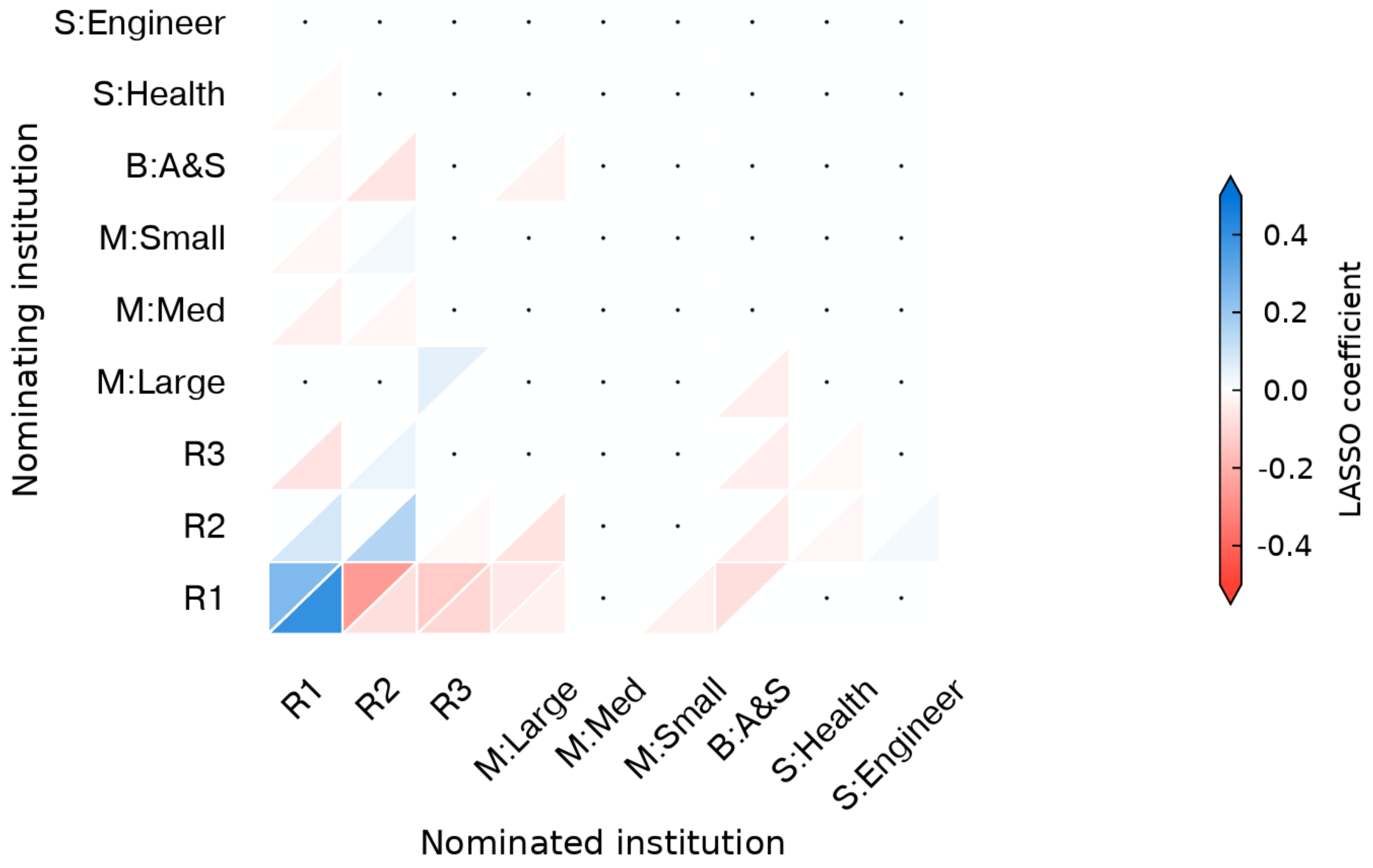
Rich peer decision patterns (categorical features)



◸ : higher prestige
◿ : lower prestige

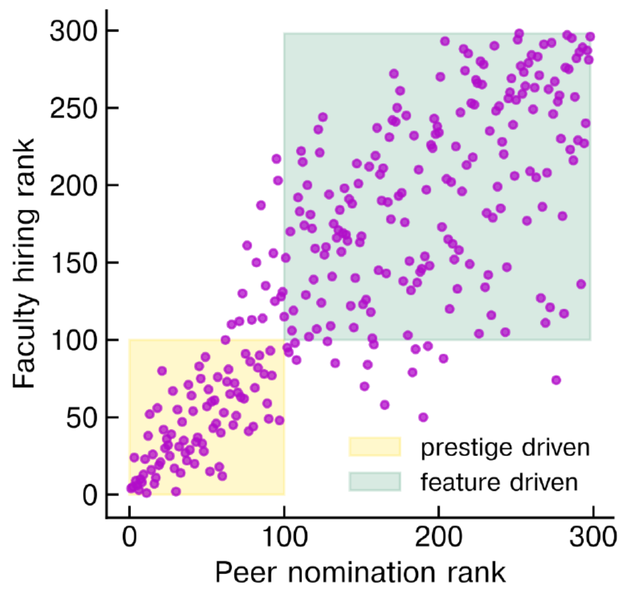


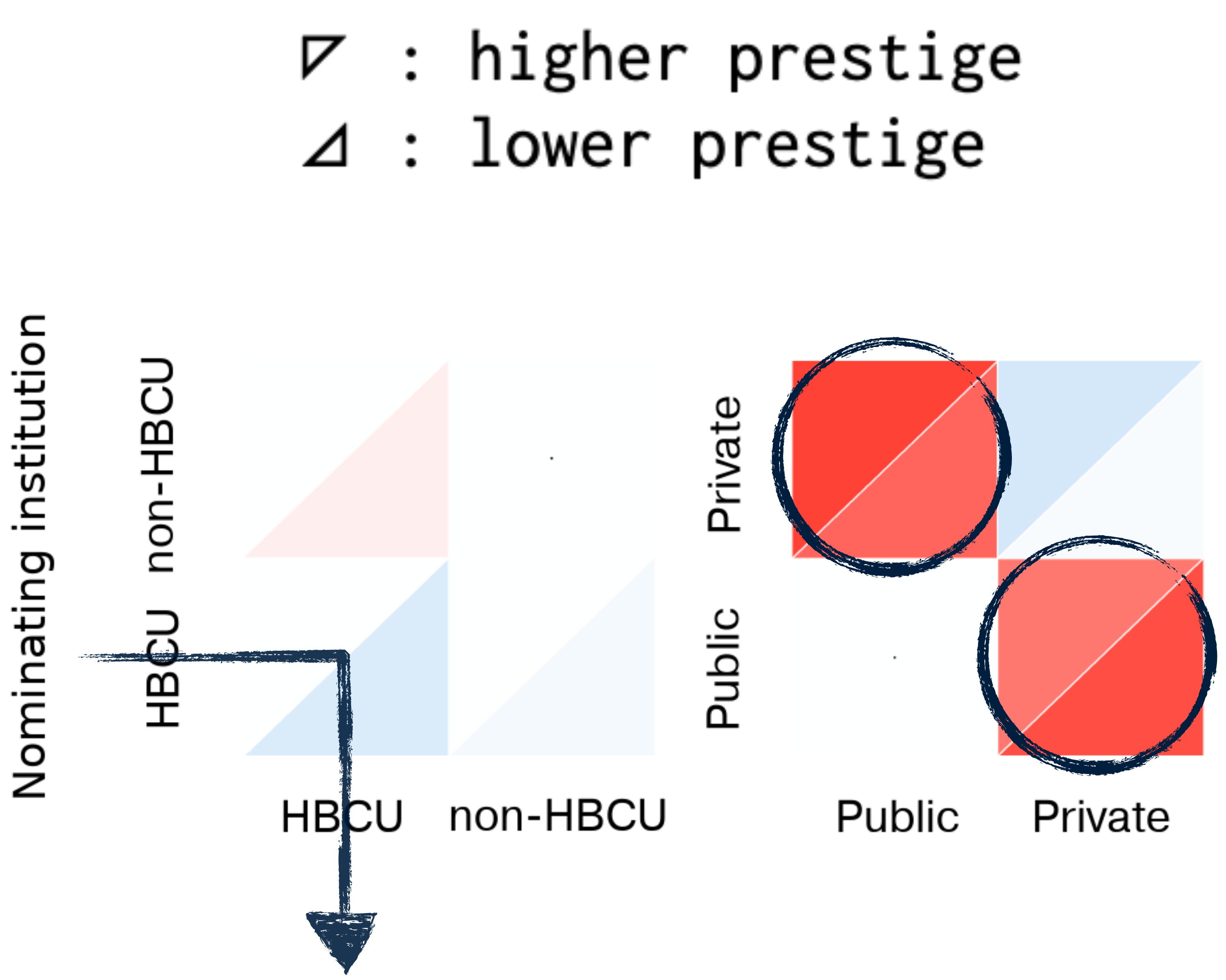
HBCU tends (slightly) to choose other HBCUs.

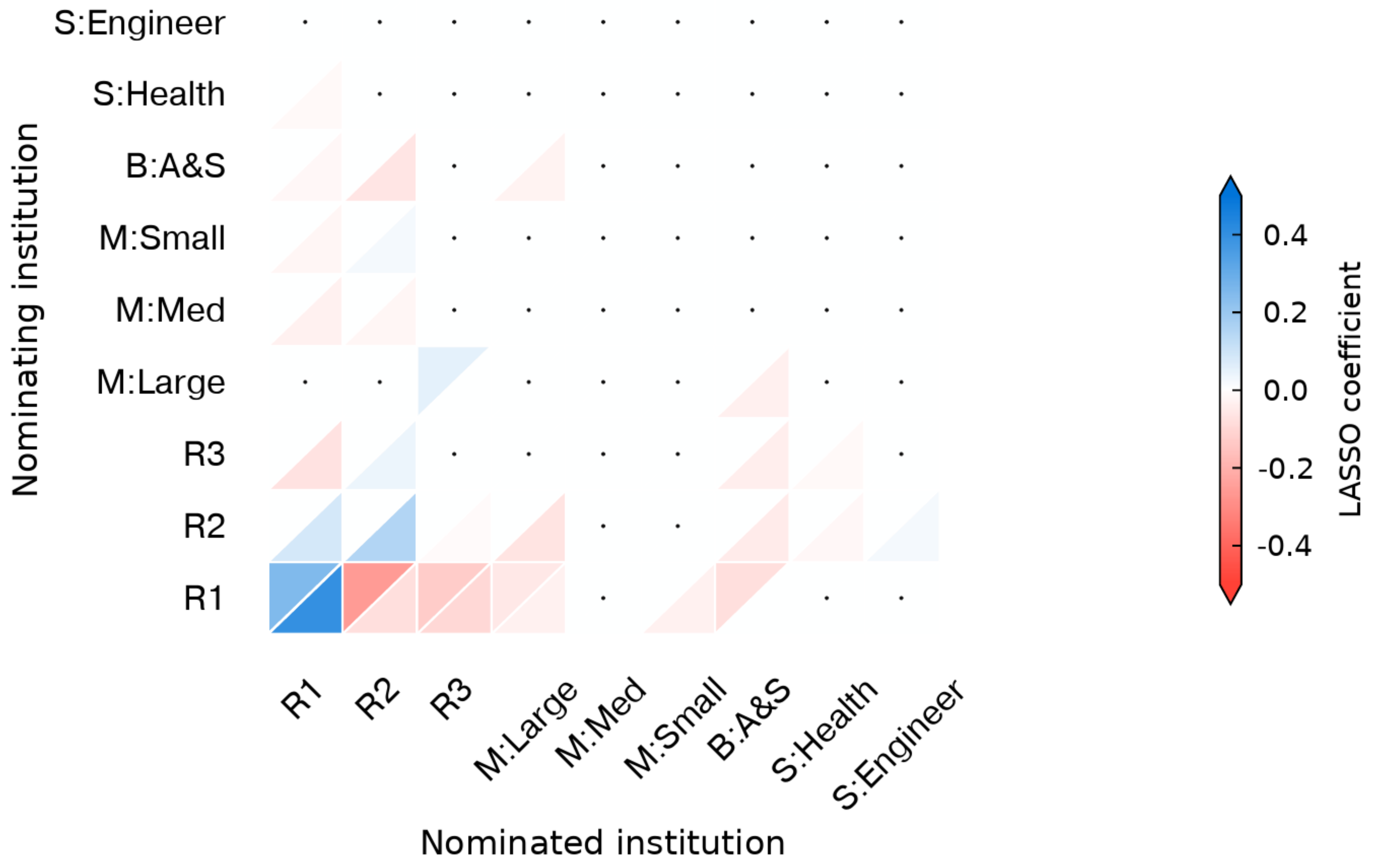


*Model validated via the average negative log likelihood (ANLL): stratified model (0.863); separate model (0.864); common model (0.912)

Rich peer decision patterns (categorical features)





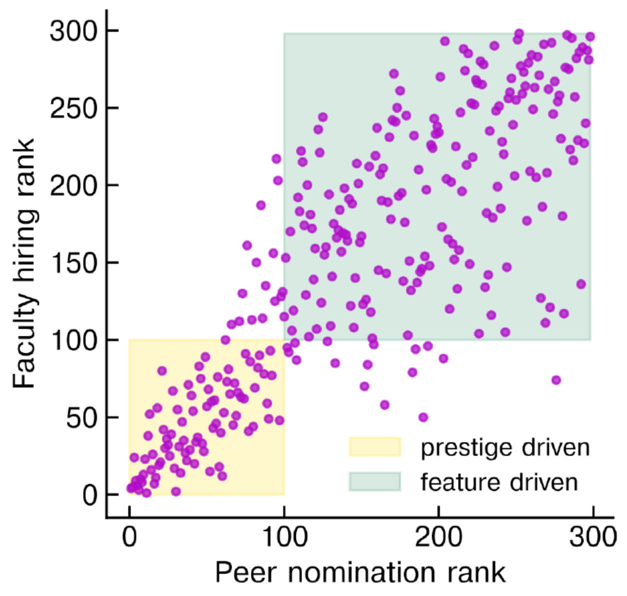


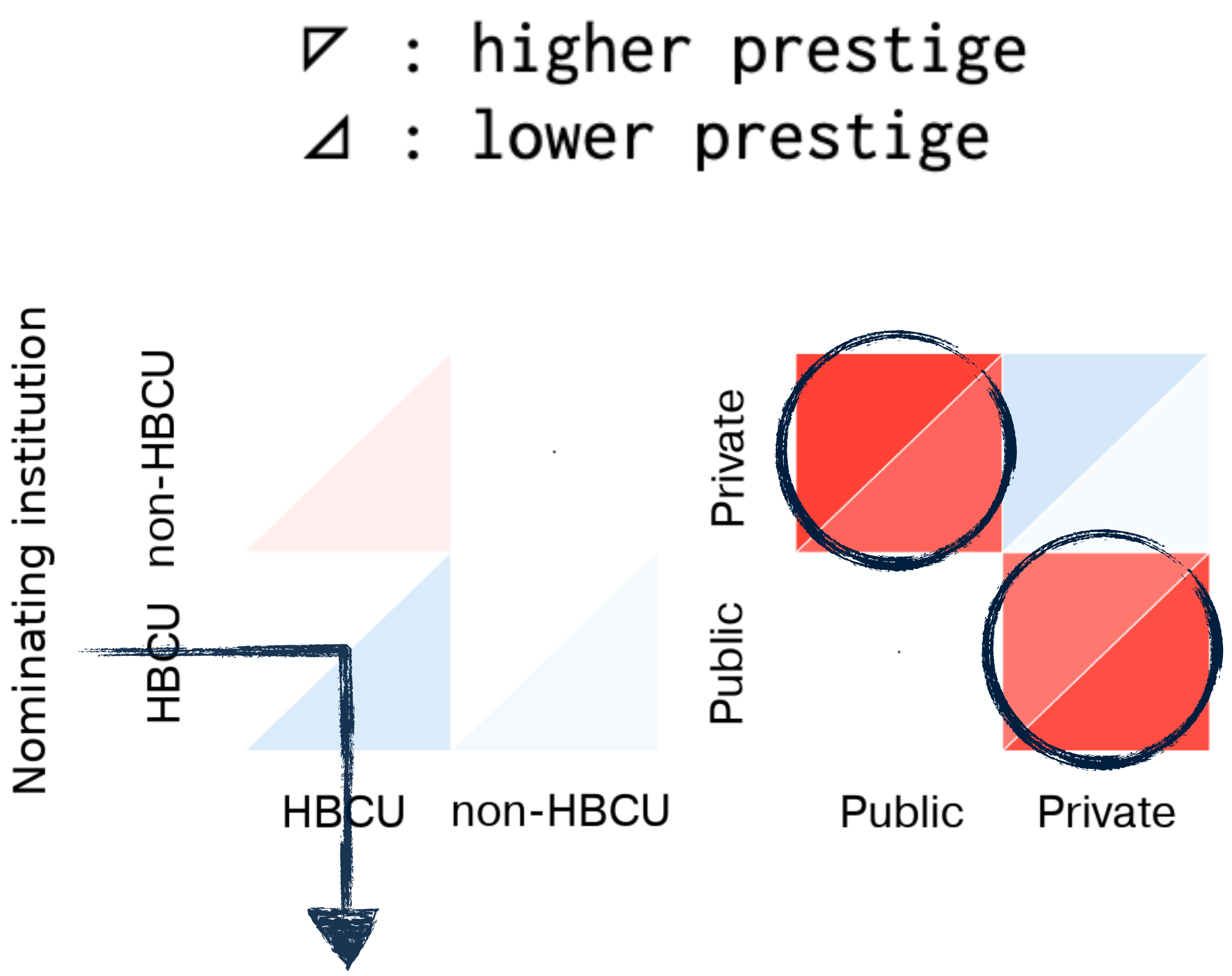
HBCU tends (slightly) to choose other HBCUs.

Private schools don't choose public schools (vv).

*Model validated via the average negative log likelihood (ANLL): stratified model (0.863); separate model (0.864); common model (0.912)

Rich peer decision patterns (categorical features)

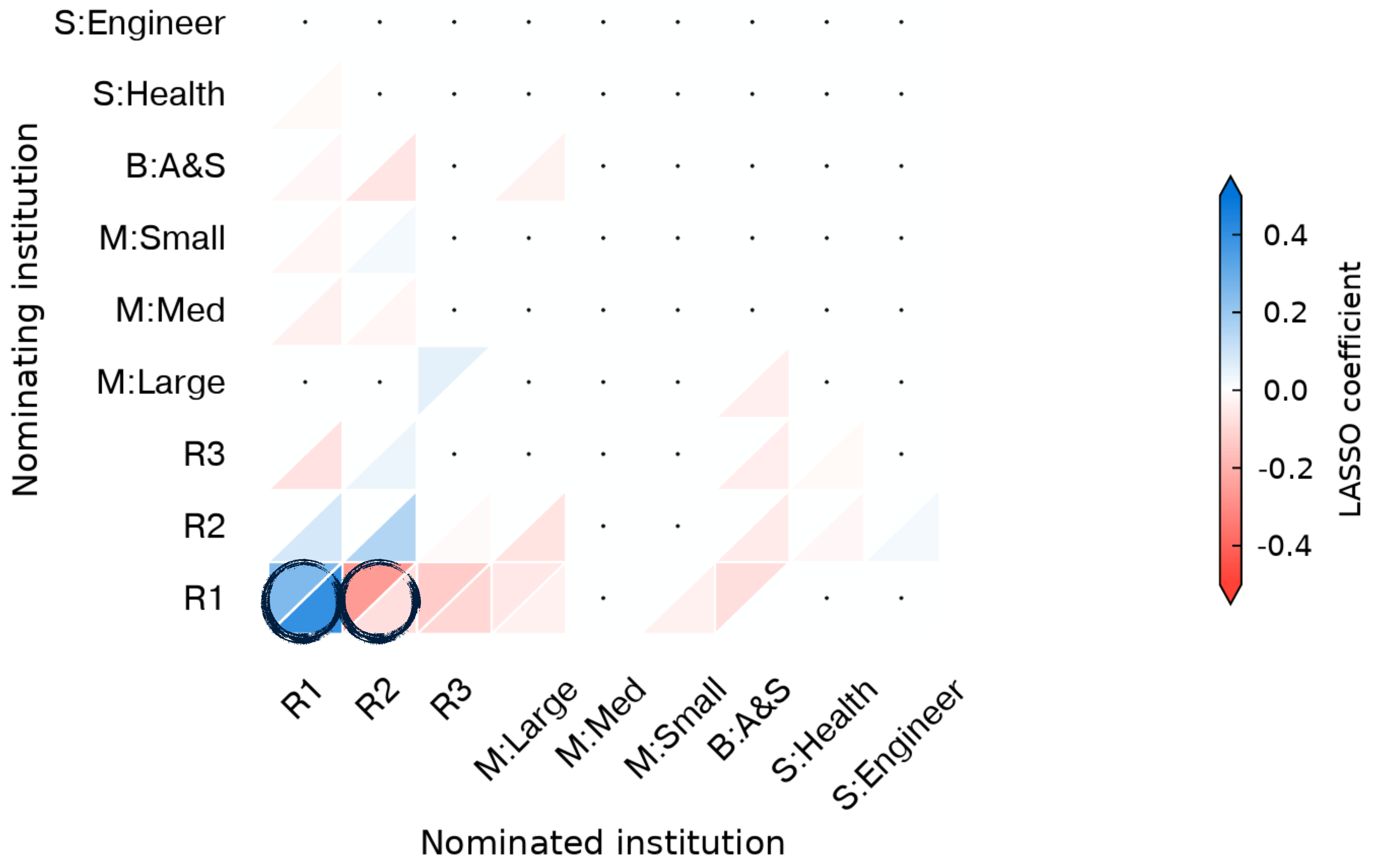




HBCU tends (slightly) to choose other HBCUs.

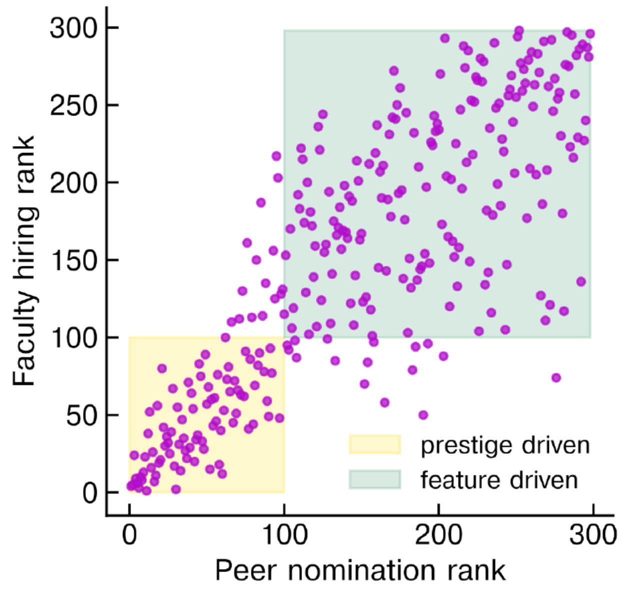
Private schools don't choose public schools (vv).

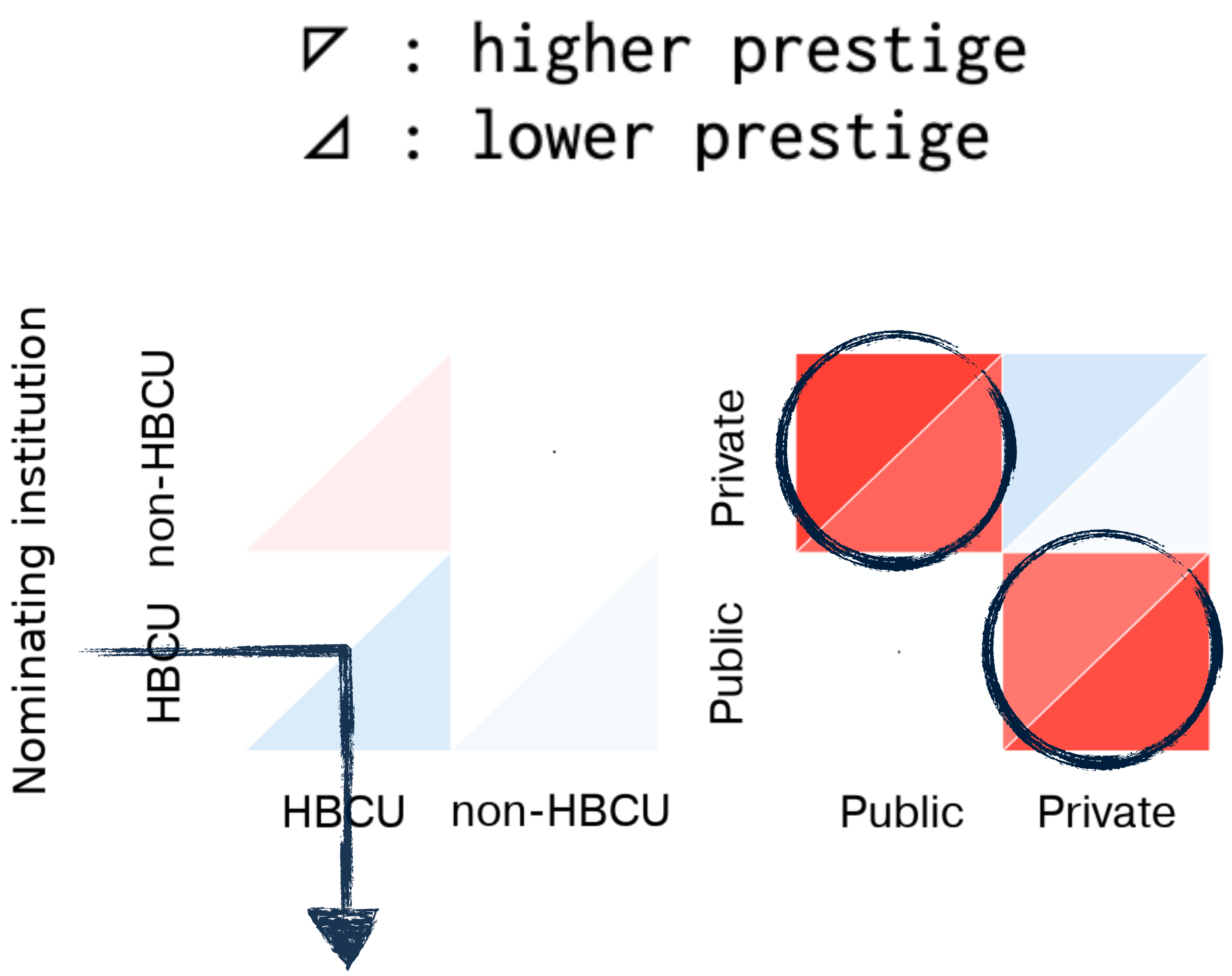
R1 chooses other R1, but **not** R2.

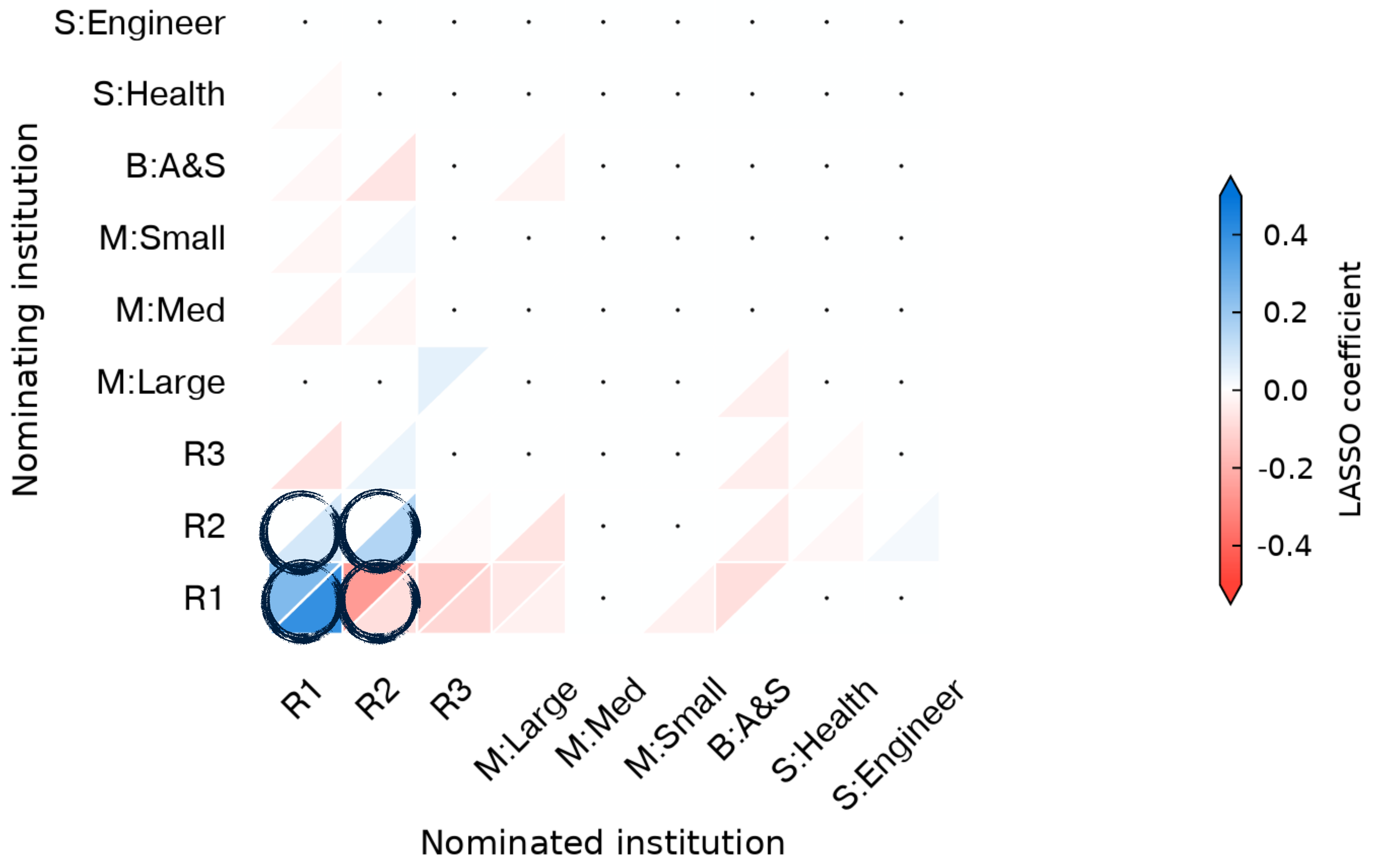


*Model validated via the average negative log likelihood (ANLL): stratified model (0.863); separate model (0.864); common model (0.912)

Rich peer decision patterns (categorical features)







HBCU tends (slightly) to choose other HBCUs.

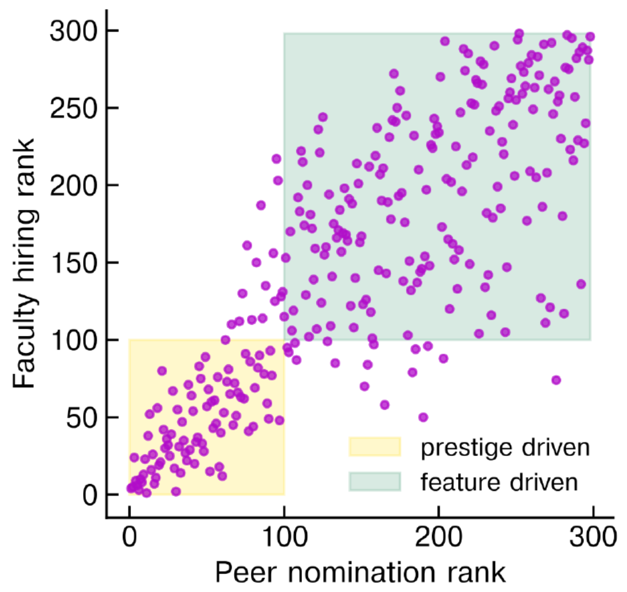
Private schools don't choose public schools (vv).

R1 chooses other R1, but **not** R2.

R2 chooses other R2, but also chooses R1.

*Model validated via the average negative log likelihood (ANLL): stratified model (0.863); separate model (0.864); common model (0.912)

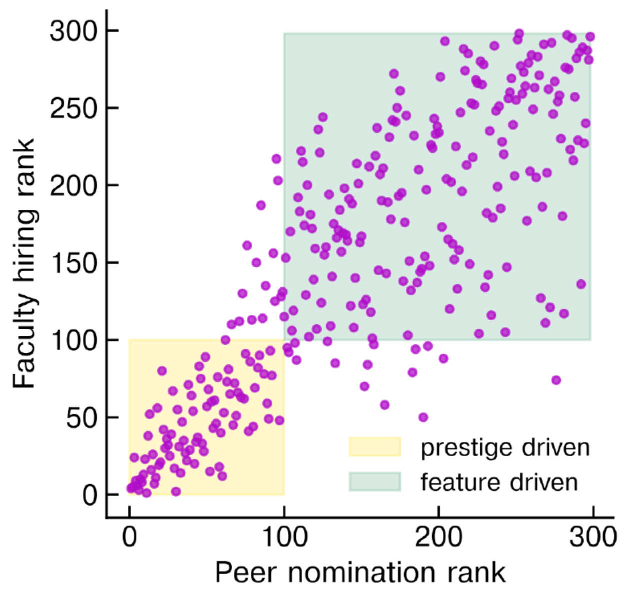
Prestige determines peer decision (numeric features)



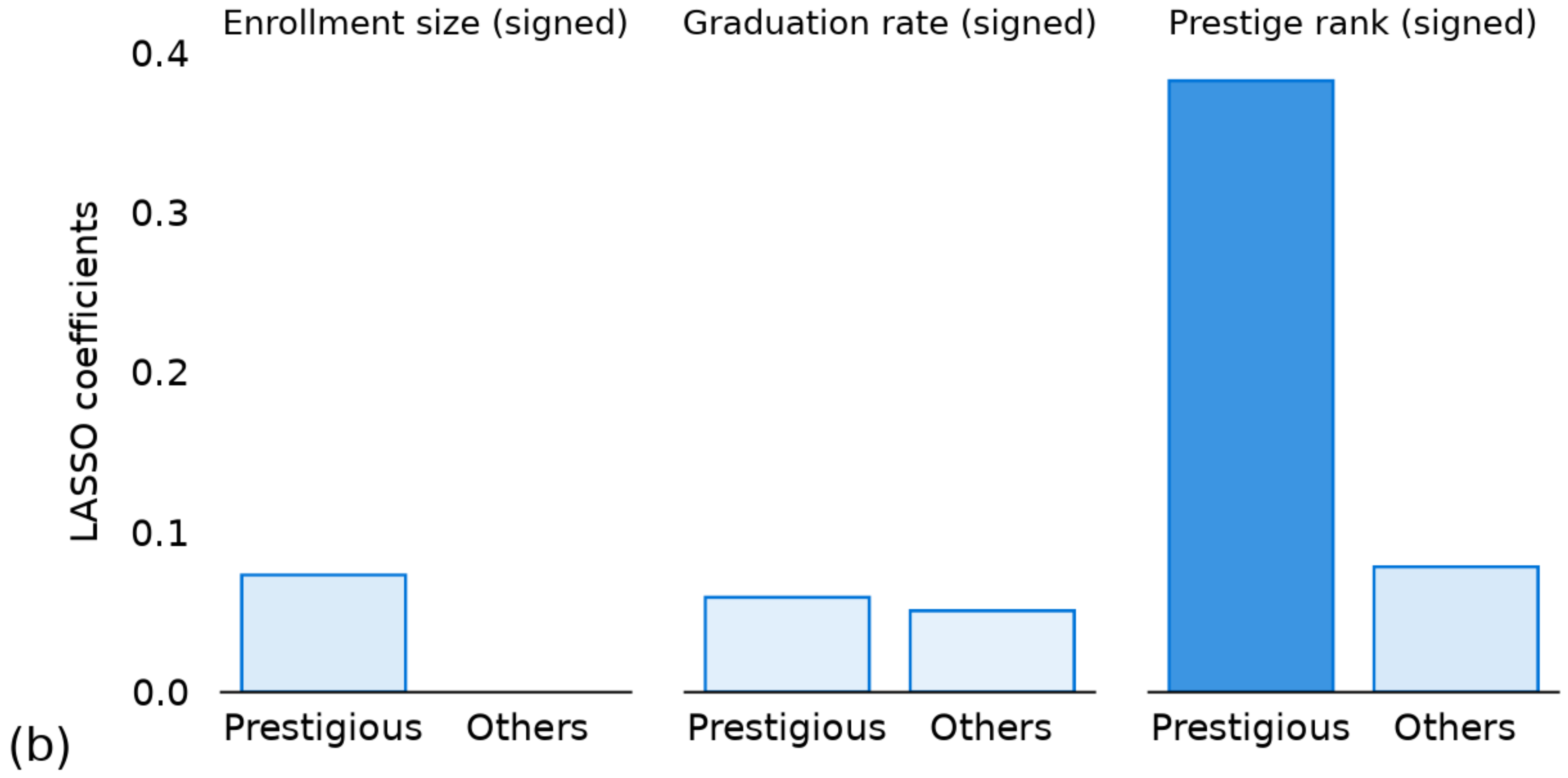
$i \to j$	Adm	Size	Grad	Prestige
Similarity (unsigned) $(x_j - x_i)^2$	∅	∅	∅	∅
Difference (signed) $x_j - x_i$	∅	✓	✓	✓

*Model validated via the average negative log likelihood (ANLL): stratified model (0.863); separate model (0.864); common model (0.912)

Prestige determines peer decision (numeric features)

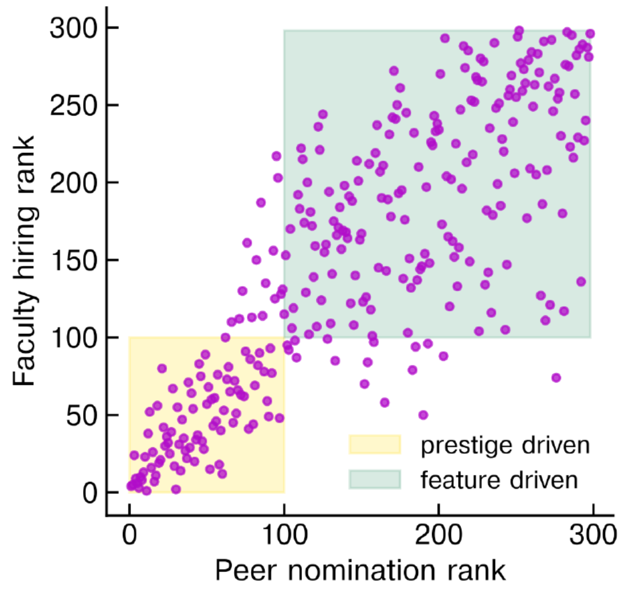


$i \to j$	Adm	Size	Grad	Prestige
Similarity (unsigned) $(x_j - x_i)^2$	∅	∅	∅	∅
Difference (signed) $x_j - x_i$	∅	✓	✓	✓

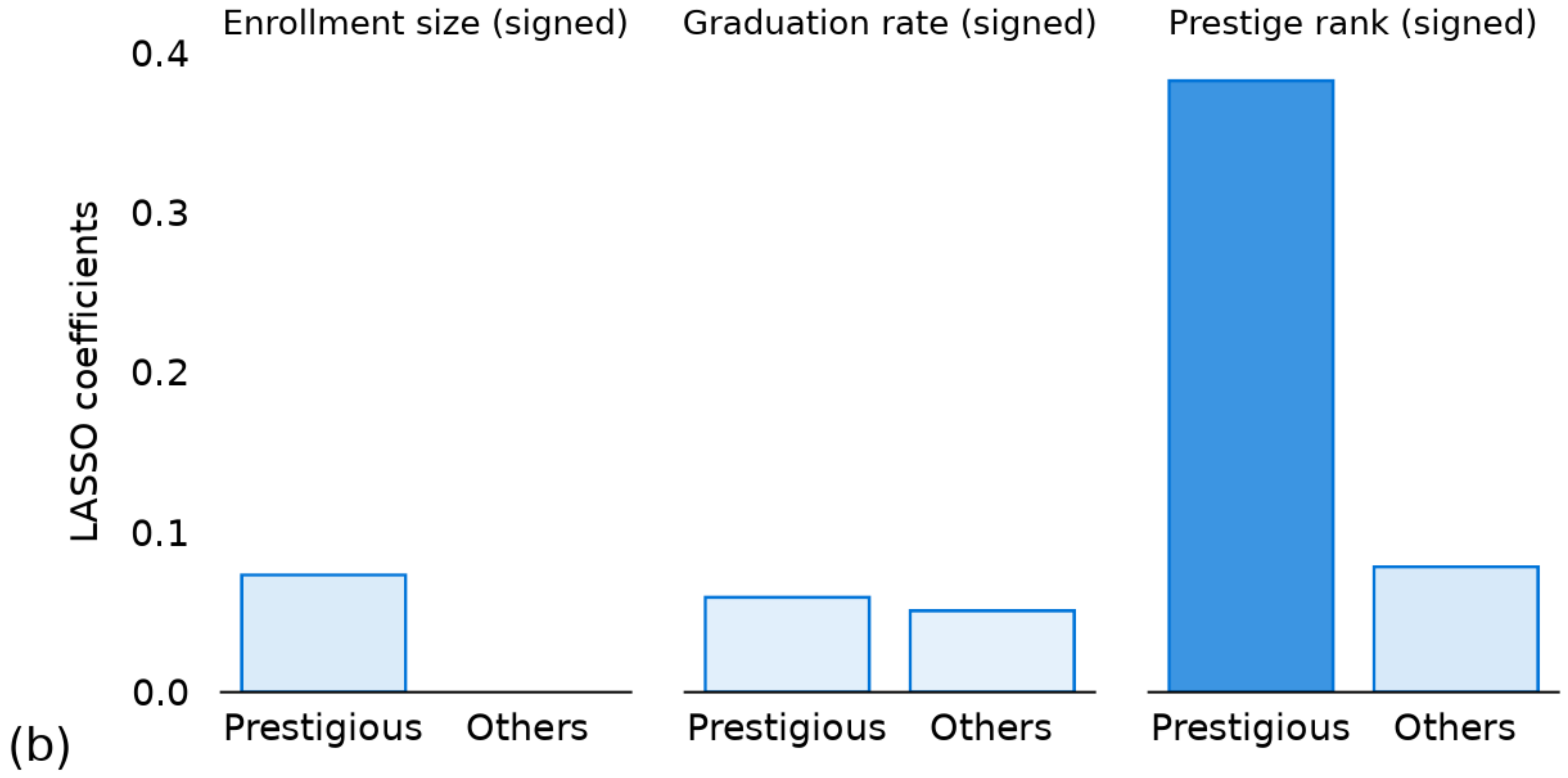


*Model validated via the average negative log likelihood (ANLL): stratified model (0.863); separate model (0.864); common model (0.912)

Prestige determines peer decision (numeric features)



$i \to j$	Adm	Size	Grad	Prestige
Similarity (unsigned) $(x_j - x_i)^2$	∅	∅	∅	∅
Difference (signed) $x_j - x_i$	∅	✓	✓	✓

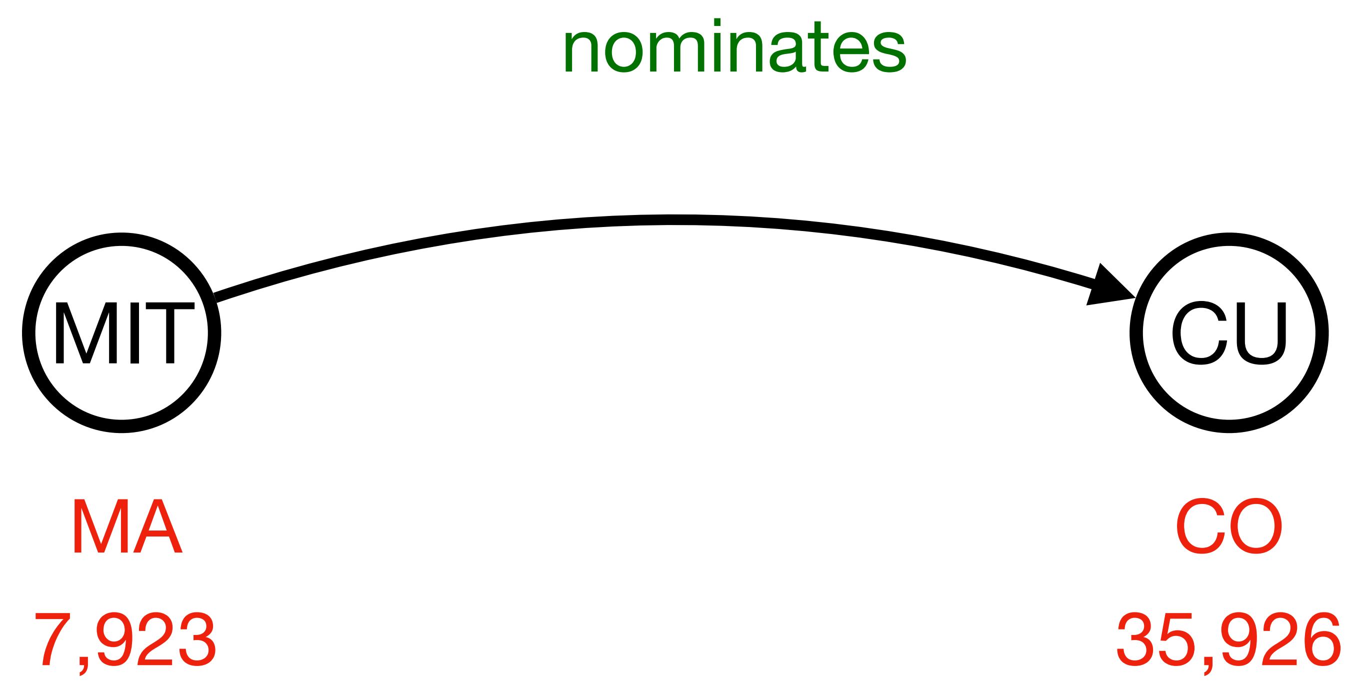


Outcome: Aspiration for a higher prestige rank is a significant predictive variable (for the elite institutions)

*Model validated via the average negative log likelihood (ANLL): stratified model (0.863); separate model (0.864); common model (0.912)

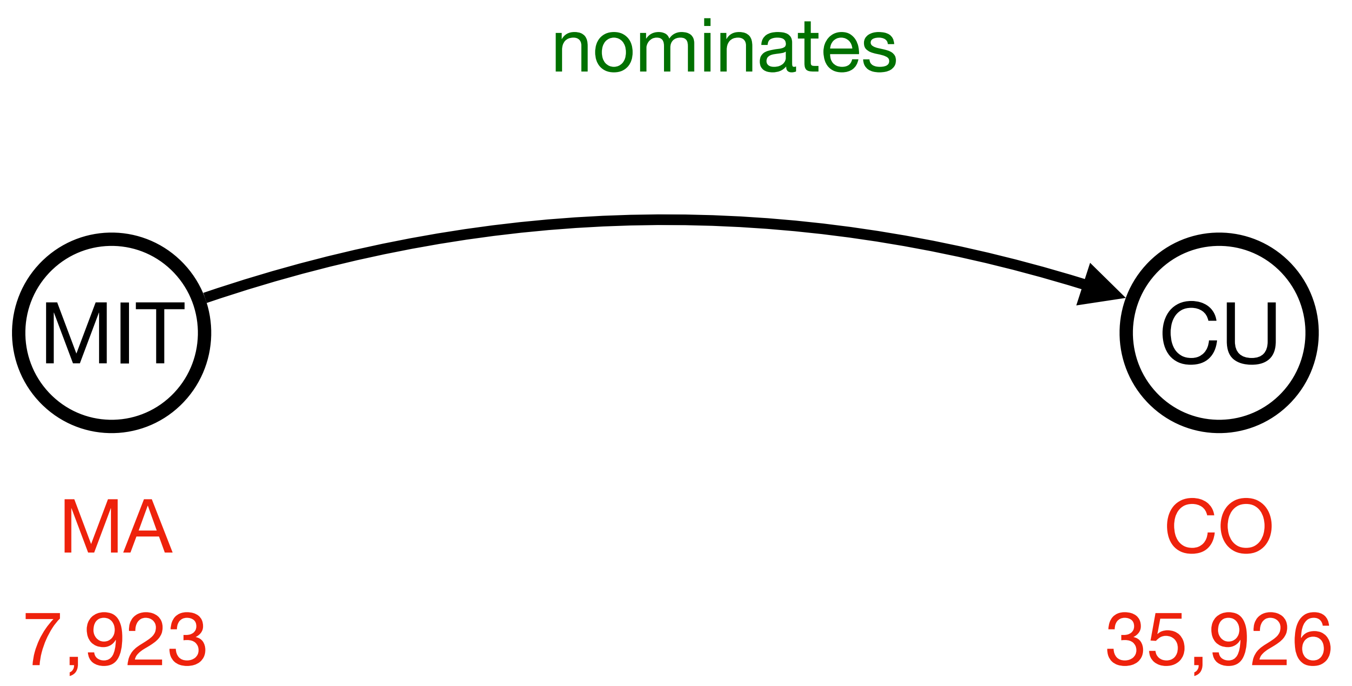
Wait… What about *other* schools?

Nomination network



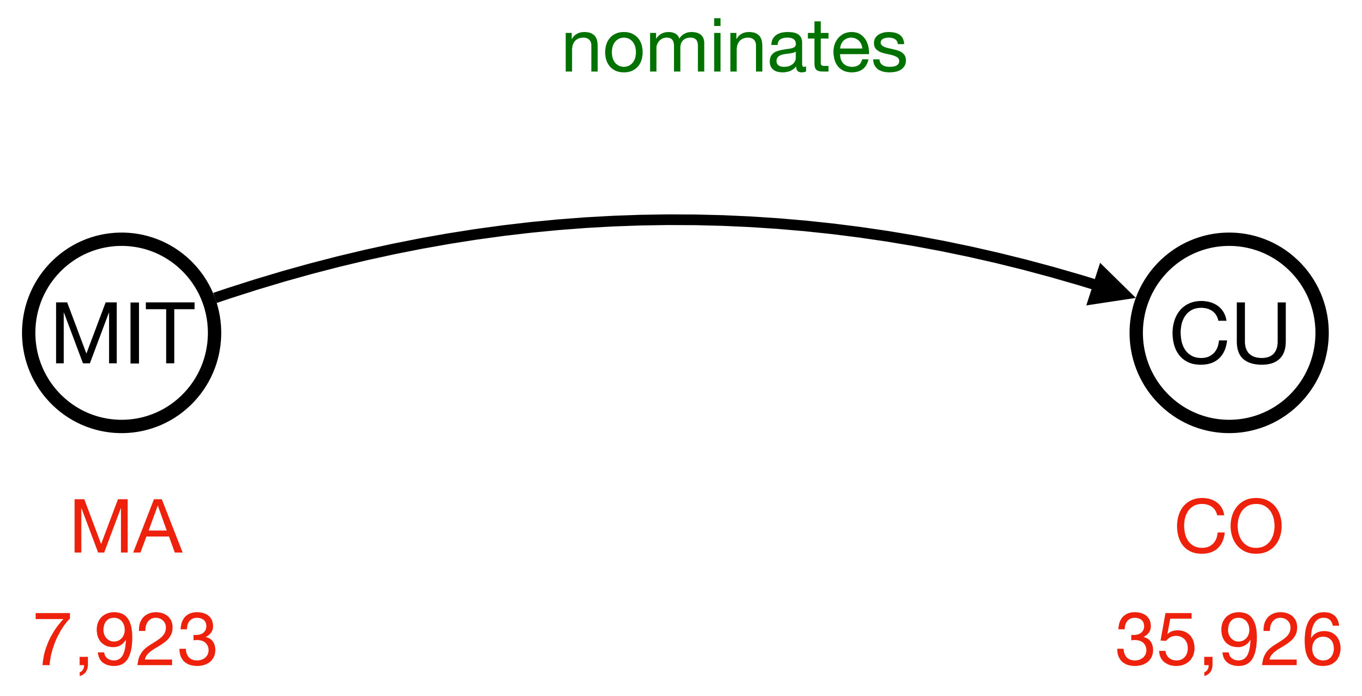
Wait… What about *other* schools?

Nomination network

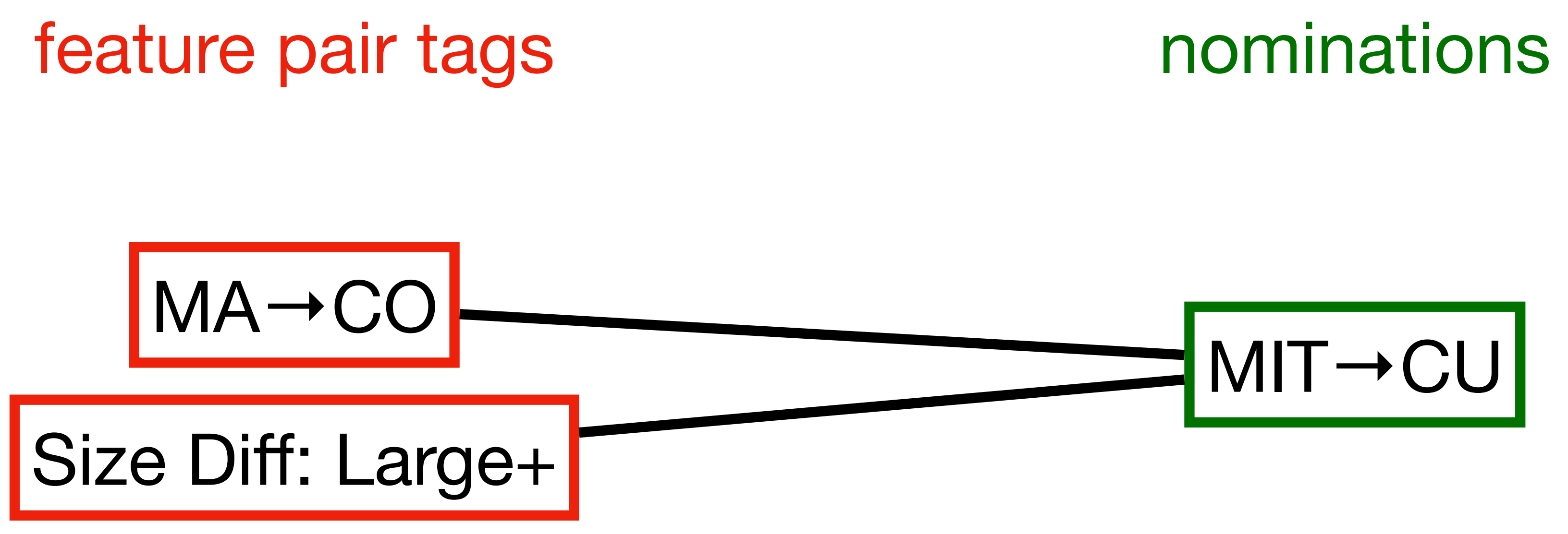


Wait… What about *other* schools?

Nomination network

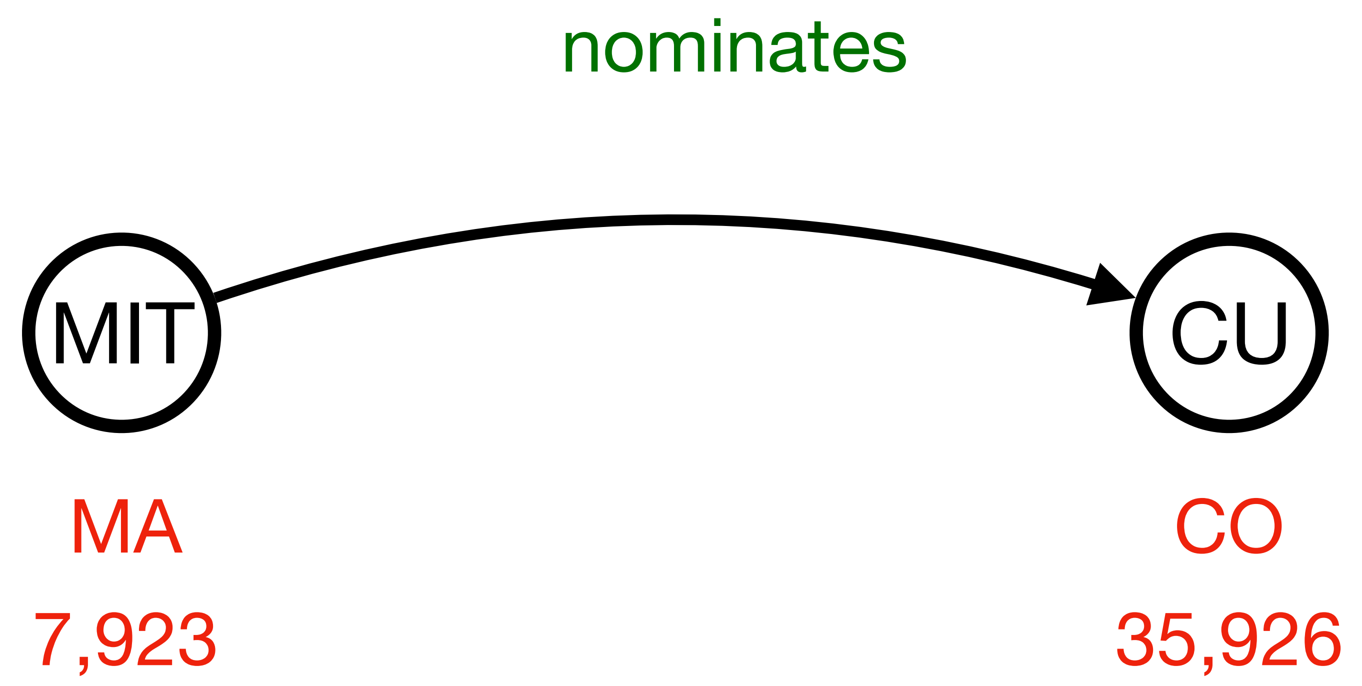


Bipartite expanded feature network

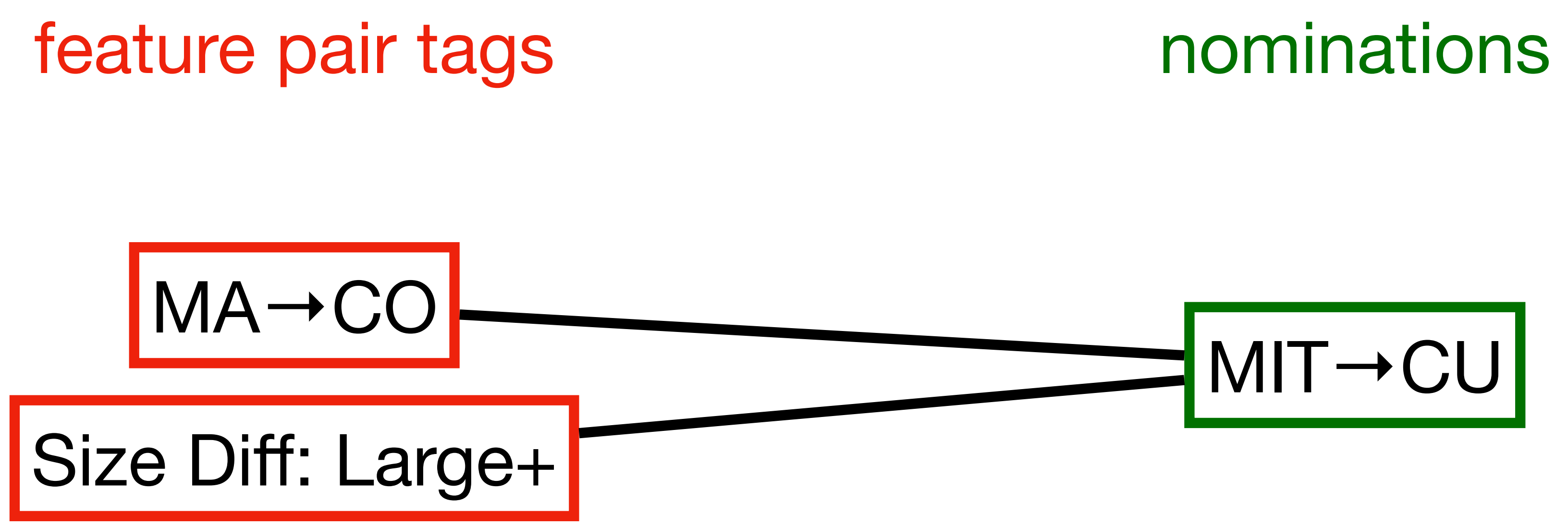


Wait… What about *other* schools?

Nomination network



Bipartite expanded feature network



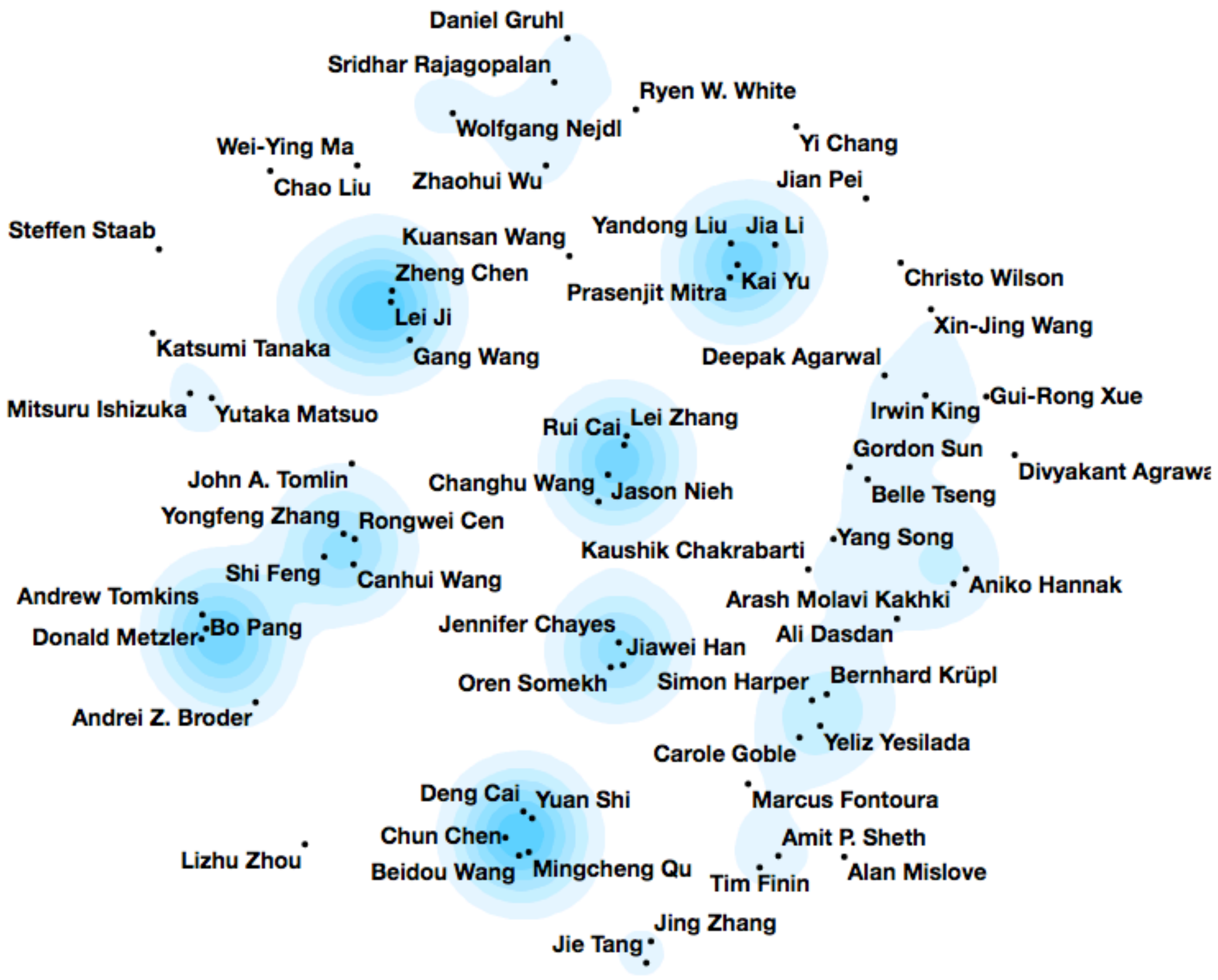
Method (clustering via the bipartite stochastic block model)

Conventional.

Bipartite network. of conferences and authors.

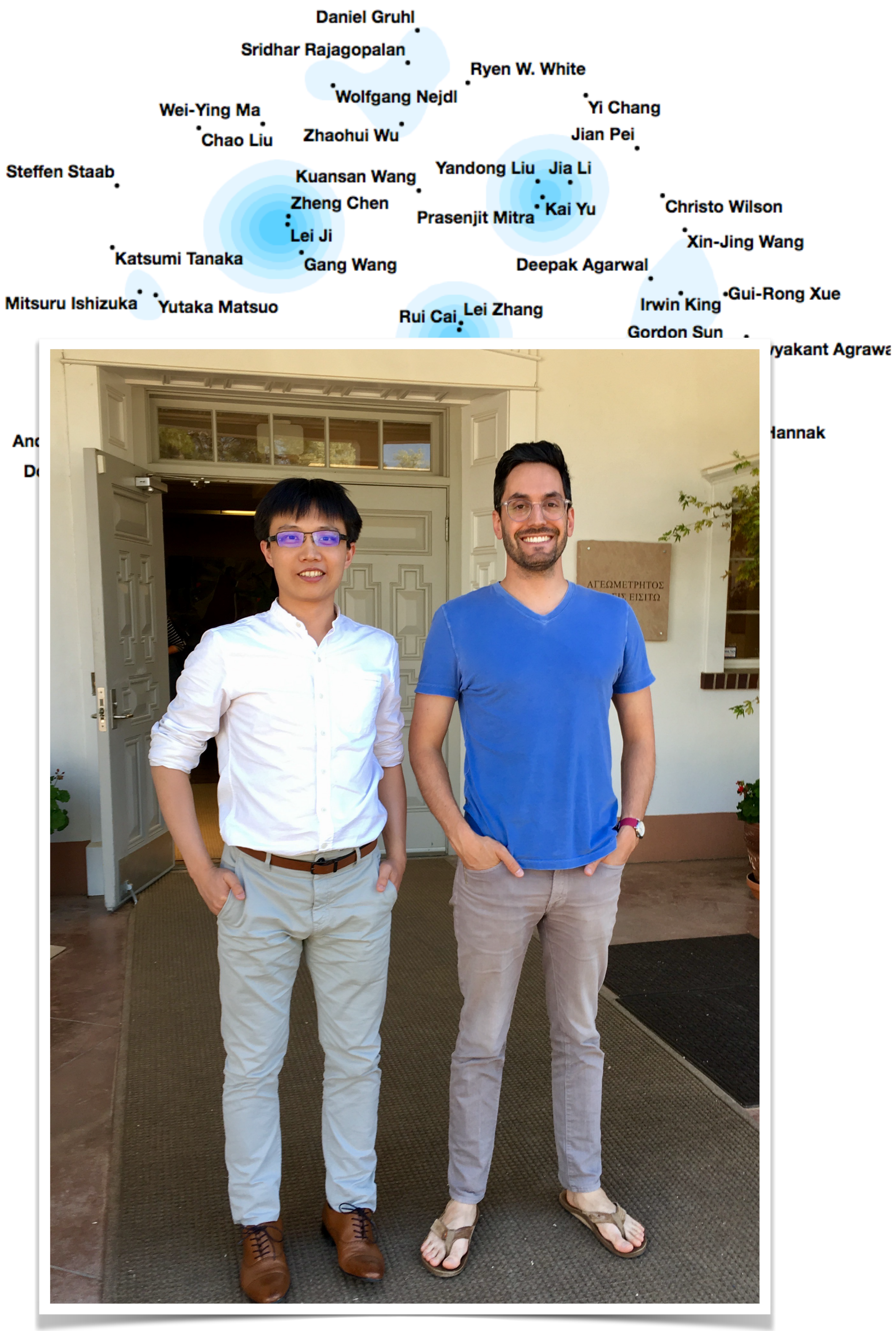
Nonnegative matrix factorization. Clustering who attend what conference.

Problems. *How many groups to choose? Sometimes just outputs 1 group (e.g., scarce data)* 😭



Lin *et al.*, SIAM Conference on Data Mining (2010)

Method (clustering via the bipartite stochastic block model)



Conventional.

Bipartite network. of conferences and authors.

Nonnegative matrix factorization. Clustering who attend what conference.

Problems. *How many groups to choose? Sometimes just outputs 1 group (e.g., scarce data)* 😭

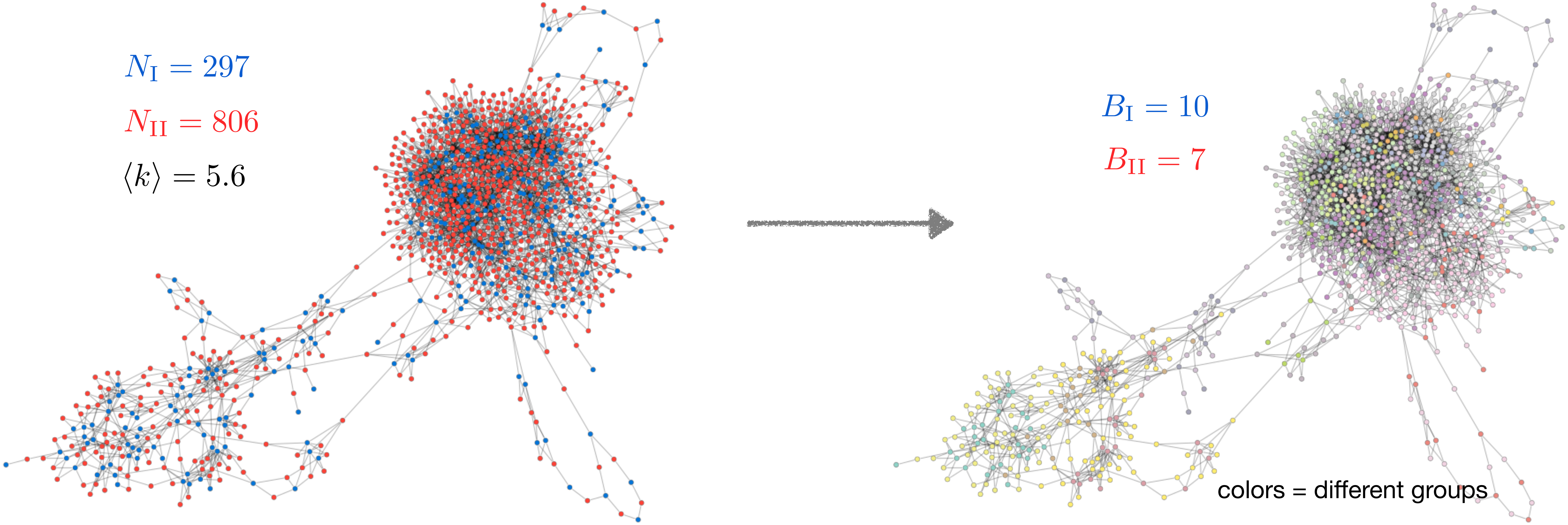
Novel.

Nonparametric Bayesian modeling. Probabilistic generative model that detects communities. 😎

Communities with statistical validation (e.g., description length, posterior likelihood)

Known: The bipartite structure, i.e., which node is of which type. Edges are given (measurement is perfect).

Unknown: Of which group should a node be assigned to.



SBMs formalize the concept of "large-scale structure"

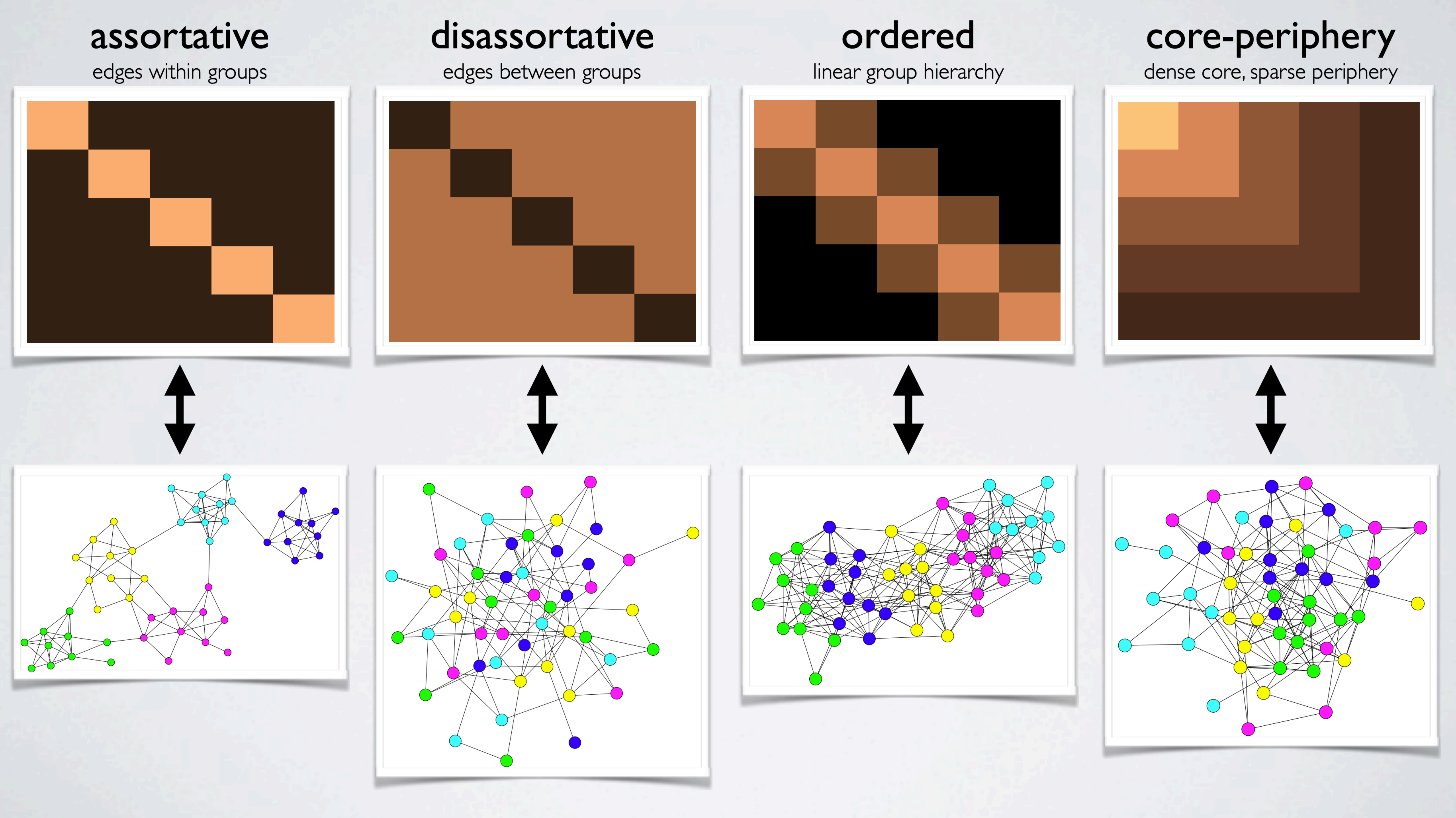


Image credit: Aaron Clauset

Bipartite stochastic block model

$$P(\boldsymbol{A},\boldsymbol{k},\boldsymbol{e},\boldsymbol{b}) = \underbrace{P(\boldsymbol{A} \mid \boldsymbol{k},\boldsymbol{e},\boldsymbol{b})}_{\text{Likelihood}}\underbrace{P(\boldsymbol{k} \mid \boldsymbol{e},\boldsymbol{b})P(\boldsymbol{e} \mid \boldsymbol{b})P(\boldsymbol{b})}_{\text{Prior}}$$

Bipartite stochastic block model

$$P_{\text{bi}}(\boldsymbol{A},\boldsymbol{k},\boldsymbol{e},\boldsymbol{b}) = \frac{\prod_i k_i! \prod_{r<s} e_{rs}!}{\prod_r e_r! \prod_{i<j} A_{ij}!} \prod_r \frac{\prod_k \eta_k^r!}{n_r!} \frac{1}{q(e_r,n_r)} \left(\!\!\binom{B_{\text{I}}B_{\text{II}}}{E}\!\!\right)^{-1} \frac{\prod_r n_r!}{N_{\text{I}}!N_{\text{II}}!} \binom{N_{\text{I}}-1}{B_{\text{I}}-1}^{-1} \binom{N_{\text{II}}-1}{B_{\text{II}}-1}^{-1} \frac{1}{N_{\text{I}}N_{\text{II}}}.$$

$$P(\boldsymbol{A},\boldsymbol{k},\boldsymbol{e},\boldsymbol{b}) = \underbrace{P(\boldsymbol{A} \mid \boldsymbol{k},\boldsymbol{e},\boldsymbol{b})}_{\text{Likelihood}} \underbrace{P(\boldsymbol{k} \mid \boldsymbol{e},\boldsymbol{b}) P(\boldsymbol{e} \mid \boldsymbol{b}) P(\boldsymbol{b})}_{\text{Prior}}$$

Bipartite stochastic block model

$$P_{\mathrm{bi}}(\boldsymbol{A},\boldsymbol{k},\boldsymbol{e},\boldsymbol{b}) = \frac{\prod_i k_i! \prod_{r<s} e_{rs}!}{\prod_r e_r! \prod_{i<j} A_{ij}!} \prod_r \frac{\prod_k \eta_k^r!}{n_r!} \frac{1}{q(e_r, n_r)} \left(\!\!\binom{B_{\mathrm{I}} B_{\mathrm{II}}}{E}\!\!\right)^{-1} \frac{\prod_r n_r!}{N_{\mathrm{I}}! N_{\mathrm{II}}!} \binom{N_{\mathrm{I}}-1}{B_{\mathrm{I}}-1}^{-1} \binom{N_{\mathrm{II}}-1}{B_{\mathrm{II}}-1}^{-1} \frac{1}{N_{\mathrm{I}} N_{\mathrm{II}}}.$$

$$P(\boldsymbol{A},\boldsymbol{k},\boldsymbol{e},\boldsymbol{b}) = \underbrace{P(\boldsymbol{A} \mid \boldsymbol{k},\boldsymbol{e},\boldsymbol{b})}_{\text{Likelihood}} \underbrace{P(\boldsymbol{k} \mid \boldsymbol{e},\boldsymbol{b}) P(\boldsymbol{e} \mid \boldsymbol{b}) P(\boldsymbol{b})}_{\text{Prior}} = P(\mathbf{A}, \{b_i\})$$

Optimized with Markov chain Monte Carlo (Metropolis-Hastings)

Bipartite stochastic block model

$$P_{\mathrm{bi}}(\boldsymbol{A},\boldsymbol{k},\boldsymbol{e},\boldsymbol{b}) = \frac{\prod_i k_i! \prod_{r<s} e_{rs}!}{\prod_r e_r! \prod_{i<j} A_{ij}!} \prod_r \frac{\prod_k \eta_k^r!}{n_r!} \frac{1}{q(e_r,n_r)} \left(\!\!\binom{B_{\mathrm{I}}B_{\mathrm{II}}}{E}\!\!\right)^{-1} \frac{\prod_r n_r!}{N_{\mathrm{I}}!N_{\mathrm{II}}!} \binom{N_{\mathrm{I}}-1}{B_{\mathrm{I}}-1}^{-1} \binom{N_{\mathrm{II}}-1}{B_{\mathrm{II}}-1}^{-1} \frac{1}{N_{\mathrm{I}}N_{\mathrm{II}}}.$$

$$P(\boldsymbol{A},\boldsymbol{k},\boldsymbol{e},\boldsymbol{b}) = \underbrace{P(\boldsymbol{A} \mid \boldsymbol{k},\boldsymbol{e},\boldsymbol{b})}_{\text{Likelihood}} \underbrace{P(\boldsymbol{k} \mid \boldsymbol{e},\boldsymbol{b}) P(\boldsymbol{e} \mid \boldsymbol{b}) P(\boldsymbol{b})}_{\text{Prior}} = P(\mathbf{A}, \{b_i\})$$

Optimized with Markov chain Monte Carlo (Metropolis-Hastings)

Enhanced model sensitivity: More detailed prior (e.g., we know *a priori* that the network is bipartite) gives more sensitive model to detect structure.

Bipartite stochastic block model

$$P_{\mathrm{bi}}(\boldsymbol{A},\boldsymbol{k},\boldsymbol{e},\boldsymbol{b}) = \frac{\prod_i k_i! \prod_{r<s} e_{rs}!}{\prod_r e_r! \prod_{i<j} A_{ij}!} \prod_r \frac{\prod_k \eta_k^r!}{n_r!} \frac{1}{q(e_r,n_r)} \left(\!\!\binom{B_{\mathrm{I}} B_{\mathrm{II}}}{E}\!\!\right)^{-1} \frac{\prod_r n_r!}{N_{\mathrm{I}}! N_{\mathrm{II}}!} \binom{N_{\mathrm{I}}-1}{B_{\mathrm{I}}-1}^{-1} \binom{N_{\mathrm{II}}-1}{B_{\mathrm{II}}-1}^{-1} \frac{1}{N_{\mathrm{I}} N_{\mathrm{II}}}.$$

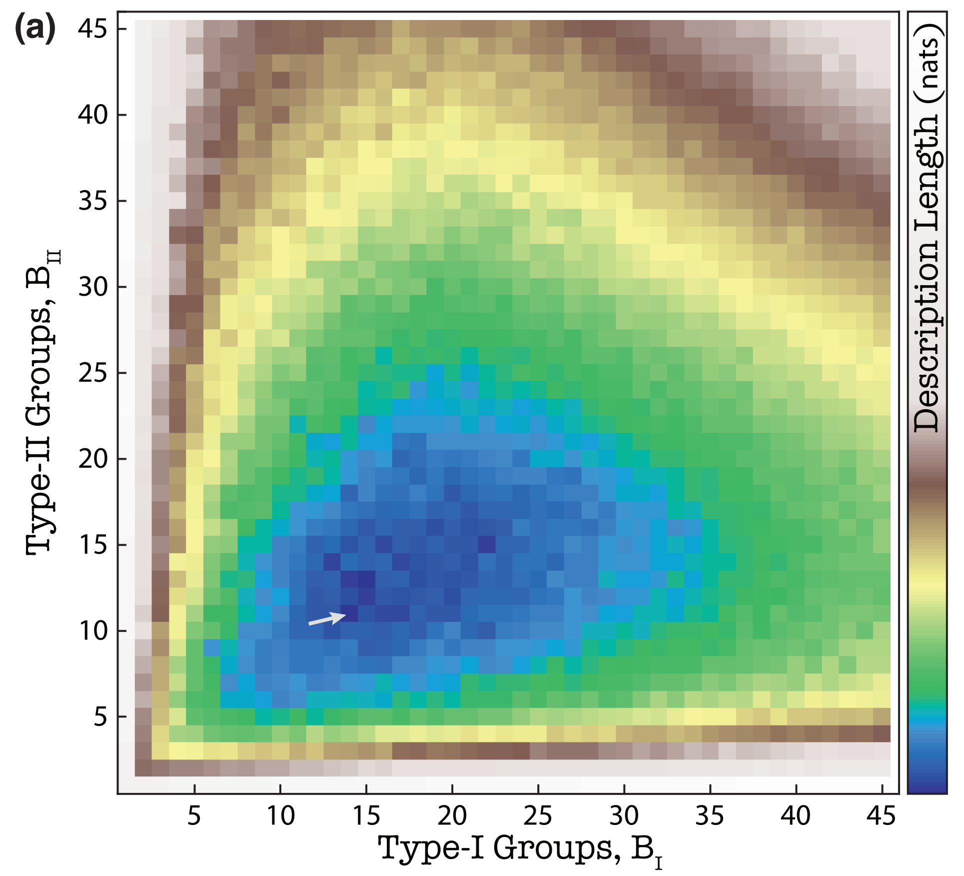
$$P(\boldsymbol{A},\boldsymbol{k},\boldsymbol{e},\boldsymbol{b}) = \underbrace{P(\boldsymbol{A} \mid \boldsymbol{k},\boldsymbol{e},\boldsymbol{b})}_{\text{Likelihood}} \underbrace{P(\boldsymbol{k} \mid \boldsymbol{e},\boldsymbol{b}) P(\boldsymbol{e} \mid \boldsymbol{b}) P(\boldsymbol{b})}_{\text{Prior}} = P(\mathbf{A}, \{b_i\})$$

Optimized with Markov chain Monte Carlo (Metropolis-Hastings)

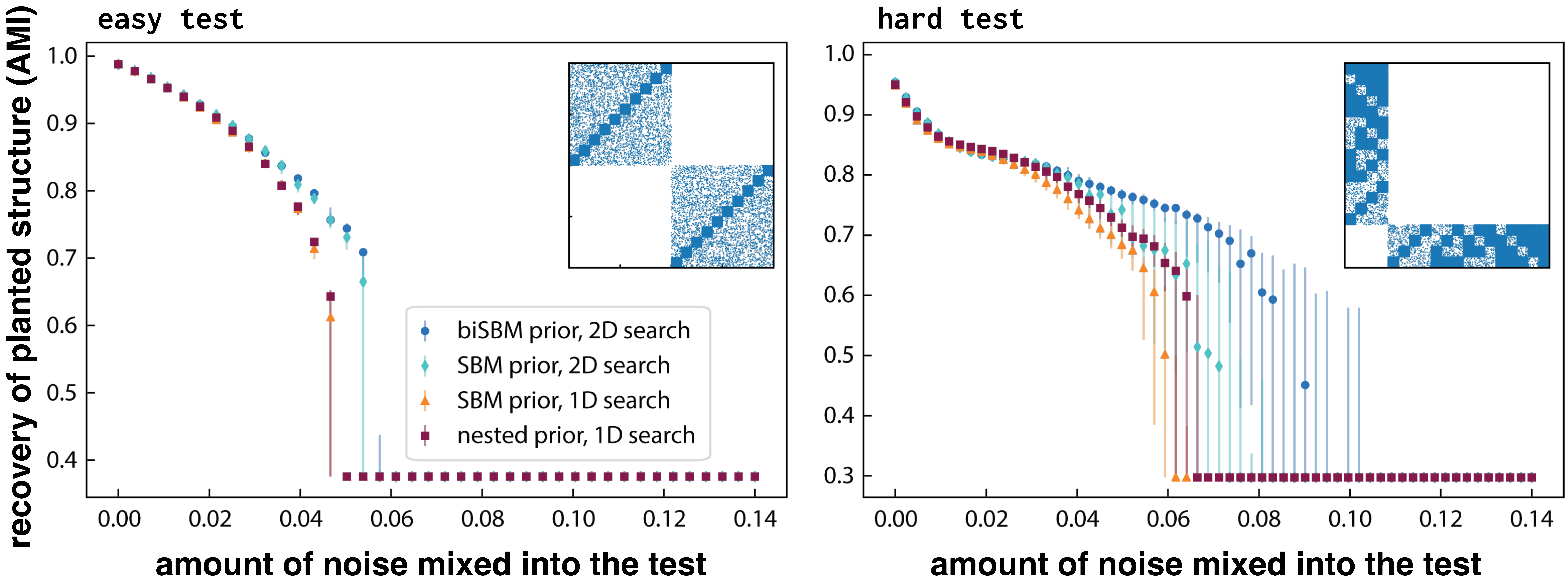
Enhanced model sensitivity: More detailed prior (e.g., we know *a priori* that the network is bipartite) gives more sensitive model to detect structure.

New algorithm comes with the new prior: Prior may introduce new variables, hence a new algorithm is needed.

Use dynamic programming to navigate the landscape



Testing performance: bipartite priors & search matter



effect of bipartite priors

effect of our search algorithm

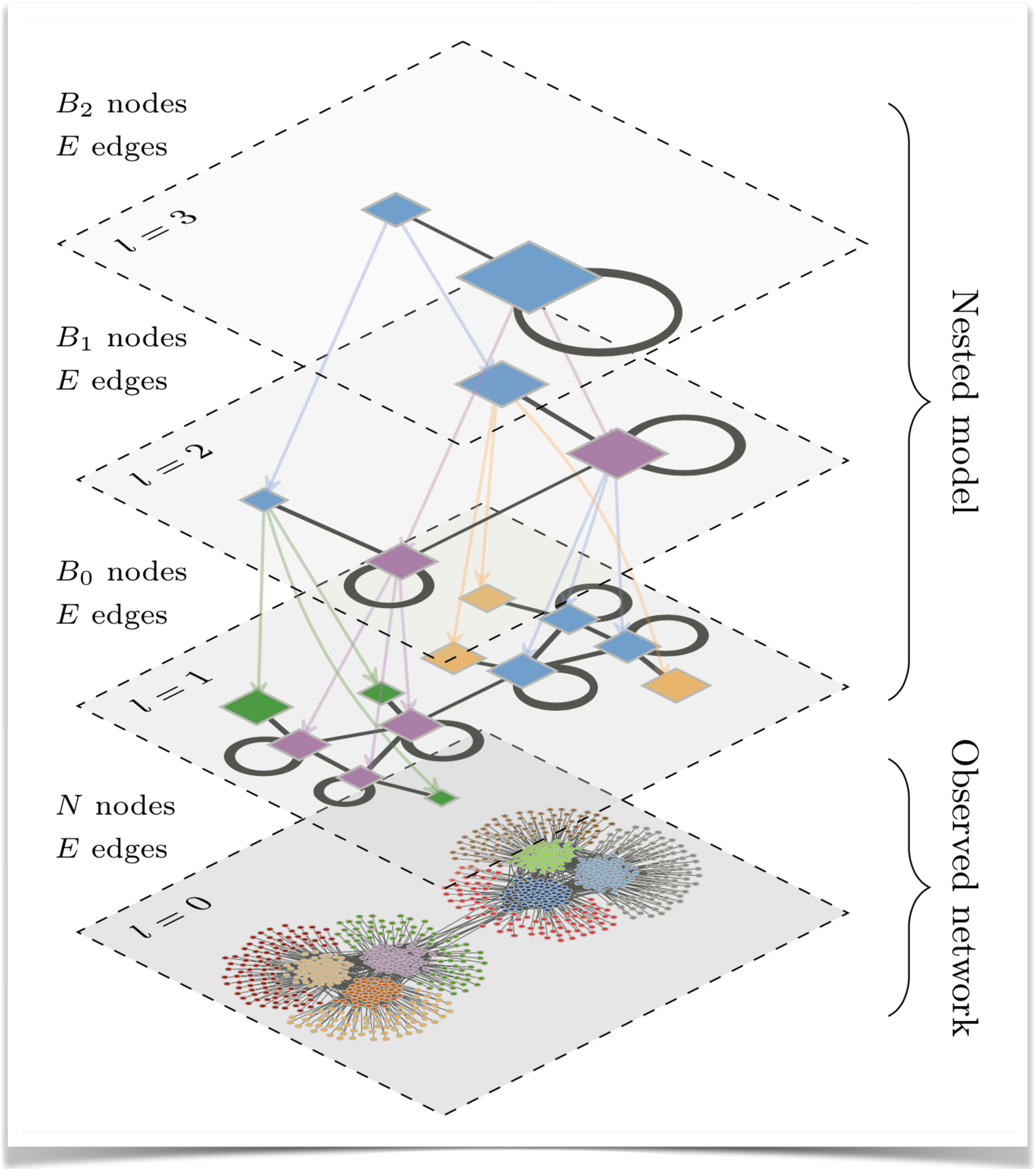
Yen & Larremore, Phys. Rev. E 102, 032309 (2020)

Comparison with hierarchical SBM

Table I. Results for 24 empirical networks. Number of nodes n_{I}, n_{II}, mean degree $\langle k \rangle$, number of type-I groups B_{I}, and number of type-II groups B_{II}, and description length per edge Σ/E. Superscripts: b-biSBM, g-SBM, h-hSBM. L indicates the number of levels found by the hSBM. Reported values indicate best of 100 independent runs. Unless otherwise noted, data are accessible from the Colorado Index of Complex Networks (ICON) [52]. The confidence level is marked with asterisks[a].

Dataset	N_{I}	N_{II}	$\langle k \rangle$	$(B_{\mathrm{I}}^{\mathrm{b}}, B_{\mathrm{II}}^{\mathrm{b}})$	$(B_{\mathrm{I}}^{\mathrm{g}}, B_{\mathrm{II}}^{\mathrm{g}})$	$(B_{\mathrm{I}}^{\mathrm{h}}, B_{\mathrm{II}}^{\mathrm{h}})$	$\langle L+1 \rangle$	Σ^{b}/E	Σ^{h}/E
Southern women interactions [53]	18	14	5.56	(1, 1)	(1, 1)	(1, 1)	2.0	$\mathbf{2.15}^{*}$	2.26
Joern plant-herbivore web [54]	22	52	4.97	(2, 2)	(1, 1)	(1, 1)	2.0	$\mathbf{2.64}^{*}$	2.74
Swingers and parties [55]	57	39	4.83	(1, 1)	(1, 1)	(1, 1)	2.0	$\mathbf{2.92}^{*}$	2.97
McMullen pollination web [56]	54	105	2.57	(2, 2)	(2, 2)	(1, 1)	2.0	$\mathbf{2.87}^{*}$	3.02
Ndrangheta criminals [57]	156	47	4.48	(3, 4)	(3, 3)	(3, 4)	2.87	$\mathbf{3.44}^{*}$	3.49
Abu Sayyaf kidnappings[b] [58]	246	105	2.28	(2, 2)	(1, 1)	(1, 1)	2.0	$\mathbf{4.50}^{*}$	4.54
Virus-host interactome [59]	53	307	2.52	(2, 2)	(1, 1)	(1, 1)	2.0	$\mathbf{3.78}^{*}$	3.81
Clements-Long plant-pollinator [60]	275	96	4.98	(1, 1)	(1, 1)	(1, 1)	2.0	$\mathbf{3.45}^{*}$	3.47
Human musculoskeletal system [61]	173	270	4.30	(7, 8)	(5, 5)	(8, 8)	4.01	$\mathbf{3.94}$	3.94
Mexican drug trafficking[b] [62]	765	10	16.1	(12, 8)	(8, 7)	(10, 6)	3.11	$\mathbf{1.26}^{*}$	1.29
Country-language network [63]	254	614	2.89	(4, 5)	(2, 2)	(4, 3)	2.11	$\mathbf{4.53}^{*}$	4.56
Malaria gene similarity [44]	297	806	5.38	(15, 16)	(6, 6)	(25, 20)	4.95	4.73	$\mathbf{4.67}^{*}$
Protein complex-drug [64]	739	680	5.20	(20, 22)	(14, 14)	(35, 39)	5.06	3.65	$\mathbf{3.50}^{**}$
Robertson plant-pollinator [65]	456	1428	16.2	(20, 18)	(11, 11)	(20, 19)	4.0	$\mathbf{3.10}^{*}$	3.10
Human gene-disease network [66]	1419	516	4.06	(13, 14)	(9, 9)	(35, 36)	5.04	5.02	$\mathbf{4.80}^{**}$
Food ingredients-flavors web [67]	1525	1107	27.9	(27, 69)	(20, 29)	(42, 130)	4.91	2.55	$\mathbf{2.51}^{**}$
Wikipedia doc-word network [68]	63	3140	24.8	(22, 206)	(18, 23)	(29, 71)	4.17	1.58	$\mathbf{1.51}^{**}$
Foursquare check-ins [69]	2060	2876	11.0	(65, 66)	(40, 40)	(244, 248)	5.2	5.92	$\mathbf{5.09}^{**}$
Ancient metabolic network [70]	5651	5252	4.22	(18, 22)	(5, 5)	(17, 21)	4.26	$\mathbf{5.68}^{**}$	5.82
Marvel Universe characters [71]	6486	12942	9.95	(68, 72)	(67, 62)	(365, 314)	6.24	4.70	$\mathbf{4.42}^{***}$
Reuters news stories [72]	19757	38677	33.5	(396, 440)	(87, 108)	(294, 463)	6.25	4.22	$\mathbf{4.16}^{***}$
IMDb movie-actor dataset[c]	53158	39768	6.49	(91, 92)	(69, 68)	(264, 265)	6.22	7.40	$\mathbf{7.30}^{***}$
YouTube group memberships [73]	94238	30087	4.72	(62, 66)	(37, 38)	(221, 238)	5.9	$\mathbf{7.07}^{**}$	7.13
DBpedia writer network [74]	89355	46213	2.13	(22, 26)	(2, 3)	(2, 3)	2.16	$\mathbf{10.32}^{**}$	10.41

[a] Via the posterior odds ratio: $* : \Lambda < 10^{-2}$; $** : \Lambda < 10^{-100}$; $*** : \Lambda < 10^{-10000}$.
[b] Temporal data with timestamps are aggregated, making a multigraph.
[c] Data available at http://www.imdb.com/interfaces. IMDb copyright permits redistribution of data only in unaltered form.



See Peixoto, Phys. Rev. X, 4, 011047 (2014) for the hierarchical SBM.

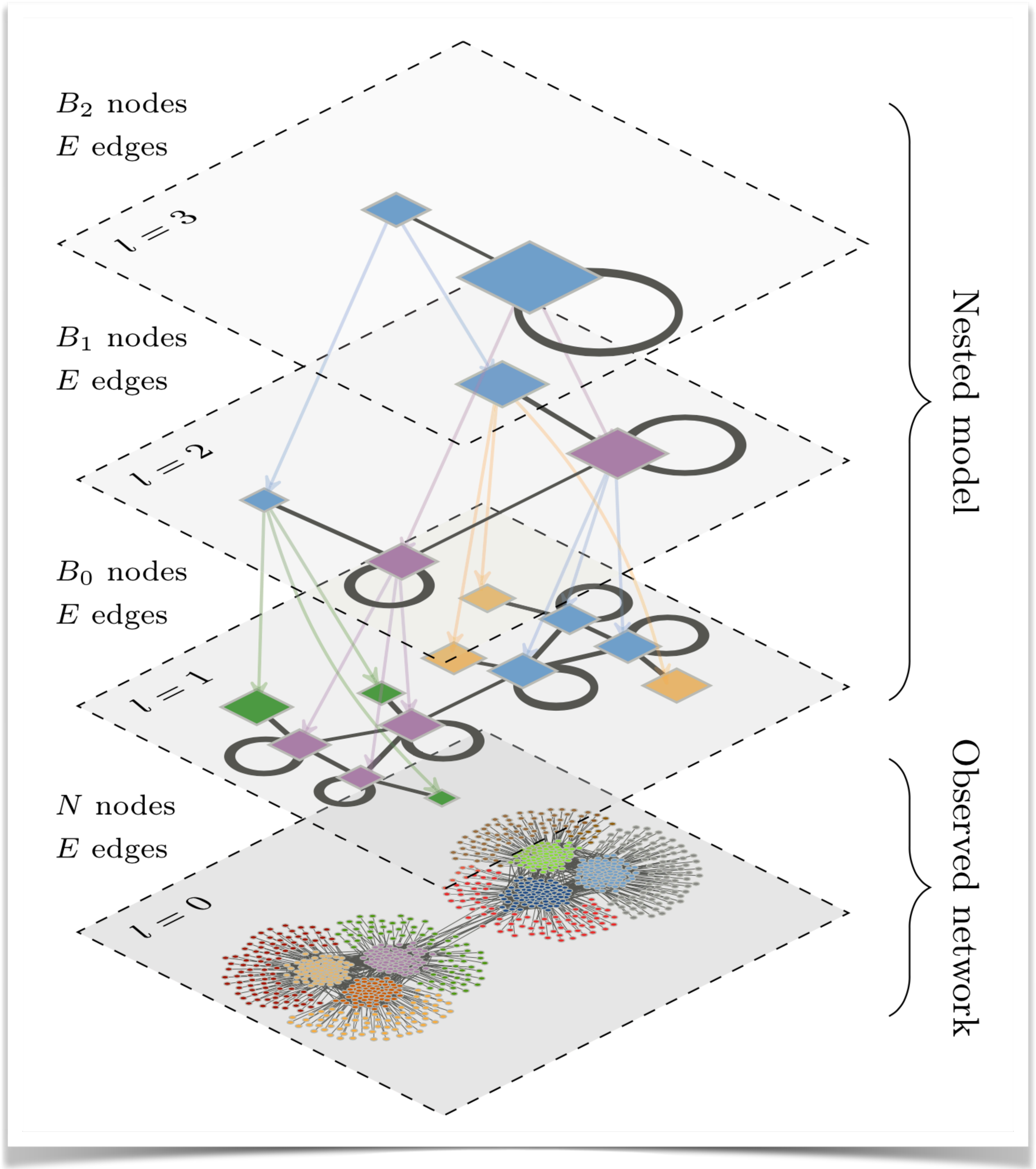
Comparison with hierarchical SBM

Avg. description length, the lower the better

Table I. Results for 24 empirical networks. Number of nodes n_{I}, n_{II}, mean degree $\langle k \rangle$, number of typ … B_{II}, and description length per edge Σ/E. Superscripts: b-biSBM, g-SBM, h-hSBM. L indicates the number of levels found by the hSBM. Reported values indicate best of 100 independent runs. Unless otherwise noted, data are accessible … plex Networks (ICON) [52]. The confidence level is marked with asterisks[a].

Dataset	N_{I}	N_{II}	$\langle k \rangle$	$(B_{\mathrm{I}}^{\mathrm{b}}, B_{\mathrm{II}}^{\mathrm{b}})$	$(B_{\mathrm{I}}^{\mathrm{g}}, B_{\mathrm{II}}^{\mathrm{g}})$	$(B_{\mathrm{I}}^{\mathrm{h}}, B_{\mathrm{II}}^{\mathrm{h}})$	$\langle L+1 \rangle$	Σ^{b}/E	Σ^{h}/E
Southern women interactions [53]	18	14	5.56	(1, 1)	(1, 1)	(1, 1)	2.0	**2.15***	2.26
Joern plant-herbivore web [54]	22	52	4.97	(2, 2)	(1, 1)	(1, 1)	2.0	**2.64***	2.74
Swingers and parties [55]	57	39	4.83	(1, 1)	(1, 1)	(1, 1)	2.0	**2.92***	2.97
McMullen pollination web [56]	54	105	2.57	(2, 2)	(2, 2)	(1, 1)	2.0	**2.87***	3.02
Ndrangheta criminals [57]	156	47	4.48	(3, 4)	(3, 3)	(3, 4)	2.87	**3.44***	3.49
Abu Sayyaf kidnappings[b] [58]	246	105	2.28	(2, 2)	(1, 1)	(1, 1)	2.0	**4.50***	4.54
Virus-host interactome [59]	53	307	2.52	(2, 2)	(1, 1)	(1, 1)	2.0	**3.78***	3.81
Clements-Long plant-pollinator [60]	275	96	4.98	(1, 1)	(1, 1)	(1, 1)	2.0	**3.45***	3.47
Human musculoskeletal system [61]	173	270	4.30	(7, 8)	(5, 5)	(8, 8)	4.01	**3.94**	3.94
Mexican drug trafficking[b] [62]	765	10	16.1	(12, 8)	(8, 7)	(10, 6)	3.11	**1.26***	1.29
Country-language network [63]	254	614	2.89	(4, 5)	(2, 2)	(4, 3)	2.11	**4.53***	4.56
Malaria gene similarity [44]	297	806	5.38	(15, 16)	(6, 6)	(25, 20)	4.95	4.73	**4.67***
Protein complex-drug [64]	739	680	5.20	(20, 22)	(14, 14)	(35, 39)	5.06	3.65	**3.50****
Robertson plant-pollinator [65]	456	1428	16.2	(20, 18)	(11, 11)	(20, 19)	4.0	**3.10***	3.10
Human gene-disease network [66]	1419	516	4.06	(13, 14)	(9, 9)	(35, 36)	5.04	5.02	**4.80****
Food ingredients-flavors web [67]	1525	1107	27.9	(27, 69)	(20, 29)	(42, 130)	4.91	2.55	**2.51****
Wikipedia doc-word network [68]	63	3140	24.8	(22, 206)	(18, 23)	(29, 71)	4.17	1.58	**1.51****
Foursquare check-ins [69]	2060	2876	11.0	(65, 66)	(40, 40)	(244, 248)	5.2	5.92	**5.09****
Ancient metabolic network [70]	5651	5252	4.22	(18, 22)	(5, 5)	(17, 21)	4.26	**5.68****	5.82
Marvel Universe characters [71]	6486	12942	9.95	(68, 72)	(67, 62)	(365, 314)	6.24	4.70	**4.42*****
Reuters news stories [72]	19757	38677	33.5	(396, 440)	(87, 108)	(294, 463)	6.25	4.22	**4.16*****
IMDb movie-actor dataset[c]	53158	39768	6.49	(91, 92)	(69, 68)	(264, 265)	6.22	7.40	**7.30*****
YouTube group memberships [73]	94238	30087	4.72	(62, 66)	(37, 38)	(221, 238)	5.9	**7.07****	7.13
DBpedia writer network [74]	89355	46213	2.13	(22, 26)	(2, 3)	(2, 3)	2.16	**10.32****	10.41

[a] Via the posterior odds ratio: $* : \Lambda < 10^{-2}$; $** : \Lambda < 10^{-100}$; $*** : \Lambda < 10^{-10000}$.
[b] Temporal data with timestamps are aggregated, making a multigraph.
[c] Data available at http://www.imdb.com/interfaces. IMDb copyright permits redistribution of data only in unaltered form.



See Peixoto, Phys. Rev. X, 4, 011047 (2014) for the hierarchical SBM.

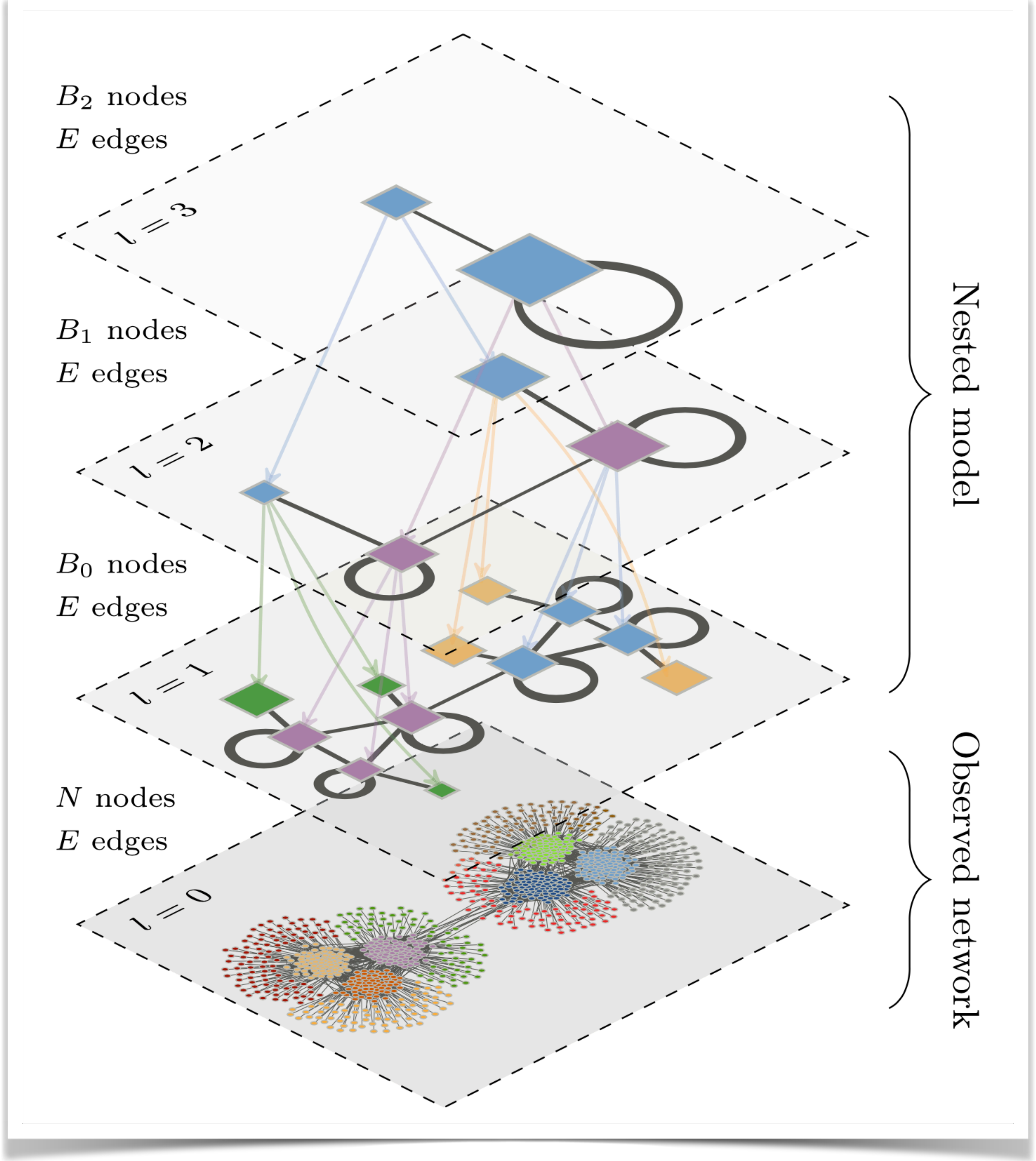
Comparison with hierarchical SBM

Table I. Results for 24 empirical networks. Number of nodes n_{I}, n_{II}, mean degree $\langle k \rangle$, number of typ... B_{II}, and description length per edge Σ/E. Superscripts: b-biSBM, g-SBM, h-hSBM. L indicates the number of levels found by the hSBM. Reported values indicate best of 100 independent runs. Unless otherwise noted, data are accessible ... Networks (ICON) [52]. The confidence level is marked with asterisks[a].

Avg. description length, the lower the better

Dataset	N_{I}	N_{II}	$\langle k \rangle$	$(B^b_{\mathrm{I}}, B^b_{\mathrm{II}})$	$(B^g_{\mathrm{I}}, B^g_{\mathrm{II}})$	$(B^h_{\mathrm{I}}, B^h_{\mathrm{II}})$	$\langle L+1 \rangle$	Σ^b/E	Σ^h/E
Southern women interactions [53]	18	14	5.56	(1, 1)	(1, 1)	(1, 1)	2.0	**2.15***	2.26
Joern plant-herbivore web [54]	22	52	4.97	(2, 2)	(1, 1)	(1, 1)	2.0	**2.64***	2.74
Swingers and parties [55]	57	39	4.83	(1, 1)	(1, 1)	(1, 1)	2.0	**2.92***	2.97
McMullen pollination web [56]	54	105	2.57	(2, 2)	(2, 2)	(1, 1)	2.0	**2.87***	3.02
Ndrangheta criminals [57]	156	47	4.48	(3, 4)	(3, 3)	(3, 4)	2.87	**3.44***	3.49
Abu Sayyaf kidnappings[b] [58]	246	105	2.28	(2, 2)	(1, 1)	(1, 1)	2.0	**4.50***	4.54
Virus-host interactome [59]	53	307	2.52	(2, 2)	(1, 1)	(1, 1)	2.0	**3.78***	3.81
Clements-Long plant-pollinator [60]	275	96	4.98	(1, 1)	(1, 1)	(1, 1)	2.0	**3.45***	3.47
Human musculoskeletal system [61]	173	270	4.30	(7, 8)	(5, 5)	(8, 8)	4.01	**3.94**	3.94
Mexican drug trafficking[b] [62]	765	10	16.1	(12, 8)	(8, 7)	(10, 6)	3.11	**1.26***	1.29
Country-language network [63]	254	614	2.89	(4, 5)	(2, 2)	(4, 3)	2.11	**4.53***	4.56
Malaria gene similarity [44]	297	806	5.38	(15, 16)	(6, 6)	(25, 20)	4.95	4.73	**4.67***
Protein complex-drug [64]	739	680	5.20	(20, 22)	(14, 14)	(35, 39)	5.06	3.65	**3.50****
Robertson plant-pollinator [65]	456	1428	16.2	(20, 18)	(11, 11)	(20, 19)	4.0	**3.10***	3.10
Human gene-disease network [66]	1419	516	4.06	(13, 14)	(9, 9)	(35, 36)	5.04	5.02	**4.80****
Food ingredients-flavors web [67]	1525	1107	27.9	(27, 69)	(20, 29)	(42, 130)	4.91	2.55	**2.51****
Wikipedia doc-word network [68]	63	3140	24.8	(22, 206)	(18, 23)	(29, 71)	4.17	1.58	**1.51****
Foursquare check-ins [69]	2060	2876	11.0	(65, 66)	(40, 40)	(244, 248)	5.2	5.92	**5.09****
Ancient metabolic network [70]	5651	5252	4.22	(18, 22)	(5, 5)	(17, 21)	4.26	**5.68****	5.82
Marvel Universe characters [71]	6486	12942	9.95	(68, 72)	(67, 62)	(365, 314)	6.24	4.70	**4.42*****
Reuters news stories [72]	19757	38677	33.5	(396, 440)	(87, 108)	(294, 463)	6.25	4.22	**4.16*****
IMDb movie-actor dataset[c]	53158	39768	6.49	(91, 92)	(69, 68)	(264, 265)	6.22	7.40	**7.30*****
YouTube group memberships [73]	94238	30087	4.72	(62, 66)	(37, 38)	(221, 238)	5.9	**7.07****	7.13
DBpedia writer network [74]	89355	46213	2.13	(22, 26)	(2, 3)	(2, 3)	2.16	**10.32****	10.41

[a] Via the posterior odds ratio: $* : \Lambda < 10^{-2}$; $** : \Lambda < 10^{-100}$; $*** : \Lambda < 10^{-10000}$.
[b] Temporal data with timestamps are aggregated, making a multigraph.
[c] Data available at http://www.imdb.com/interfaces. IMDb copyright permits redistribution of data only in unaltered form.



For smaller networks, bipartite SBM is better.

See Peixoto, Phys. Rev. X, 4, 011047 (2014) for the hierarchical SBM.

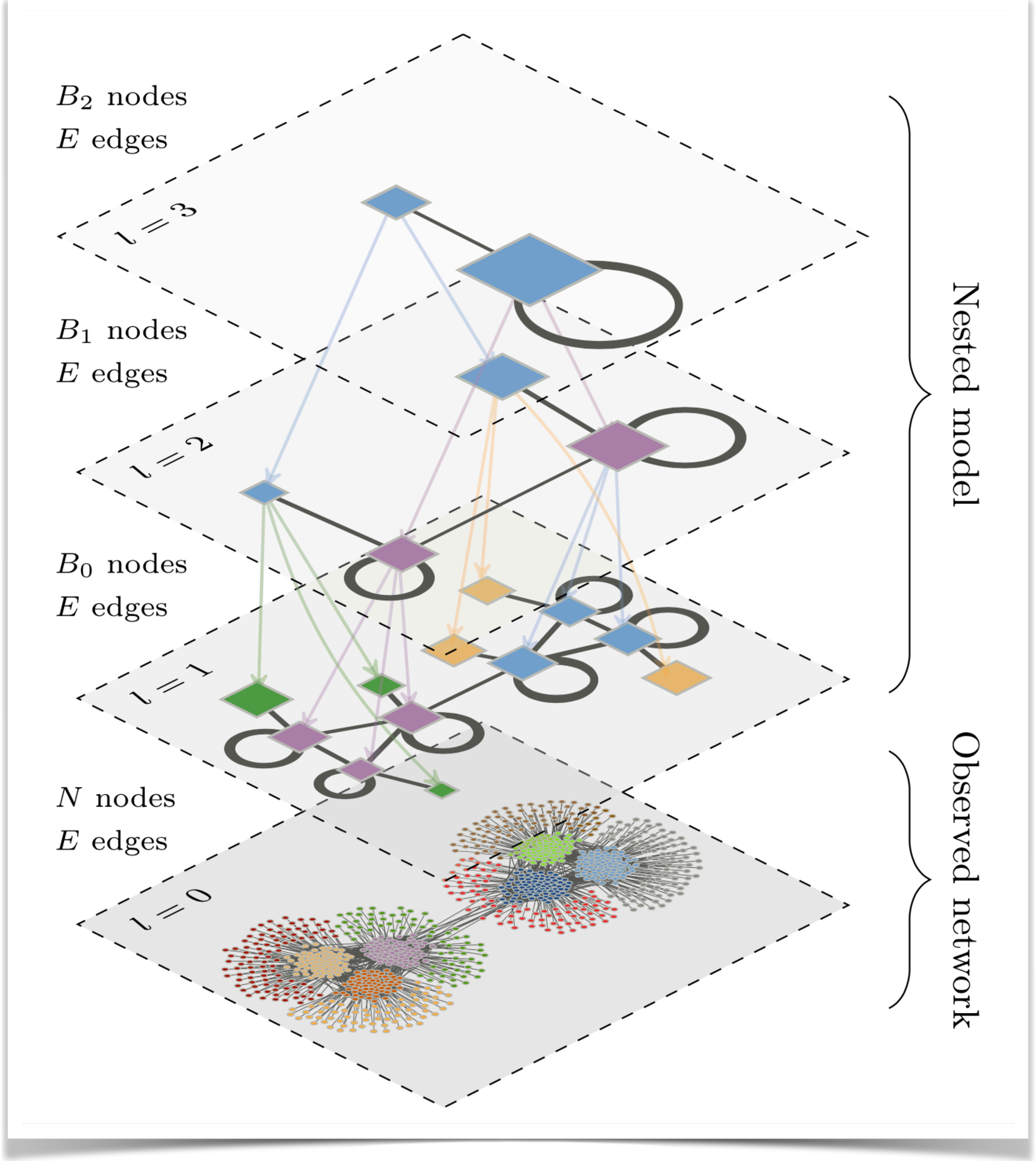
Comparison with hierarchical SBM

Table I. Results for 24 empirical networks. Number of nodes n_{I}, n_{II}, mean degree $\langle k \rangle$, number of typ… B_{II}, and description length per edge Σ/E. Superscripts: b-biSBM, g-SBM, h-hSBM. L indicates the number of levels found by the hSBM. Reported values indicate best of 100 independent runs. Unless otherwise noted, data are accessible … plex Networks (ICON) [52]. The confidence level is marked with asterisks[a].

Avg. description length, the lower the better

Dataset	N_{I}	N_{II}	$\langle k \rangle$	$(B_{\mathrm{I}}^{\mathrm{b}}, B_{\mathrm{II}}^{\mathrm{b}})$	$(B_{\mathrm{I}}^{\mathrm{g}}, B_{\mathrm{II}}^{\mathrm{g}})$	$(B_{\mathrm{I}}^{\mathrm{h}}, B_{\mathrm{II}}^{\mathrm{h}})$	$\langle L+1 \rangle$	Σ^{b}/E	Σ^{h}/E
Southern women interactions [53]	18	14	5.56	(1, 1)	(1, 1)	(1, 1)	2.0	**2.15***	2.26
Joern plant-herbivore web [54]	22	52	4.97	(2, 2)	(1, 1)	(1, 1)	2.0	**2.64***	2.74
Swingers and parties [55]	57	39	4.83	(1, 1)	(1, 1)	(1, 1)	2.0	**2.92***	2.97
McMullen pollination web [56]	54	105	2.57	(2, 2)	(2, 2)	(1, 1)	2.0	**2.87***	3.02
Ndrangheta criminals [57]	156	47	4.48	(3, 4)	(3, 3)	(3, 4)	2.87	**3.44***	3.49
Abu Sayyaf kidnappings[b] [58]	246	105	2.28	(2, 2)	(1, 1)	(1, 1)	2.0	**4.50***	4.54
Virus-host interactome [59]	53	307	2.52	(2, 2)	(1, 1)	(1, 1)	2.0	**3.78***	3.81
Clements-Long plant-pollinator [60]	275	96	4.98	(1, 1)	(1, 1)	(1, 1)	2.0	**3.45***	3.47
Human musculoskeletal system [61]	173	270	4.30	(7, 8)	(5, 5)	(8, 8)	4.01	**3.94**	3.94
Mexican drug trafficking[b] [62]	765	10	16.1	(12, 8)	(8, 7)	(10, 6)	3.11	**1.26***	1.29
Country-language network [63]	254	614	2.89	(4, 5)	(2, 2)	(4, 3)	2.11	**4.53***	4.56
Malaria gene similarity [44]	297	806	5.38	(15, 16)	(6, 6)	(25, 20)	4.95	4.73	**4.67***
Protein complex-drug [64]	739	680	5.20	(20, 22)	(14, 14)	(35, 39)	5.06	3.65	**3.50****
Robertson plant-pollinator [65]	456	1428	16.2	(20, 18)	(11, 11)	(20, 19)	4.0	**3.10***	3.10
Human gene-disease network [66]	1419	516	4.06	(13, 14)	(9, 9)	(35, 36)	5.04	5.02	**4.80****
Food ingredients-flavors web [67]	1525	1107	27.9	(27, 69)	(20, 29)	(42, 130)	4.91	2.55	**2.51****
Wikipedia doc-word network [68]	63	3140	24.8	(22, 206)	(18, 23)	(29, 71)	4.17	1.58	**1.51****
Foursquare check-ins [69]	2060	2876	11.0	(65, 66)	(40, 40)	(244, 248)	5.2	5.92	**5.09****
Ancient metabolic network [70]	5651	5252	4.22	(18, 22)	(5, 5)	(17, 21)	4.26	**5.68****	5.82
Marvel Universe characters [71]	6486	12942	9.95	(68, 72)	(67, 62)	(365, 314)	6.24	4.70	**4.42*****
Reuters news stories [72]	19757	38677	33.5	(396, 440)	(87, 108)	(294, 463)	6.25	4.22	**4.16*****
IMDb movie-actor dataset[c]	53158	39768	6.49	(91, 92)	(69, 68)	(264, 265)	6.22	7.40	**7.30*****
YouTube group memberships [73]	94238	30087	4.72	(62, 66)	(37, 38)	(221, 238)	5.9	**7.07****	7.13
DBpedia writer network [74]	89355	46213	2.13	(22, 26)	(2, 3)	(2, 3)	2.16	**10.32****	10.41

[a] Via the posterior odds ratio: $^{*}: \Lambda < 10^{-2}$; $^{**}: \Lambda < 10^{-100}$; $^{***}: \Lambda < 10^{-10000}$.
[b] Temporal data with timestamps are aggregated, making a multigraph.
[c] Data available at http://www.imdb.com/interfaces. IMDb copyright permits redistribution of data only in unaltered form.



For smaller networks, bipartite SBM is better.

For larger+sparser networks, bipartite SBM is better.

See Peixoto, Phys. Rev. X, 4, 011047 (2014) for the hierarchical SBM.

Implications

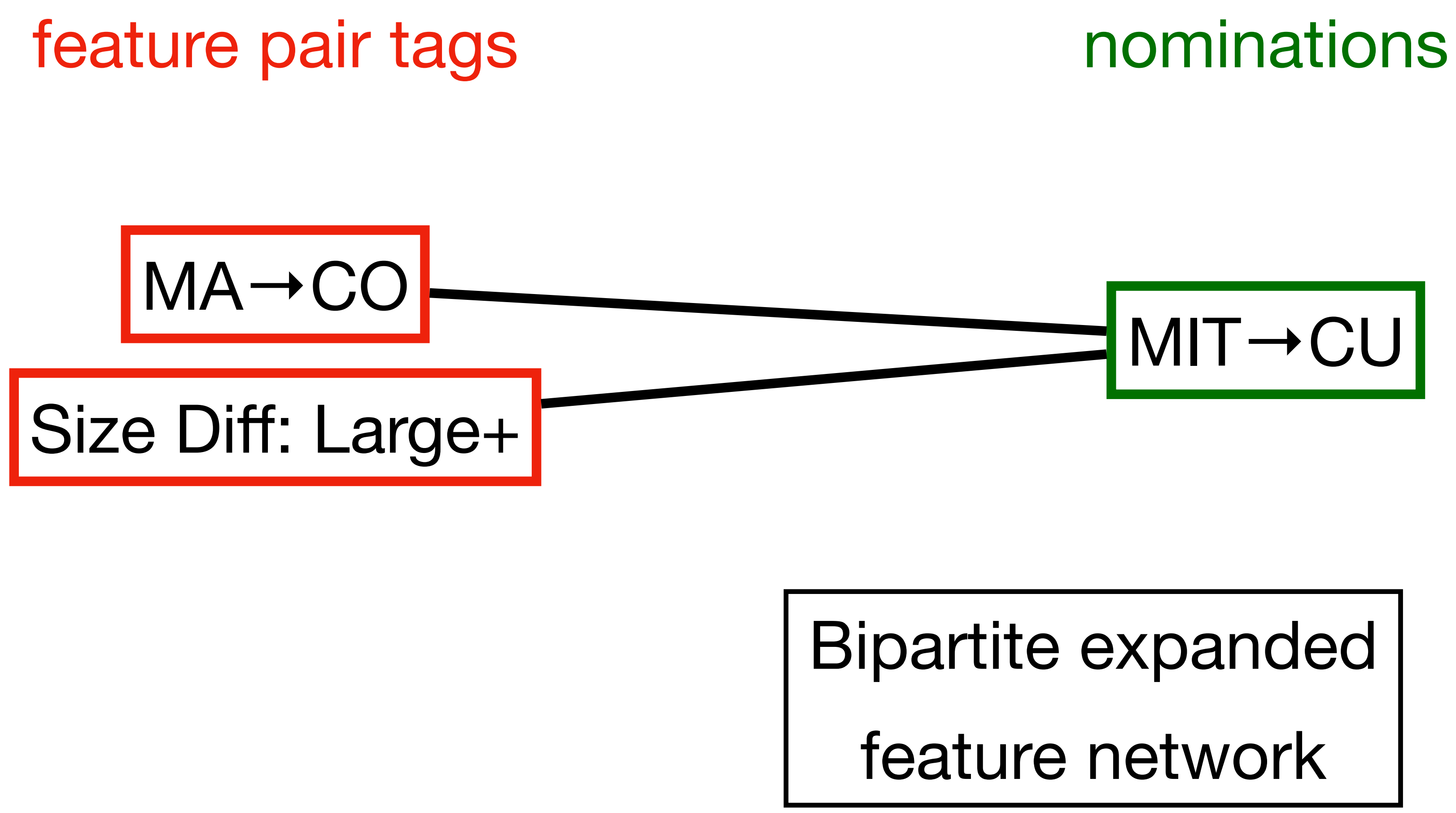
Bayesian framework. data becomes a part of the model, algorithm becomes data-dependent, i.e., some data may be harder to fit (opportunities for improvement).

Always choose a simpler model. If not possible to attain (e.g., no good definition of "simple-ness"), choose a more flexible one.

A handful of groups explain most nomination behaviors

Type-I nodes: feature pair tags, to cluster as B_{I} groups

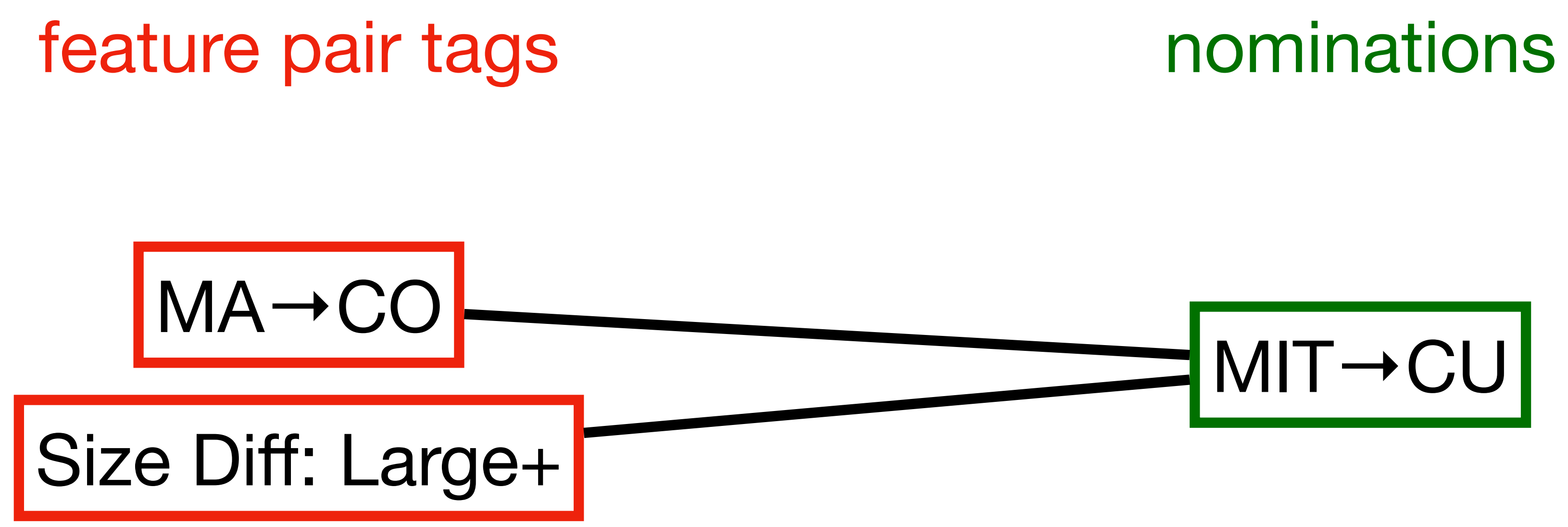
Type-II nodes: nominations, to cluster as B_{II} groups



A handful of groups explain most nomination behaviors

Type-I nodes: feature pair tags, to cluster as B_I groups

Type-II nodes: nominations, to cluster as B_{II} groups



Bipartite expanded feature network

Code	Type of Institutions	N	N_{I}	N_{II}	E	B_{I}	B_{II}	Σ/E
15	Doctoral Universities: Highest Research Activity	117	946	1956	13692	5	2	2.45
16	Doctoral Universities: Higher Research Activity	112	833	1757	12299	8	2	2.67
17	Doctoral Universities: Moderate Research Activity	102	596	1815	12705	11	3	2.68
18	Master's Colleges & Universities: Larger Programs	251	958	4118	28825	16	3	2.73
19	Master's Colleges & Universities: Medium Programs	129	656	2082	14574	14	3	2.71
20	Master's Colleges & Universities: Small Programs	78	502	1169	8183	6	2	2.70
21	Baccalaureate Colleges: Arts & Sciences Focus	176	778	3105	21735	4	2	2.48
22	Baccalaureate Colleges: Diverse Fields	144	698	2129	14903	12	3	2.69
23	Baccalaureate/Associate's Colleges: Mixed Baccalaureate/Associate's	126	32	232	1624	4	2	2.45
24	Special Focus Four-Year: Faith-Related Institutions	1	9	2	14	1	1	1.31
26	Special Focus Four-Year: Other Health Professions Schools	5	53	42	293	1	1	2.55
27	Special Focus Four-Year: Engineering Schools	3	45	24	168	1	1	2.53
28	Special Focus Four-Year: Other Technology-Related Schools	3	53	38	266	1	1	2.37
29	Special Focus Four-Year: Business & Management Schools	6	47	122	854	1	1	2.27
30	Special Focus Four-Year: Arts, Music & Design Schools	18	138	196	1372	1	1	2.41
32	Special Focus Four-Year: Other Special Focus Institutions	2	38	30	209	1	1	2.34

total = 1179

A handful of groups explain most nomination behaviors

Type-I nodes: feature pair tags, to cluster as B_I groups

Type-II nodes: nominations, to cluster as B_{II} groups

feature pair tags — nominations

MA→CO — MIT→CU

Size Diff: Large+ — MIT→CU

Bipartite expanded feature network

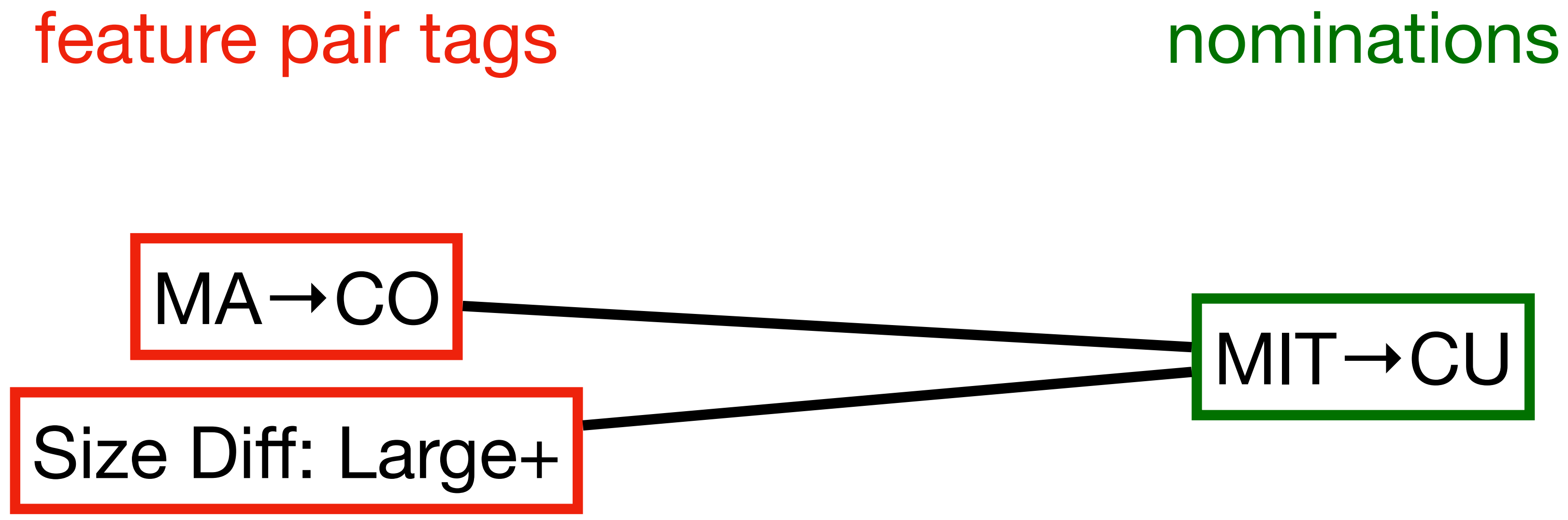
Code	Type of Institutions	N	N_I	N_II	E	B_I	B_II	Σ/E
15	Doctoral Universities: Highest Research Activity	117	946	1956	13692	5	2	2.45
16	Doctoral Universities: Higher Research Activity	112	833	1757	12299	8	2	2.67
17	Doctoral Universities: Moderate Research Activity	102	596	1815	12705	11	3	2.68
18	Master's Colleges & Universities: Larger Programs	251	958	4118	28825	16	3	2.73
19	Master's Colleges & Universities: Medium Programs	129	656	2082	14574	14	3	2.71
20	Master's Colleges & Universities: Small Programs	78	502	1169	8183	6	2	2.70
21	Baccalaureate Colleges: Arts & Sciences Focus	176	778	3105	21735	4	2	2.48
22	Baccalaureate Colleges: Diverse Fields	144	698	2129	14903	12	3	2.69
23	Baccalaureate/Associate's Colleges: Mixed Baccalaureate/Associate's	126	32	232	1624	4	2	2.45
24	Special Focus Four-Year: Faith-Related Institutions	1	9	2	14	1	1	1.31
26	Special Focus Four-Year: Other Health Professions Schools	5	53	42	293	1	1	2.55
27	Special Focus Four-Year: Engineering Schools	3	45	24	168	1	1	2.53
28	Special Focus Four-Year: Other Technology-Related Schools	3	53	38	266	1	1	2.37
29	Special Focus Four-Year: Business & Management Schools	6	47	122	854	1	1	2.27
30	Special Focus Four-Year: Arts, Music & Design Schools	18	138	196	1372	1	1	2.41
32	Special Focus Four-Year: Other Special Focus Institutions	2	38	30	209	1	1	2.34

total = 1179

A handful of groups explain most nomination behaviors

Type-I nodes: feature pair tags, to cluster as B_I groups

Type-II nodes: nominations, to cluster as B_{II} groups



Bipartite expanded feature network

Code	Type of Institutions	N	N_I	N_{II}	E	B_I	B_{II}	Σ/E
15	Doctoral Universities: Highest Research Activity	117	946	1956	13692	5	2	2.45
16	Doctoral Universities: Higher Research Activity	112	833	1757	12299	8	2	2.67
17	Doctoral Universities: Moderate Research Activity	102	596	1815	12705	11	3	2.68
18	Master's Colleges & Universities: Larger Programs	251	958	4118	28825	16	3	2.73
19	Master's Colleges & Universities: Medium Programs	129	656	2082	14574	14	3	2.71
20	Master's Colleges & Universities: Small Programs	78	502	1169	8183	6	2	2.70
21	Baccalaureate Colleges: Arts & Sciences Focus	176	778	3105	21735	4	2	2.48
22	Baccalaureate Colleges: Diverse Fields	144	698	2129	14903	12	3	2.69
23	Baccalaureate/Associate's Colleges: Mixed Baccalaureate/Associate's	126	32	232	1624	4	2	2.45
24	Special Focus Four-Year: Faith-Related Institutions	1	9	2	14	1	1	1.31
26	Special Focus Four-Year: Other Health Professions Schools	5	53	42	293	1	1	2.55
27	Special Focus Four-Year: Engineering Schools	3	45	24	168	1	1	2.53
28	Special Focus Four-Year: Other Technology-Related Schools	3	53	38	266	1	1	2.37
29	Special Focus Four-Year: Business & Management Schools	6	47	122	854	1	1	2.27
30	Special Focus Four-Year: Arts, Music & Design Schools	18	138	196	1372	1	1	2.41
32	Special Focus Four-Year: Other Special Focus Institutions	2	38	30	209	1	1	2.34

total = 1179

A handful of groups explain most nomination behaviors (cont.)

Type	Top feature pairs	Group 1	Group 2
R1	non-HBCU to non-HBCU, to higher admission rate, to other R1	10.4%	36.8%
	78.9% intrastate, 93.5% interstate, non-HBCU to HBCU, to similar enrollment size, to R2 or to M:Large	2.5%	22.1%
	Public to Public/Private, to lower grad rate	0%	13.2%
	5.3% intrastate, to higher grad rate, to similar admission rate	3.9%	6.1%
	15.8% intrastate, 6.5% interstate, Private to Public/Private	4.9%	0.1%

R2	non-HBCU to non-HBCU, to higher admission rate, to R1, to M:Large	28.4%	12.7%
	to lower enrollment size, to higher graduation rate, to similar admission rate, to other R2	19.7%	4.7%
	69.6% intrastate, 90.7% interstate, HBCU to HBCU/non-HBCU, non-HBCU to HBCU, to M:Small, to Bacc. Colleges (diverse fields), Private to Public/Private for-profit	11.5%	1.3%
	Public to Public	9.6%	0%
	to higher admission rate, to R3	1.9%	2.1%
	Private to Private	0%	3.7%
	17.4% intrastate, 4% interstate, to M:Med, S:Engineer, Public to Private	0.6%	1.8%
	13% intrastate, 5.3% interstate, to B:A&S, to Special Focus Four-Year (Arts/Music/Design)	0%	1.9%

*See dissertation for clustering of other Carnegie classifications

A handful of groups explain most nomination behaviors (cont.)

Type	Top feature pairs	Group 1	Group 2
R1	non-HBCU to non-HBCU, to higher admission rate, to other R1	10.4%	36.8%
	78.9% intrastate, 93.5% interstate, non-HBCU to HBCU, to similar enrollment size, to R2 or to M:Large	2.5%	22.1%
	Public to Public/Private, to lower grad rate	0%	13.2%
	5.3% intrastate, to higher grad rate, to similar admission rate	3.9%	6.1%
	15.8% intrastate, 6.5% interstate, Private to Public/Private	4.9%	0.1%

R2	non-HBCU to non-HBCU, to higher admission rate, to R1, to M:Large	28.4%	12.7%
	to lower enrollment size, to higher graduation rate, to similar admission rate, to other R2	19.7%	4.7%
	69.6% intrastate, 90.7% interstate, HBCU to HBCU/non-HBCU, non-HBCU to HBCU, to M:Small, to Bacc. Colleges (diverse fields), Private to Public/Private for-profit	11.5%	1.3%
	Public to Public	9.6%	0%
	to higher admission rate, to R3	1.9%	2.1%
	Private to Private	0%	3.7%
	17.4% intrastate, 4% interstate, to M:Med, S:Engineer, Public to Private	0.6%	1.8%
	13% intrastate, 5.3% interstate, to B:A&S, to Special Focus Four-Year (Arts/Music/Design)	0%	1.9%

Homophily. In R1 (47.2%) and R2 schools (24.4%), but R2 shows more aspiration to R1 (41.1%).

*See dissertation for clustering of other Carnegie classifications

A handful of groups explain most nomination behaviors (cont.)

Type	Top feature pairs	Group 1	Group 2
R1	non-HBCU to non-HBCU, to higher admission rate, to other R1	10.4%	36.8%
	78.9% intrastate, 93.5% interstate, non-HBCU to HBCU, to similar enrollment size, to R2 or to M:Large	2.5%	22.1%
	Public to Public/Private, to lower grad rate	0%	13.2%
	5.3% intrastate, to higher grad rate, to similar admission rate	3.9%	6.1%
	15.8% intrastate, 6.5% interstate, Private to Public/Private	4.9%	0.1%

R2	non-HBCU to non-HBCU, to higher admission rate, to R1, to M:Large	28.4%	12.7%
	to lower enrollment size, to higher graduation rate, to similar admission rate, to other R2	19.7%	4.7%
	69.6% intrastate, 90.7% interstate, HBCU to HBCU/non-HBCU, non-HBCU to HBCU, to M:Small, to Bacc. Colleges (diverse fields), Private to Public/Private for-profit	11.5%	1.3%
	Public to Public	9.6%	0%
	to higher admission rate, to R3	1.9%	2.1%
	Private to Private	0%	3.7%
	17.4% intrastate, 4% interstate, to M:Med, S:Engineer, Public to Private	0.6%	1.8%
	13% intrastate, 5.3% interstate, to B:A&S, to Special Focus Four-Year (Arts/Music/Design)	0%	1.9%

Homophily. In R1 (47.2%) and R2 schools (24.4%), but R2 shows more aspiration to R1 (41.1%).

Public schools. Tend to endorse others with a higher admission rate (R1: 13.2%; R2: 9.6%).

*See dissertation for clustering of other Carnegie classifications

A handful of groups explain most nomination behaviors (cont.)

Type	Top feature pairs	Group 1	Group 2
R1	non-HBCU to non-HBCU, to higher admission rate, to other R1	10.4%	36.8%
	78.9% intrastate, 93.5% interstate, non-HBCU to HBCU, to similar enrollment size, to R2 or to M:Large	2.5%	22.1%
	Public to Public/Private, to lower grad rate	0%	13.2%
	5.3% intrastate, to higher grad rate, to similar admission rate	3.9%	6.1%
	15.8% intrastate, 6.5% interstate, Private to Public/Private	4.9%	0.1%

R2	non-HBCU to non-HBCU, to higher admission rate, to R1, to M:Large	28.4%	12.7%
	to lower enrollment size, to higher graduation rate, to similar admission rate, to other R2	19.7%	4.7%
	69.6% intrastate, 90.7% interstate, HBCU to HBCU/non-HBCU, non-HBCU to HBCU, to M:Small, to Bacc. Colleges (diverse fields), Private to Public/Private for-profit	11.5%	1.3%
	Public to Public	9.6%	0%
	to higher admission rate, to R3	1.9%	2.1%
	Private to Private	0%	3.7%
	17.4% intrastate, 4% interstate, to M:Med, S:Engineer, Public to Private	0.6%	1.8%
	13% intrastate, 5.3% interstate, to B:A&S, to Special Focus Four-Year (Arts/Music/Design)	0%	1.9%

Homophily. In R1 (47.2%) and R2 schools (24.4%), but R2 shows more aspiration to R1 (41.1%).

Public schools. Tend to endorse others with a higher admission rate (R1: 13.2%; R2: 9.6%).

Niche effect. R1 wants to increase grad rate than other intrastate R1s; R2 aspires the same, but be fitter (lower enrollment size).

*See dissertation for clustering of other Carnegie classifications

Implications

Implications

All institutions. Name higher prestige institutions as their peers.

Implications

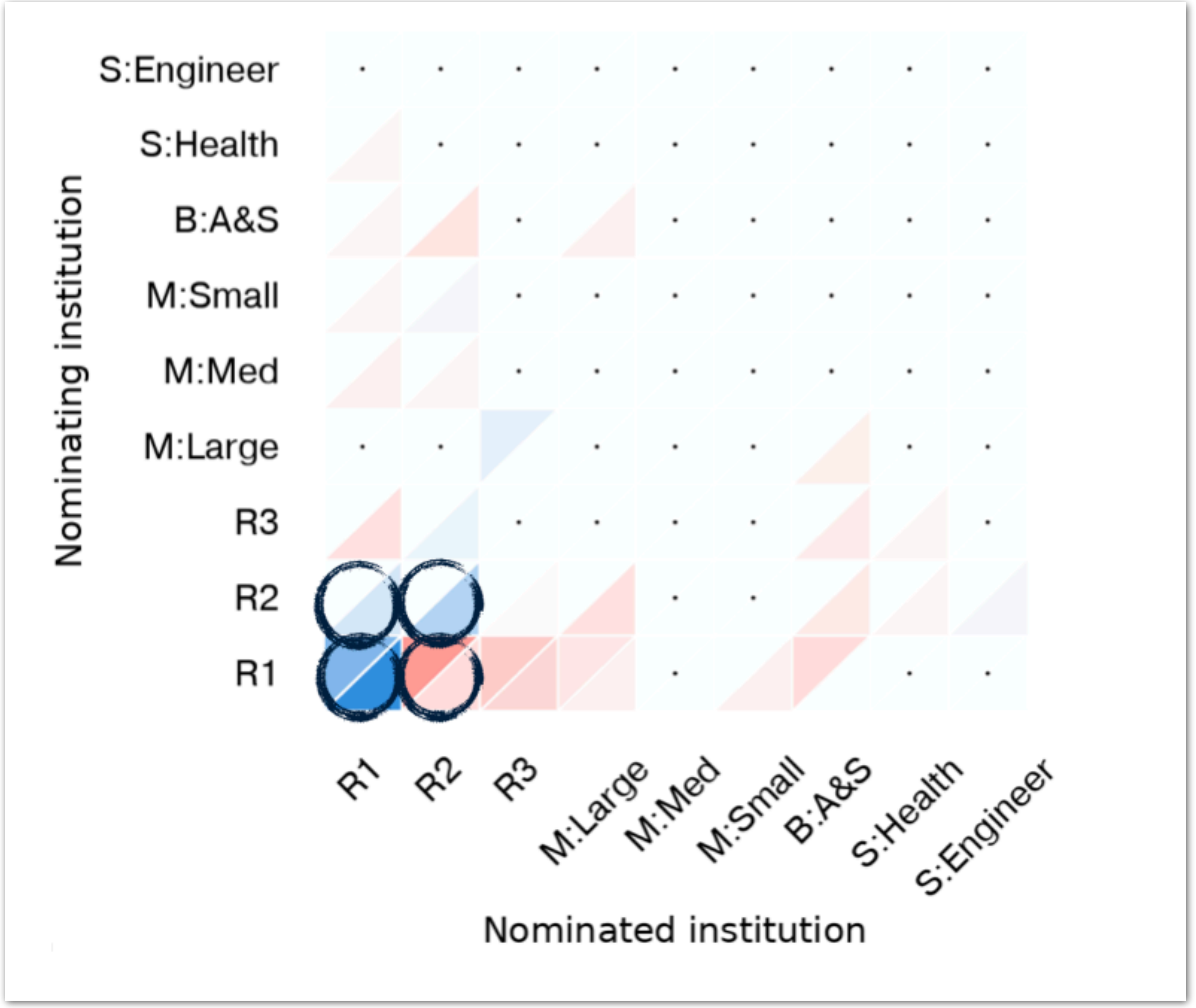
All institutions. Name higher prestige institutions as their peers.

Elite institutions. Particularly aspire prestige.

Implications

All institutions. Name higher prestige institutions as their peers.

Elite institutions. Particularly aspire prestige.

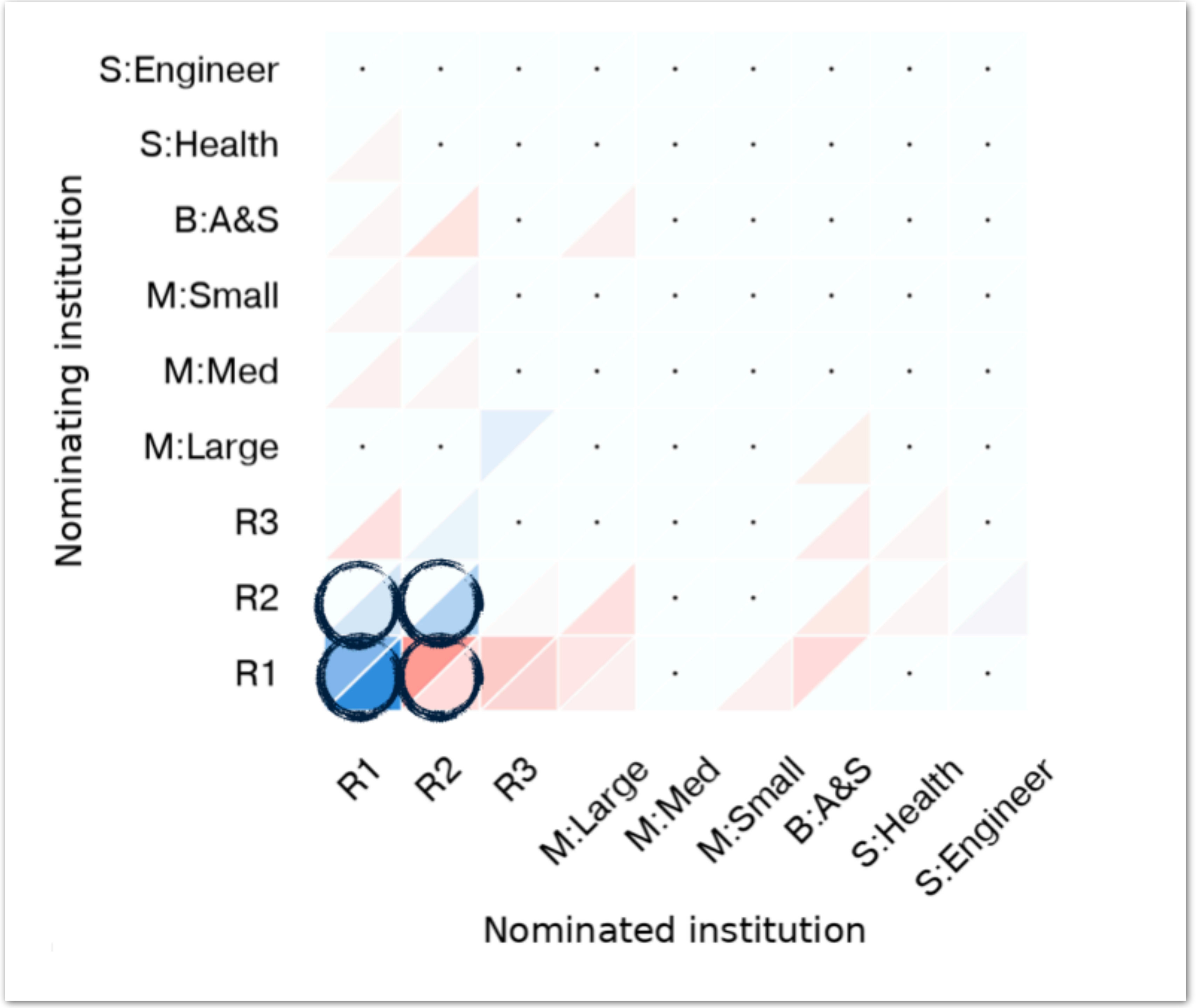


Implications

All institutions. Name higher prestige institutions as their peers.

Elite institutions. Particularly aspire prestige.

Others. Explainable by a handful of feature groups.



Outline

1. Context & motivation (chapter 1)
2. Aspiration of prestige in institutional peer selection (chapter 5)
3. Community detection in bipartite networks (chapter 2)
4. Regularized methods for efficient ranking (chapter 3)
5. A detour in higher-order structures (chapter 4)
6. Conclusions & future research (chapter 6)

Outline

Structure and inference in annotated networks

M. E. J. Newman & Aaron Clauset

Nature Communications **7**, Article number: 11863 (2016)

Network Structure, Metadata, and the Prediction of Missing Nodes and Annotations

Darko Hric, Tiago P. Peixoto, and Santo Fortunato

Phys. Rev. X **6**, 031038 – Published 12 September 2016

1. Context & motivation (chapter 1)
2. Aspiration of prestige in institutional peer selection (chapter 5)
3. Community detection in bipartite networks (chapter 2)
4. Regularized methods for efficient ranking (chapter 3)
5. A detour in higher-order structures (chapter 4)
6. Conclusions & future research (chapter 6)

Outline

Structure and inference in annotated networks

M. E. J. Newman & Aaron Clauset

Nature Communications **7**, Article number: 11863 (2016)

Network Structure, Metadata, and the Prediction of Missing Nodes and Annotations

Darko Hric, Tiago P. Peixoto, and Santo Fortunato

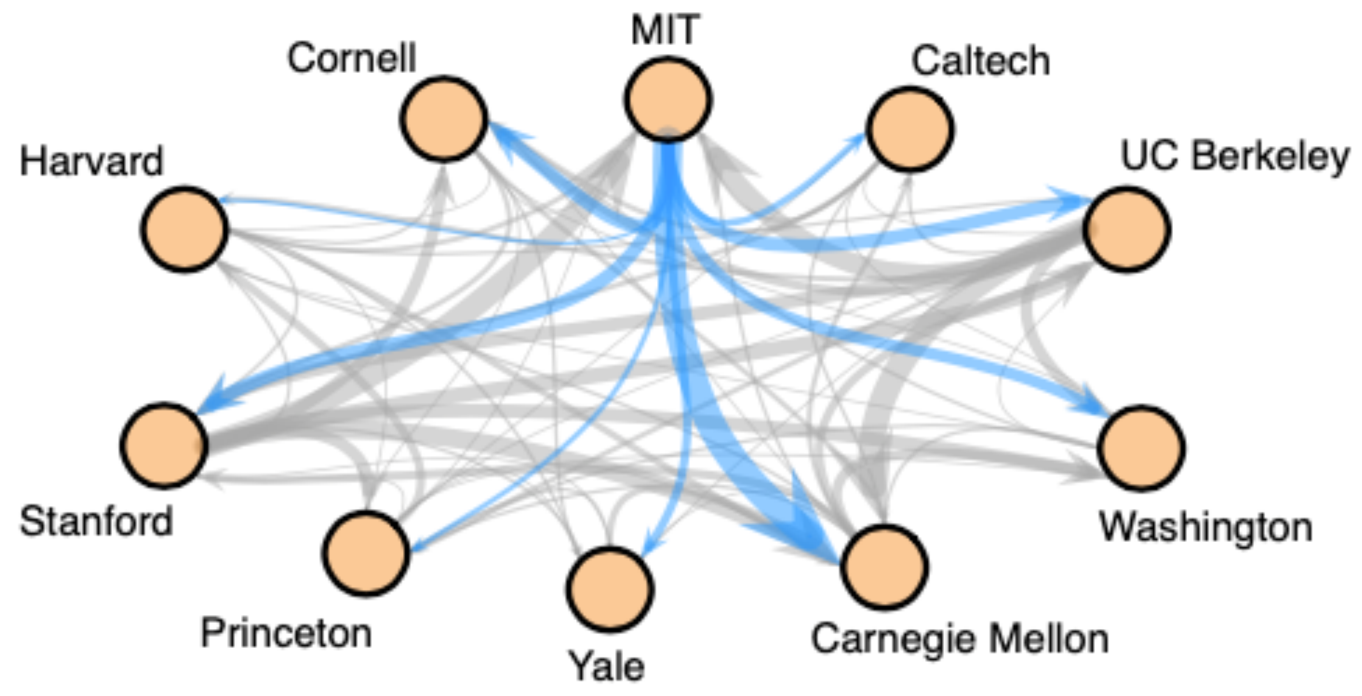
Phys. Rev. X **6**, 031038 – Published 12 September 2016

1. Context & motivation (chapter 1)
2. Aspiration of prestige in institutional peer selection (chapter 5)
3. Community detection in bipartite networks (c
4. Regularized methods for efficient ranking (ch
5. A detour in higher-order structures (chapter 4)
6. Conclusions & future research (chapter 6)

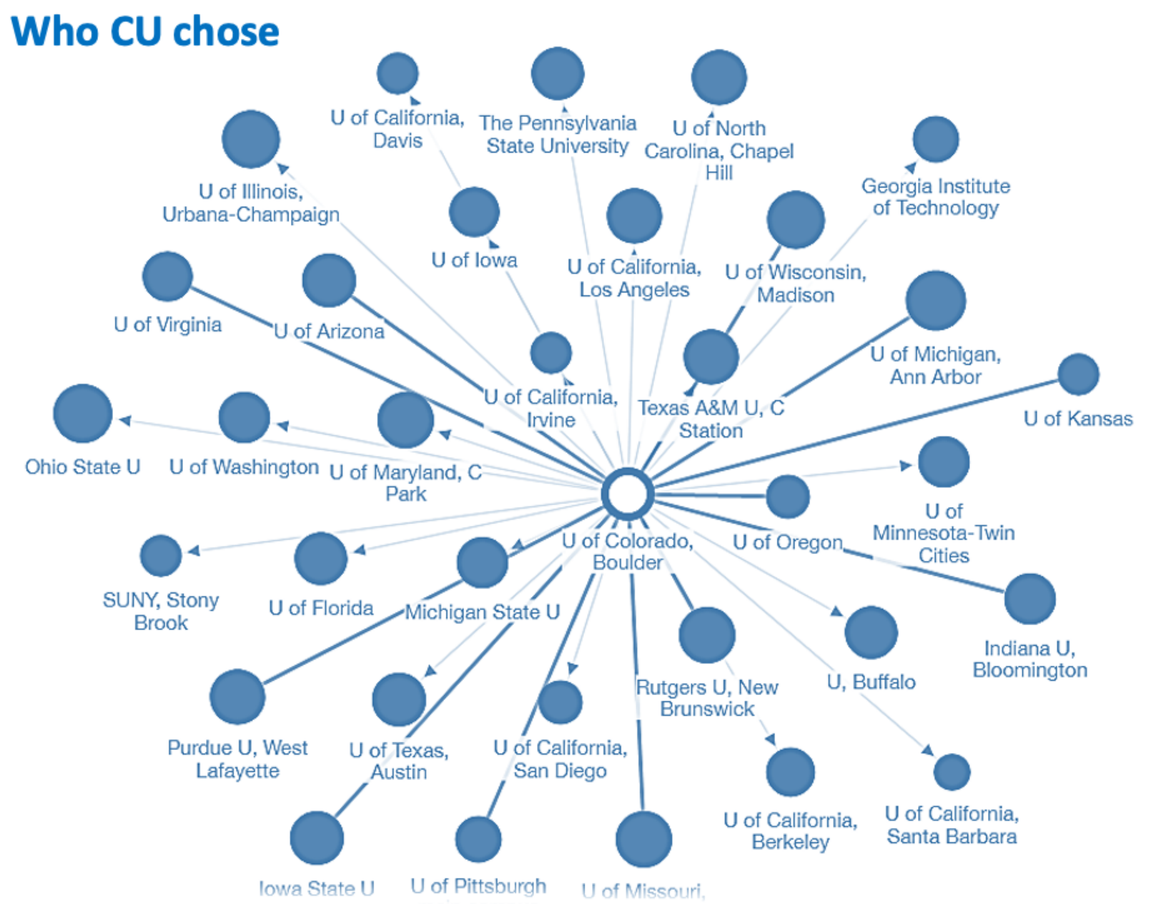
SpringRank

Each directed edge $u \to v$ implies PhD from u becomes faculty at v

Each directed edge $u \to v$ implies institution u selects institution v



Computer science faculty hiring network; http://tuvalu.santafe.edu/~aaronc/facultyhiring/. See also: *Sci. Adv.* 1(1), e1400005, 2015.



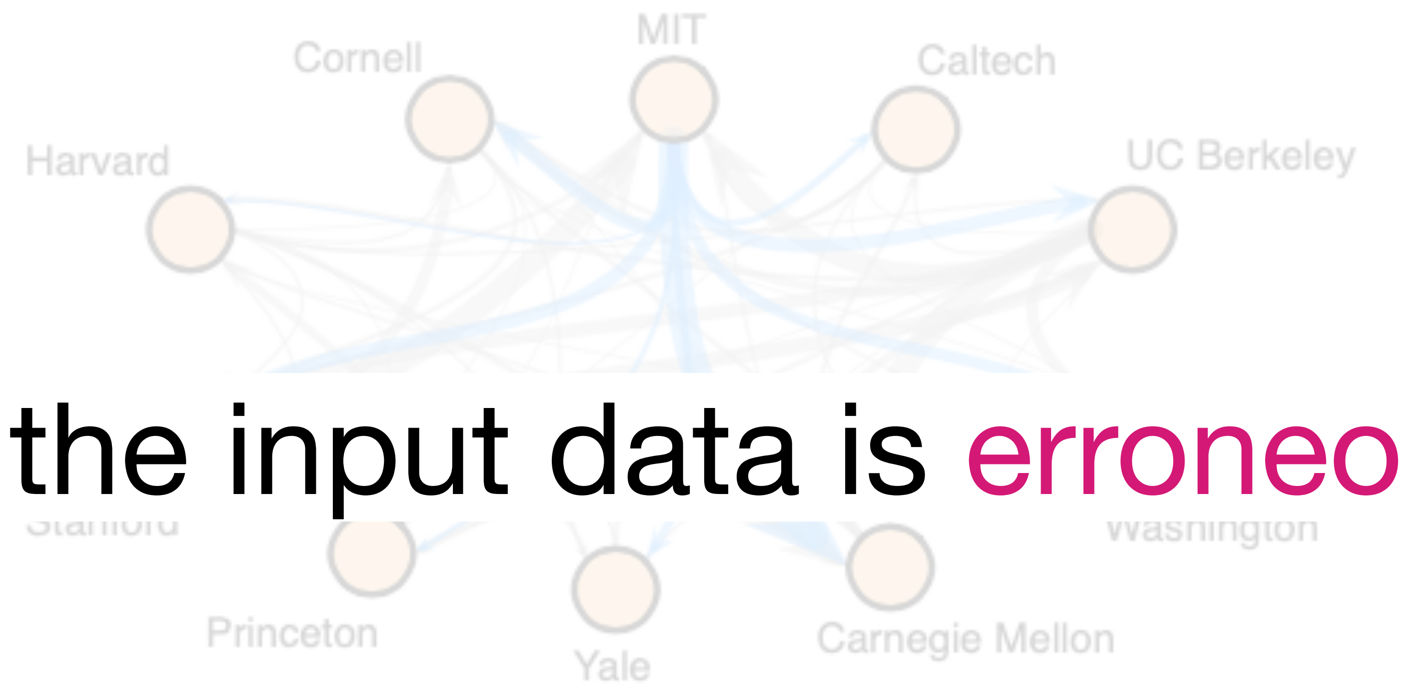
SpringRank: Let $\mathbf{A} = \{A_{ij}\}$ be the adjacency matrix of a directed multigraph. The SpringRank Model finds the ranking $\mathbf{s}^*$ of the nodes by solving:

$$\underset{s \in \mathbb{R}^N}{\text{minimize}} \frac{1}{2} \sum_{ij} A_{ij}(s_i - s_j - 1)^2 \ .$$

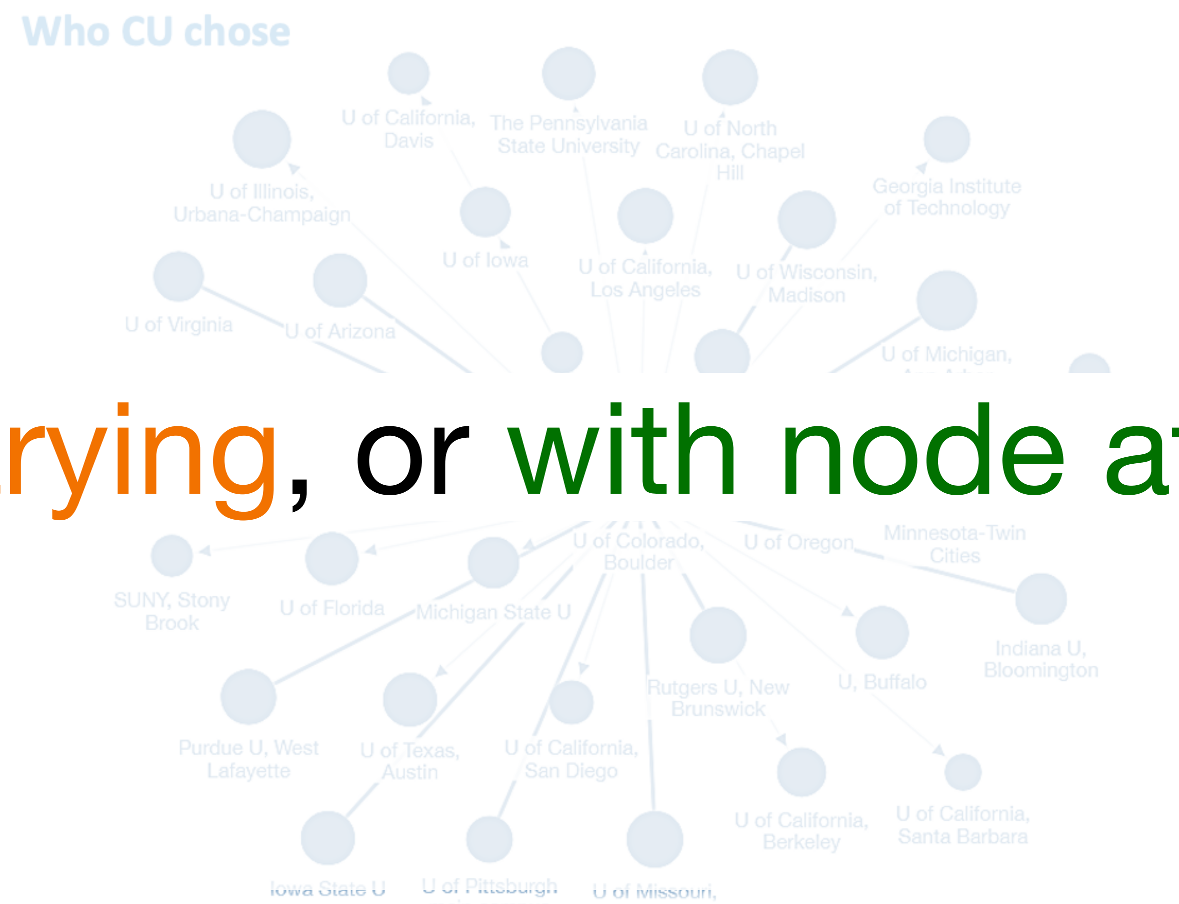
SpringRank

Each directed edge $u \to v$ implies PhD from u becomes faculty at v

Each directed edge $u \to v$ implies institution u selects institution v



Computer science faculty hiring network; http://tuvalu.santafe.edu/~aaronc/facultyhiring/. See also: *Sci. Adv.* 1(1), e1400005, 2015.



When the input data is erroneous, time-varying, or with node attributes

SpringRank: Let $\mathbf{A} = \{A_{ij}\}$ be the adjacency matrix of a directed multigraph. The SpringRank Model finds the ranking $\mathbf{s}^*$ of the nodes by solving:

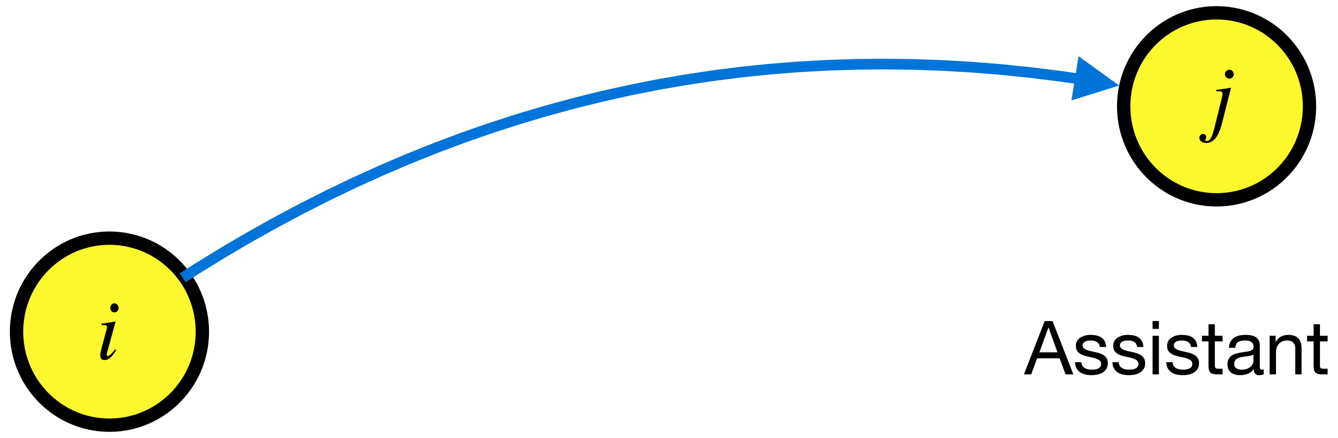
$$\underset{s \in \mathbb{R}^N}{\text{minimize}} \; \frac{1}{2} \sum_{ij} A_{ij} (s_i - s_j - 1)^2 \; .$$

Motivation

Math PhD Exchange Dataset: Proxy for university i hiring a graduate from university j at a time near t. We view this as an endorsement by j that graduates of t are of high quality.

$t = 2023$

Tzu-Chi, Maya, & Michael graduate (🤞)



Assistant Professor

University of Colorado

PhD, Computer Science

University of Chicago

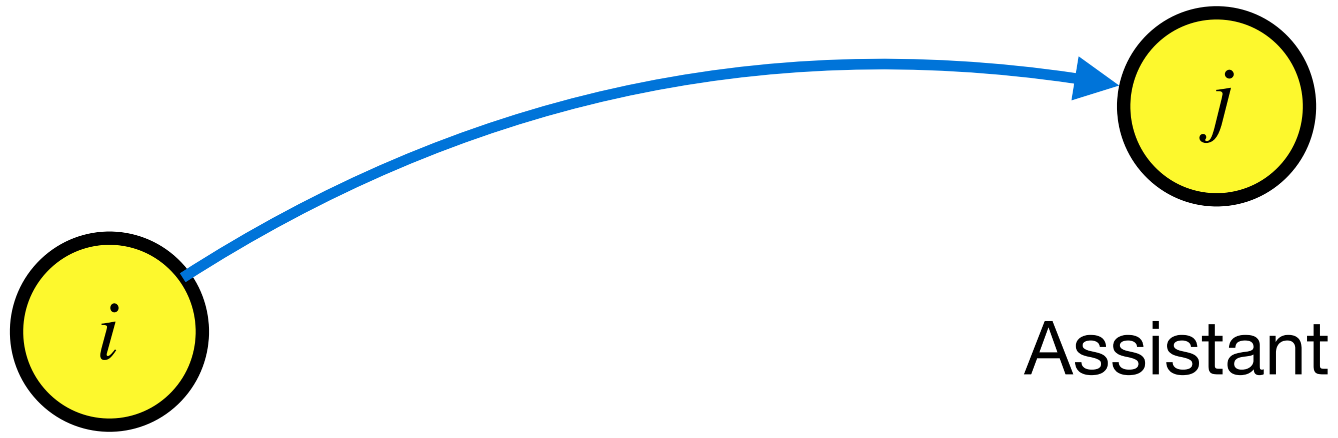
$w > 1$

Motivation

Math PhD Exchange Dataset: Proxy for university i hiring a graduate from university j at a time near t. We view this as an endorsement by j that graduates of t are of high quality.

$t = 2023$

Tzu-Chi, Maya, & Michael graduate (🤞)



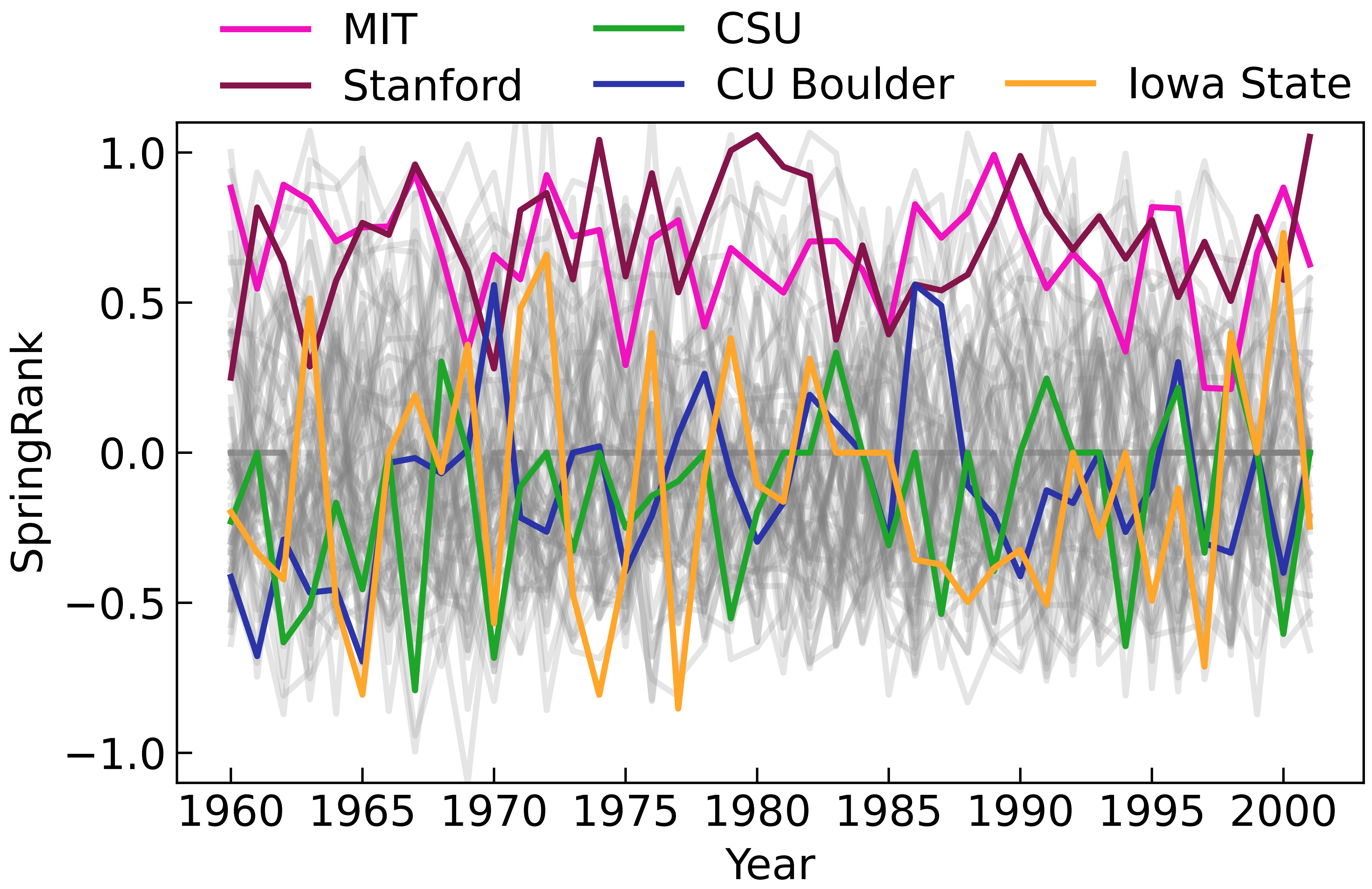
Assistant Professor

University of Colorado

PhD, Computer Science

University of Chicago

$w > 1$

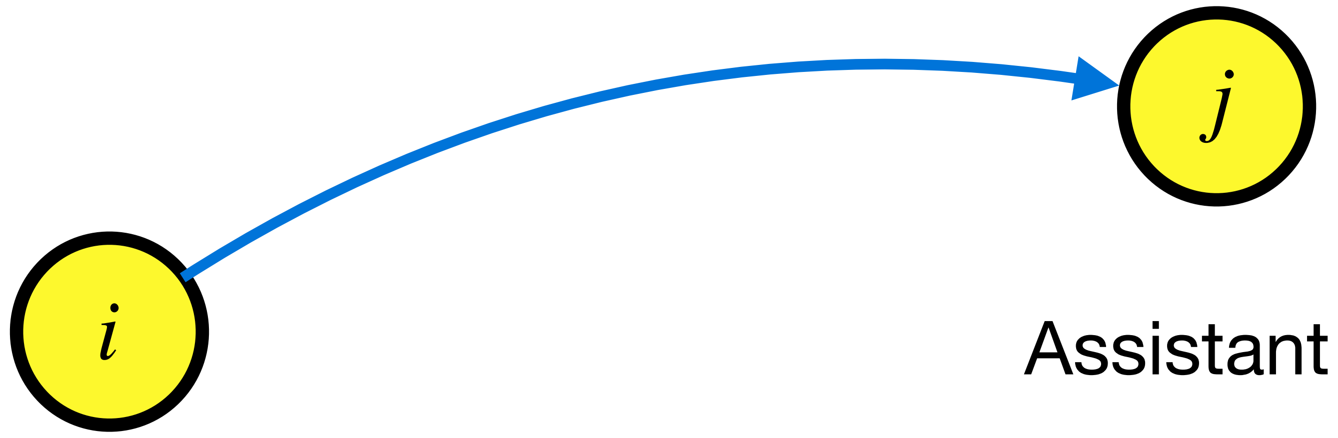


Motivation

Math PhD Exchange Dataset: Proxy for university i hiring a graduate from university j at a time near t. We view this as an endorsement by j that graduates of t are of high quality.

$t = 2023$

Tzu-Chi, Maya, & Michael graduate (🤞)



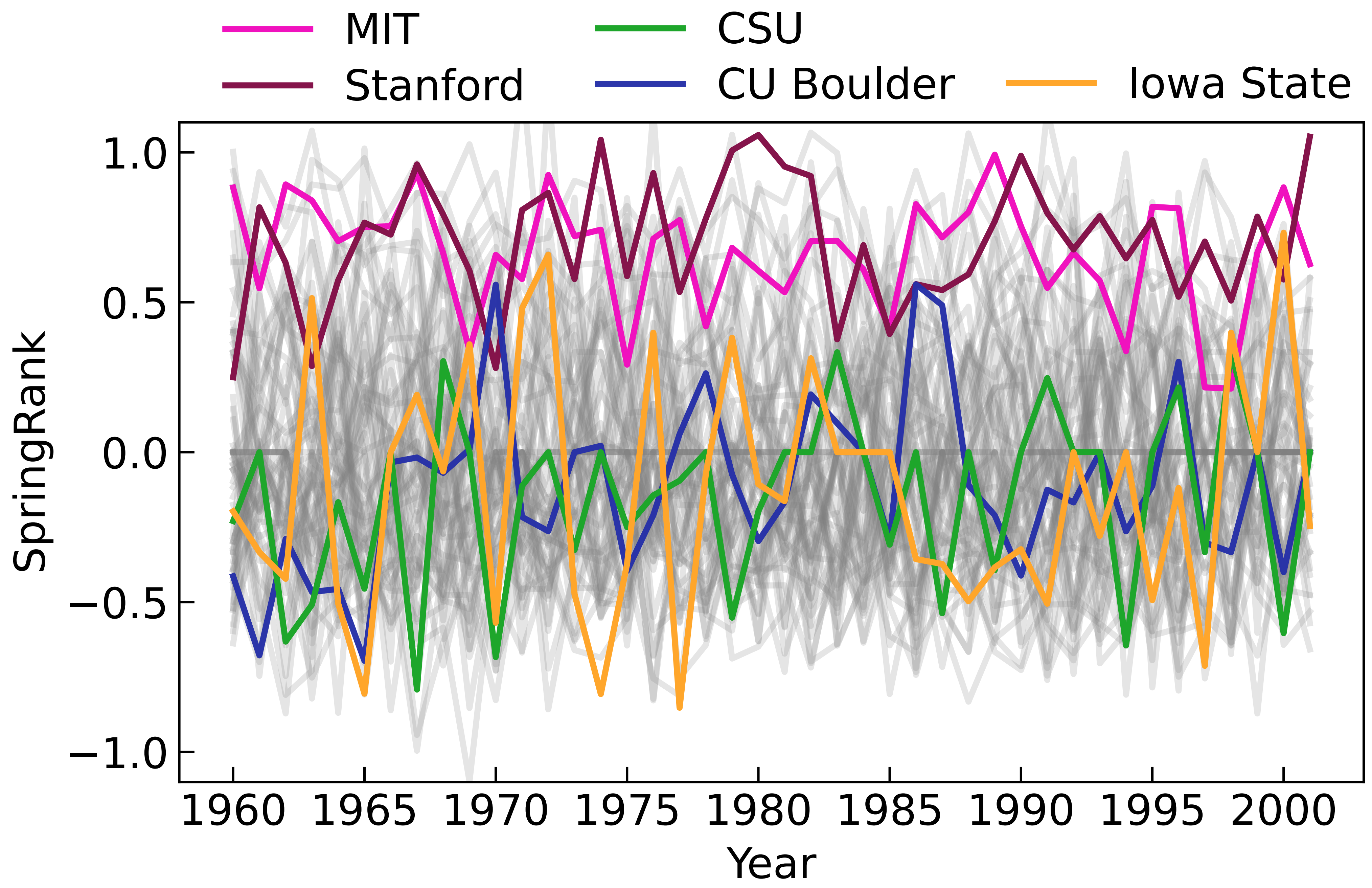
Assistant Professor

University of Colorado

PhD, Computer Science

University of Chicago

$w > 1$



Problem: Fluctuating rankings!

Idea: Optimization with Sparsity-Inducing Penalties

Sparsity through the $\ell 1$ norm: a number of coefficients of $\mathbf{w}^*$, depending on the strength of the regularization, will be exactly equal to zero.

For details, see Bach *et al.*, “Optimization with Sparsity-Inducing Penalties ” (2011).

Idea: Optimization with Sparsity-Inducing Penalties

Sparsity through the ℓ1 norm: a number of coefficients of $\mathbf{w}^*$, depending on the strength of the regularization, will be exactly equal to zero.

Sparsity through the ℓ1/ℓq norm (group norm): sometimes coefficients of $\mathbf{w}^*$ are naturally partitioned in subsets, or groups, of variables. Through regularization, some groups of coefficients will be exactly equal to zero.

For details, see Bach *et al.*, “Optimization with Sparsity-Inducing Penalties ” (2011).

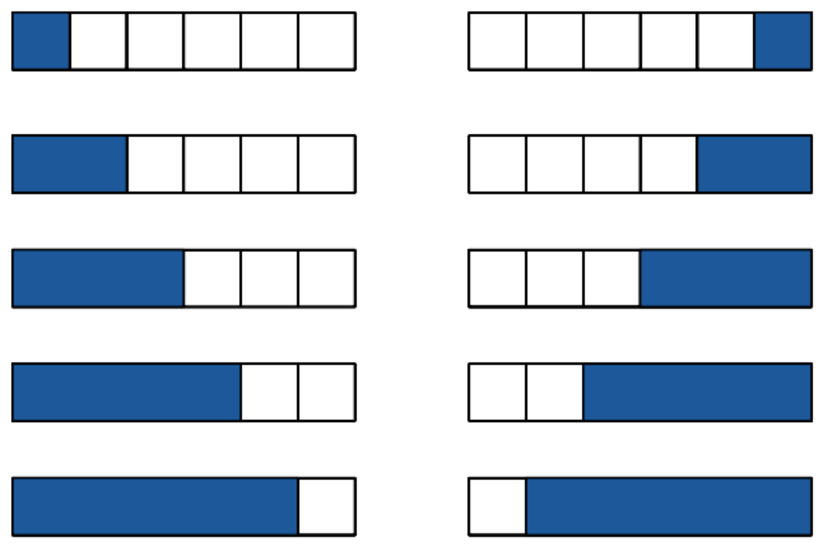
Idea: Optimization with Sparsity-Inducing Penalties

Sparsity through the $\ell 1$ norm: a number of coefficients of $\mathbf{w}^*$, depending on the strength of the regularization, will be exactly equal to zero.

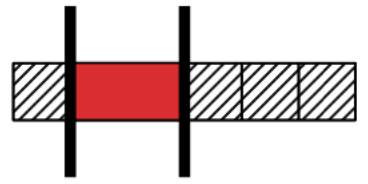
Sparsity through the $\ell 1/\ell q$ norm (group norm): sometimes coefficients of $\mathbf{w}^*$ are naturally partitioned in subsets, or groups, of variables. Through regularization, some groups of coefficients will be exactly equal to zero.

Sparsity through overlapping groups: groups can even overlap, resulting in more exotic structures to be selected.

example group setup



example coefficients selected

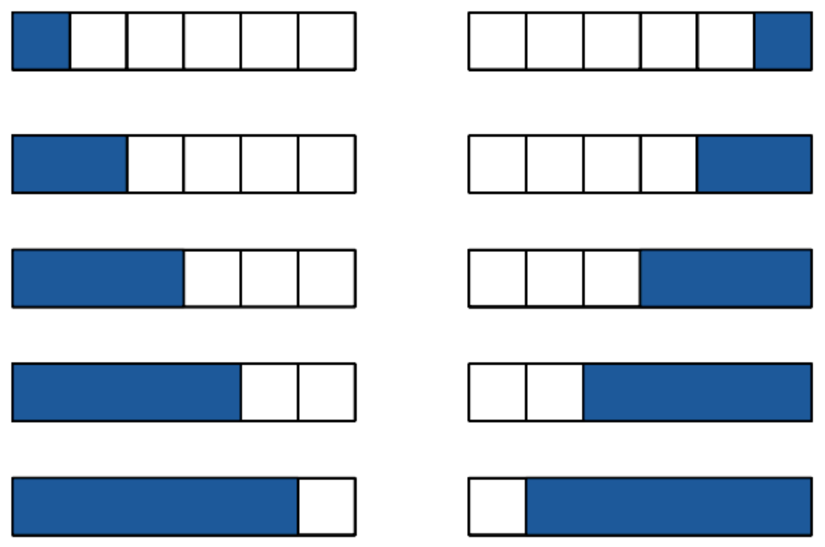


Idea: Optimization with Sparsity-Inducing Penalties

Sparsity through the ℓ1 norm: a number of coefficients of $\mathbf{w}^*$, depending on the strength of the regularization, will be exactly equal to zero.

Sparsity through the ℓ1/ℓq norm (group norm): sometimes coefficients of $\mathbf{w}^*$ are naturally partitioned in subsets, or groups, of variables. Through regularization, some groups of coefficients will be exactly equal to zero.

example group setup

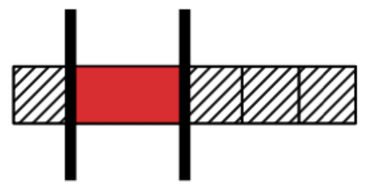


Sparsity through overlapping groups: groups can even overlap, resulting in more exotic structures to be selected.

Examples. Regression, graphical modeling, de-noising, dictionary learning, etc.

(common theme: convex optimization!)

example coefficients selected



For details, see Bach *et al.*, “Optimization with Sparsity-Inducing Penalties ” (2011).

Specific regularizations (1/3)

Regularize time dependency: apply $\ell 1$ or $\ell 2$ norm to the rank gradient in adjacent time window

$$\Omega_1(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_1$$

$$\Omega_2(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_2^2$$

Specific regularizations (1/3)

Regularize time dependency: apply $\ell 1$ or $\ell 2$ norm to the rank gradient in adjacent time window

$$\Omega_1(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_1$$

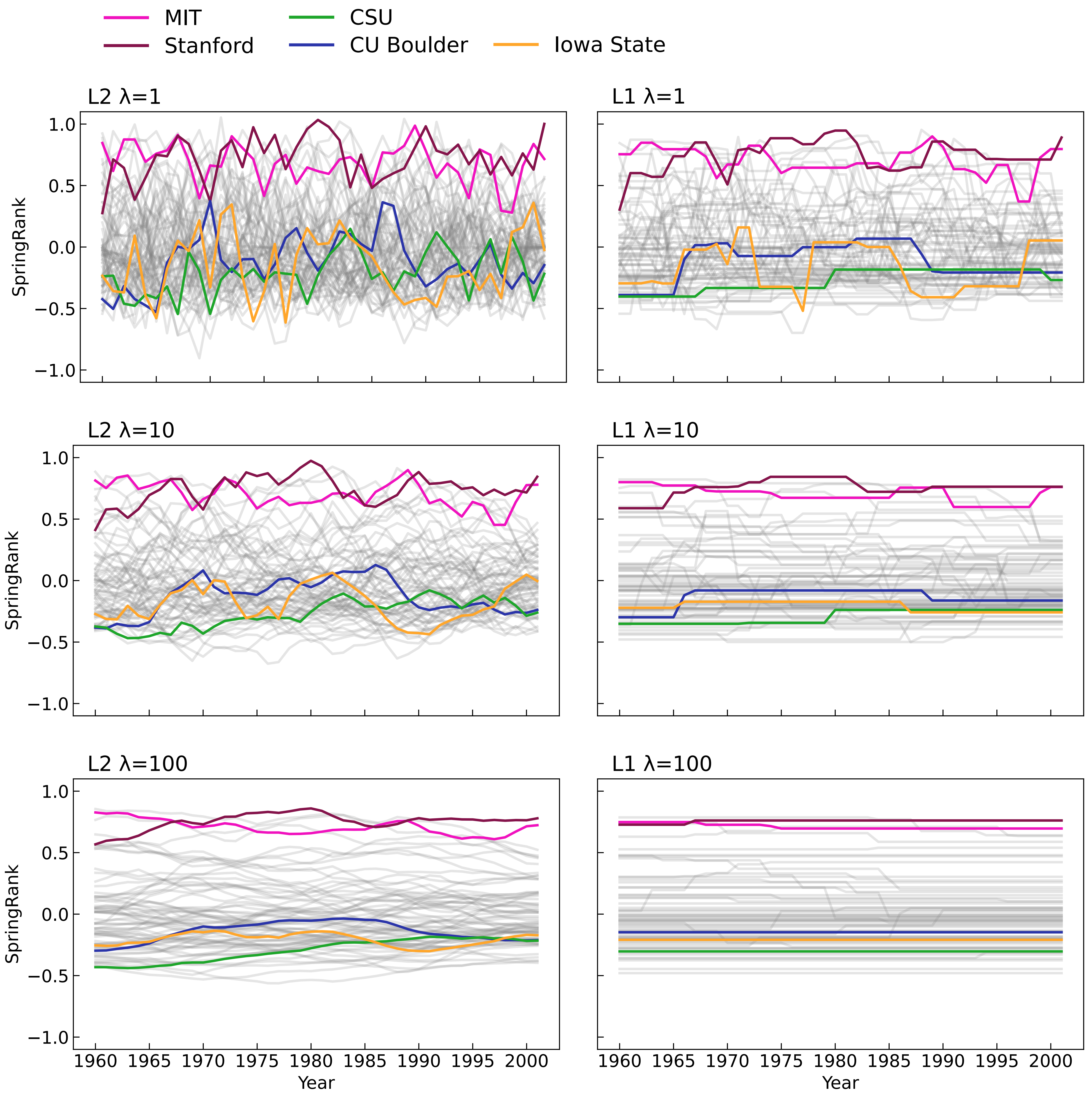
$$\Omega_2(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_2^2$$

$\ell 2$ encourages $\boldsymbol{s}_t \approx \boldsymbol{s}_{t-1}$

$\ell 1$ encourages $\boldsymbol{s}_t = \boldsymbol{s}_{t-1}$ (sparsity in ranking difference)

Result (1/3): Ranking over time

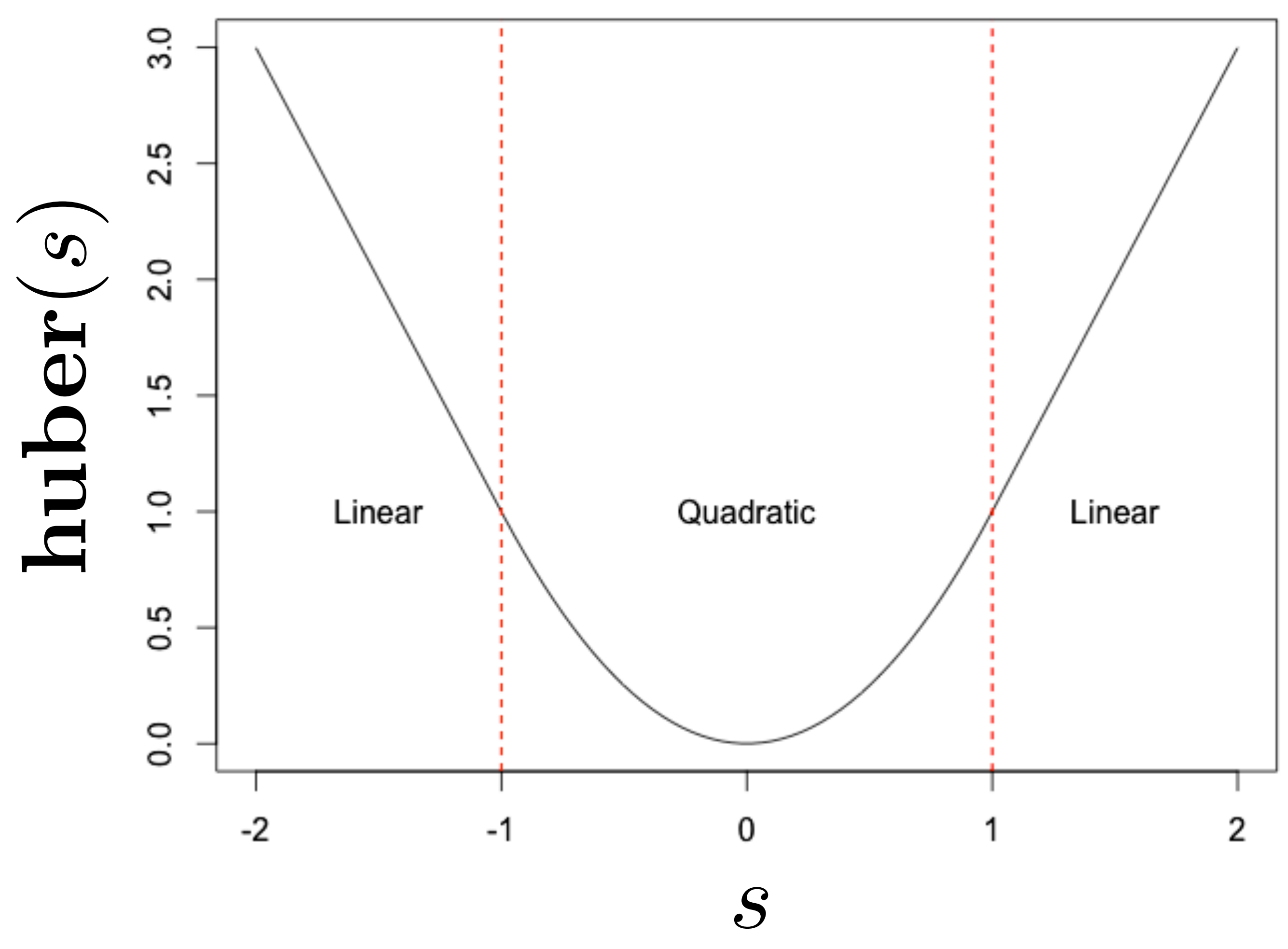
$$\underset{s}{\text{minimize}}\, \mathcal{L}(\boldsymbol{s}) + \alpha\|\boldsymbol{s}\|_2^2 + \lambda\Omega(L\boldsymbol{s})$$



Benefits. Reduce the S/N-ratio to reveal the trends (through $\ell 2$), and possibly also reveal “specific events” (through $\ell 1$).

For smaller schools (lacking data), we can assign non-zero ranks.

Specific regularizations (2/3)



Robust to outliers: substitute the quadratic loss with the Huber loss

$$\mathbf{huber}(s) = \begin{cases} (1/2)s^2 & |s| \leq 1 \\ |s| - (1/2) & |s| > 1 \end{cases} .$$

Regularize time dependency: apply $\ell 1$ or $\ell 2$ norm to the rank gradient in adjacent time window

$$\Omega_1(s) := \sum_t \|s_t - s_{t-1}\|_1$$

$$\Omega_2(s) := \sum_t \|s_t - s_{t-1}\|_2^2$$

Result (2/3): Robust ranking

$$\underset{s}{\text{minimize}}\,\mathcal{L}(\boldsymbol{s}) + \alpha\|\boldsymbol{s}\|_2^2 + \lambda\Omega(L\boldsymbol{s})$$

Error bar averages over 100 random nodes.

Result (2/3): Robust ranking

$$\underset{s}{\text{minimize}}\, \mathcal{L}(\boldsymbol{s}) + \alpha \|\boldsymbol{s}\|_2^2 + \lambda \Omega(L\boldsymbol{s})$$

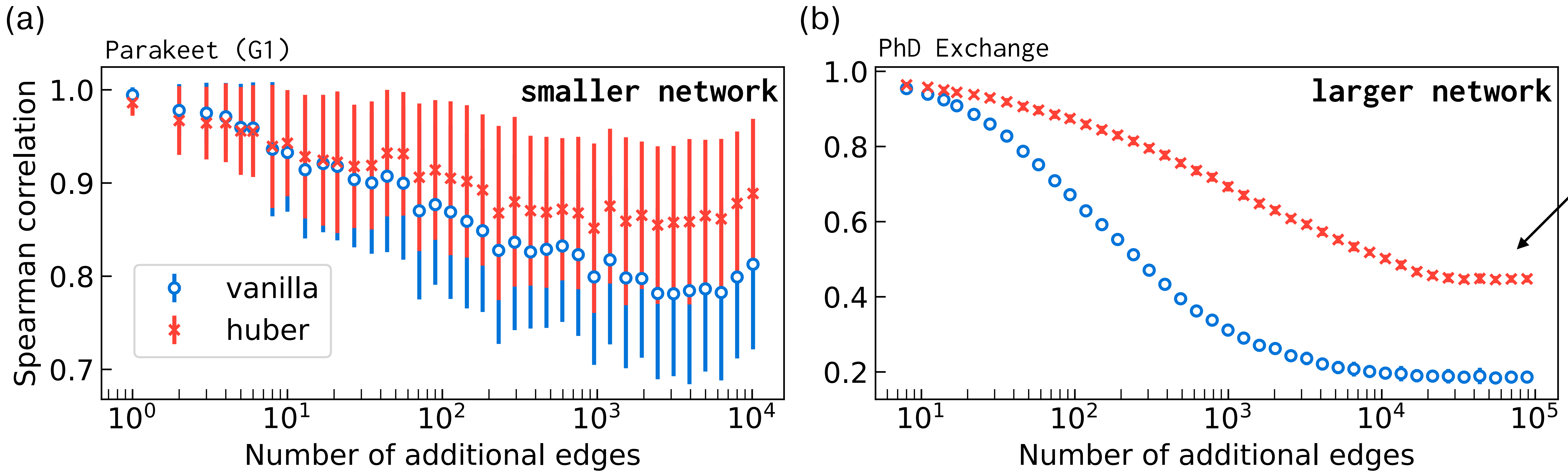
"Single-point mutation" style noise: A random node receives edges $|E|$ pointed from other nodes.

Error bar averages over 100 random nodes.

Result (2/3): Robust ranking

$$\underset{s}{\text{minimize}}\, \mathcal{L}(\boldsymbol{s}) + \alpha\|\boldsymbol{s}\|_2^2 + \lambda\Omega(L\boldsymbol{s})$$

“Single-point mutation” style noise: A random node receives edges $|E|$ pointed from other nodes.



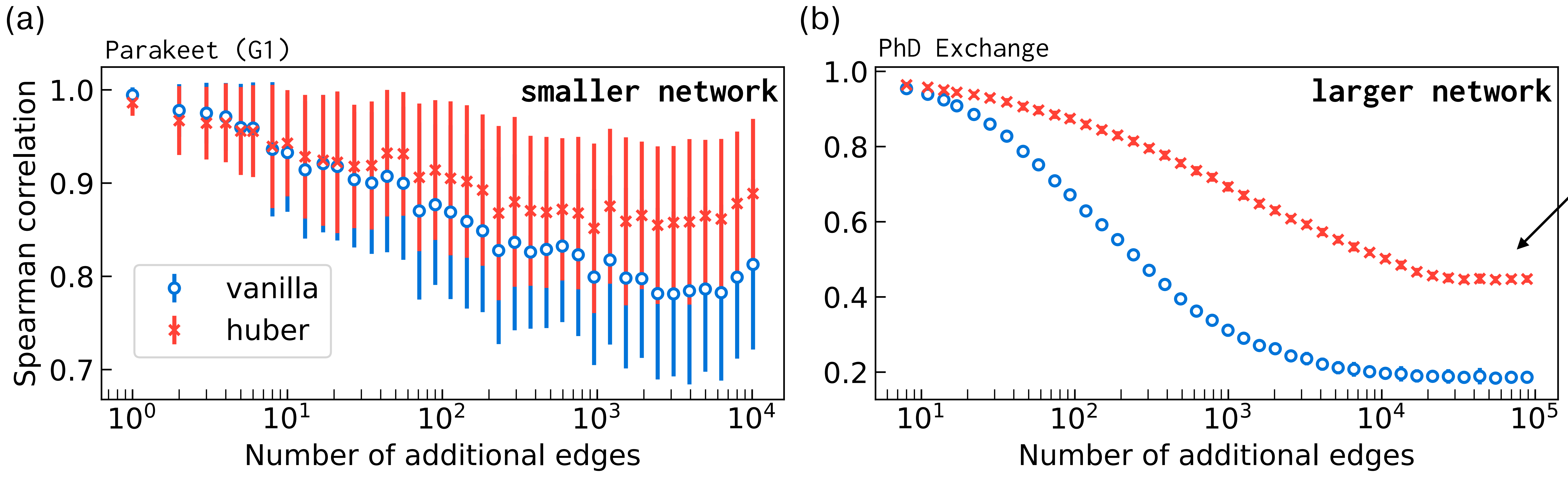
Higher correlation implies ranking information is retained

Error bar averages over 100 random nodes.

Result (2/3): Robust ranking

$$\underset{s}{\text{minimize}}\, \mathcal{L}(\boldsymbol{s}) + \alpha \|\boldsymbol{s}\|_2^2 + \lambda \Omega(L\boldsymbol{s})$$

“Single-point mutation” style noise: A random node receives edges $|E|$ pointed from other nodes.



Higher correlation implies ranking information is retained

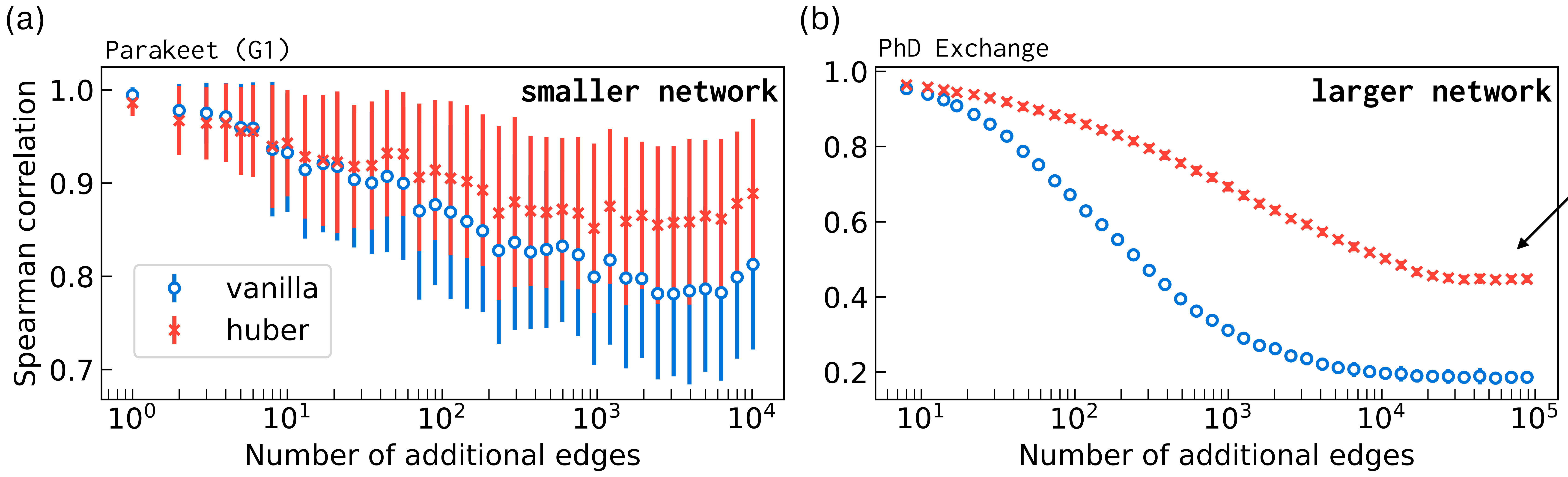
Other forms of noise: Randomly added/removed edges

Error bar averages over 100 random nodes.

Result (2/3): Robust ranking

$$\underset{s}{\text{minimize}}\, \mathcal{L}(\boldsymbol{s}) + \alpha\|\boldsymbol{s}\|_2^2 + \lambda\Omega(L\boldsymbol{s})$$

“Single-point mutation” style noise: A random node receives edges $|E|$ pointed from other nodes.



Higher correlation implies ranking information is retained

Other forms of noise: Randomly added/removed edges

Benefits: Huber loss is less sensitive to outliers in the data.

Error bar averages over 100 random nodes.

Solution method

$$\underset{s\in\mathbb{R}^N}{\text{minimize}}\ \underbrace{\frac{1}{2}\sum_{ij} A_{ij}(s_i - s_j - 1)^2}_{\mathcal{L}_{\text{quad}}} + \alpha \sum_{i=1}^{N} s_i^2$$

Generic form: two regularization functions, one makes the objective strongly convex (easier to solve), the other encourages certain structure.

$$\underset{s}{\text{minimize}}\, \mathcal{L}(\boldsymbol{s}) + \alpha\|\boldsymbol{s}\|_2^2 + \lambda\Omega(L\boldsymbol{s})$$

$$\textbf{huber}(s) = \begin{cases} (1/2)s^2 & |s| \le 1 \\ |s| - (1/2) & |s| > 1\ . \end{cases}$$

$$\Omega_1(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_1$$

$$\Omega_2(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_2^2$$

Solution method

$$\underset{s\in\mathbb{R}^N}{\text{minimize}} \underbrace{\frac{1}{2}\sum_{ij} A_{ij}(s_i - s_j - 1)^2}_{\mathcal{L}_{\text{quad}}} + \alpha\sum_{i=1}^{N} s_i^2$$

Generic form: two regularization functions, one makes the objective strongly convex (easier to solve), the other encourages certain structure.

$$\underset{s}{\text{minimize}}\, \mathcal{L}(\boldsymbol{s}) + \alpha\|\boldsymbol{s}\|_2^2 + \lambda\Omega(L\boldsymbol{s})$$

$$\textbf{huber}(s) = \begin{cases} (1/2)s^2 & |s| \le 1 \\ |s| - (1/2) & |s| > 1\ . \end{cases}$$

$$\Omega_1(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_1$$

$$\Omega_2(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_2^2$$

Solution method

Generic form: two regularization functions, one makes the objective strongly convex (easier to solve), the other encourages certain structure.

$$\underset{s}{\text{minimize}}\, \mathcal{L}(\boldsymbol{s}) + \alpha \|\boldsymbol{s}\|_2^2 + \lambda \Omega(L\boldsymbol{s})$$

$$\underset{s \in \mathbb{R}^N}{\text{minimize}}\, \underbrace{\frac{1}{2} \sum_{ij} A_{ij} (s_i - s_j - 1)^2}_{\mathcal{L}_{\text{quad}}} + \alpha \sum_{i=1}^{N} s_i^2$$

$$\textbf{huber}(s) = \begin{cases} (1/2)s^2 & |s| \leq 1 \\ |s| - (1/2) & |s| > 1 \end{cases}.$$

$$\Omega_1(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_1$$

$$\Omega_2(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_2^2$$

Solution method

$$\underset{s \in \mathbb{R}^N}{\text{minimize}} \underbrace{\frac{1}{2}\sum_{ij} A_{ij}(s_i - s_j - 1)^2}_{\mathcal{L}_{\text{quad}}} + \alpha \sum_{i=1}^{N} s_i^2$$

Generic form: two regularization functions, one makes the objective strongly convex (easier to solve), the other encourages certain structure.

$$\underset{s}{\text{minimize}}\, \mathcal{L}(\boldsymbol{s}) + \alpha \|\boldsymbol{s}\|_2^2 + \lambda \Omega(L\boldsymbol{s})$$

$$\textbf{huber}(s) = \begin{cases} (1/2)s^2 & |s| \le 1 \\ |s| - (1/2) & |s| > 1 \end{cases}.$$

$$\Omega_1(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_1$$

$$\Omega_2(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_2^2$$

Solvable by existing tools. e.g., CVXPY, which uses generic solvers.

Solution method

$$\underset{s \in \mathbb{R}^N}{\text{minimize}} \underbrace{\frac{1}{2} \sum_{ij} A_{ij} (s_i - s_j - 1)^2}_{\mathcal{L}_{\text{quad}}} + \alpha \sum_{i=1}^{N} s_i^2$$

Generic form: two regularization functions, one makes the objective strongly convex (easier to solve), the other encourages certain structure.

$$\underset{s}{\text{minimize}}\, \mathcal{L}(\boldsymbol{s}) + \alpha \|\boldsymbol{s}\|_2^2 + \lambda \Omega(L\boldsymbol{s})$$

$$\textbf{huber}(s) = \begin{cases} (1/2)s^2 & |s| \le 1 \\ |s| - (1/2) & |s| > 1 \,. \end{cases}$$

$$\Omega_1(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_1$$

$$\Omega_2(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_2^2$$

Solvable by existing tools. e.g., CVXPY, which uses generic solvers.

Issue. Graph is sparse. Need to exploit sparsity to avoid memory and computational issues for large problems.

Solution method

$$\underset{s\in\mathbb{R}^N}{\text{minimize}}\ \underbrace{\frac{1}{2}\sum_{ij} A_{ij}(s_i - s_j - 1)^2}_{\mathcal{L}_{\text{quad}}} + \alpha \sum_{i=1}^{N} s_i^2$$

Generic form: two regularization functions, one makes the objective strongly convex (easier to solve), the other encourages certain structure.

$$\underset{s}{\text{minimize}}\ \mathcal{L}(\boldsymbol{s}) + \alpha\|\boldsymbol{s}\|_2^2 + \lambda\Omega(L\boldsymbol{s})$$

$$\mathbf{huber}(s) = \begin{cases} (1/2)s^2 & |s| \le 1 \\ |s| - (1/2) & |s| > 1\ . \end{cases}$$

$$\Omega_1(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_1$$

$$\Omega_2(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_2^2$$

Solvable by existing tools. e.g., CVXPY, which uses generic solvers.

Issue. Graph is sparse. Need to exploit sparsity to avoid memory and computational issues for large problems.

Krylov methods. if we can cast the objective into a sum-of-squares form, we can use a sparse iterative solver (e.g., bi-conjugate gradient stabilized method*)

* The original SpringRank paper uses this!

Solution method

$$\underbrace{\underset{s\in\mathbb{R}^N}{\text{minimize}}\ \frac{1}{2}\sum_{ij} A_{ij}(s_i - s_j - 1)^2}_{\mathcal{L}_{\text{quad}}} + \alpha\sum_{i=1}^{N} s_i^2$$

Generic form: two regularization functions, one makes the objective strongly convex (easier to solve), the other encourages certain structure.

$$\underset{s}{\text{minimize}}\ \mathcal{L}(\boldsymbol{s}) + \alpha\|\boldsymbol{s}\|_2^2 + \lambda\Omega(L\boldsymbol{s})$$

$$\mathbf{huber}(s) = \begin{cases}(1/2)s^2 & |s| \le 1\\ |s| - (1/2) & |s| > 1\ .\end{cases}$$

$$\Omega_1(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_1$$

$$\Omega_2(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_2^2$$

Solvable by existing tools. e.g., CVXPY, which uses generic solvers.

Issue. Graph is sparse. Need to exploit sparsity to avoid memory and computational issues for large problems.

Krylov methods. if we can cast the objective into a sum-of-squares form, we can use a sparse iterative solver (e.g., bi-conjugate gradient stabilized method*)

First-order methods. we may turn to generalizations of gradient decent

* The original SpringRank paper uses this!

Dual-based proximal gradient methods

$$\underset{s}{\text{minimize}}\, \mathcal{L}(\boldsymbol{s}) + \alpha \|\boldsymbol{s}\|_2^2 + \lambda \Omega(L\boldsymbol{s})$$

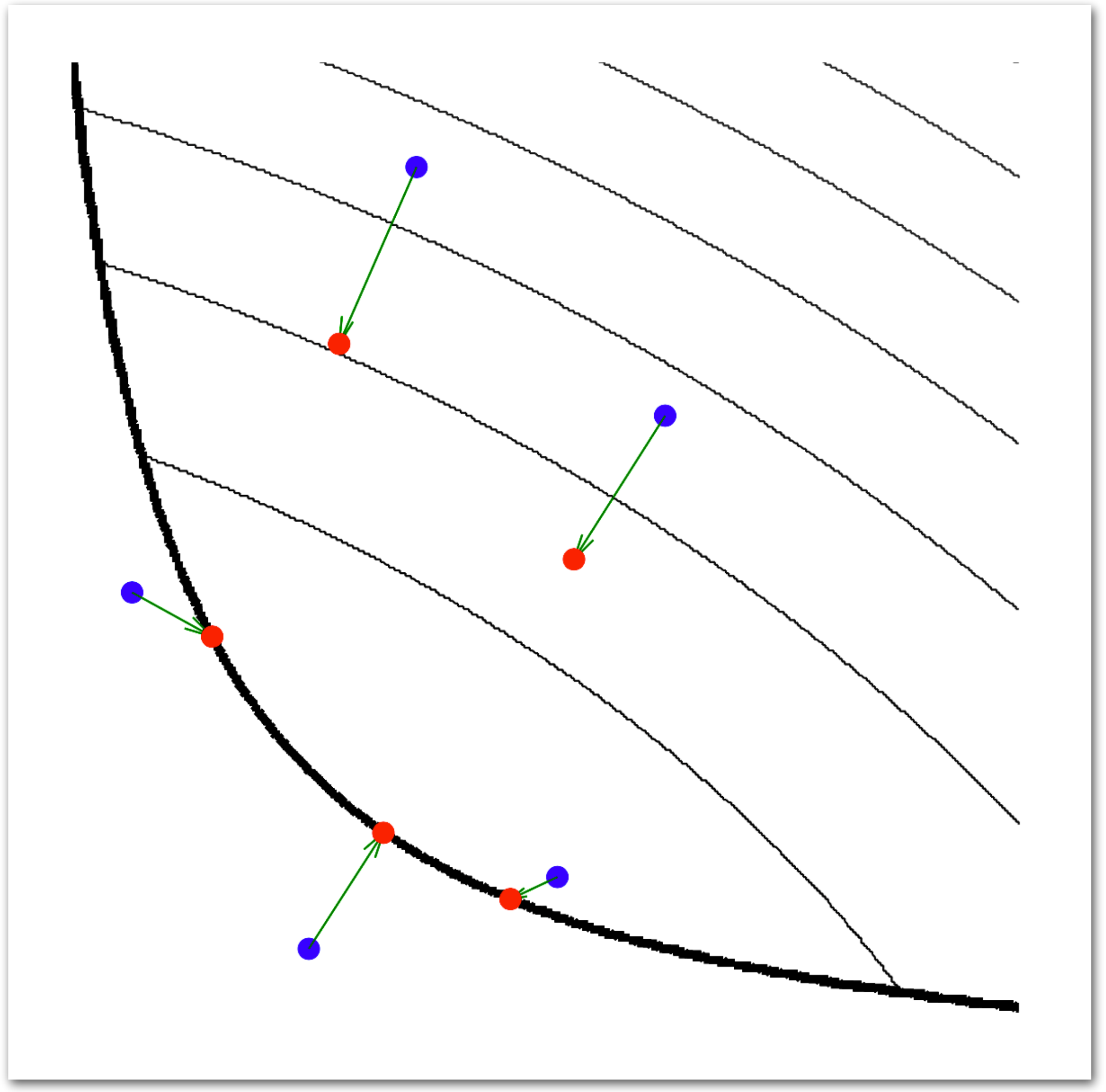
For details, see Beck, "First-Order Methods in Optimization" (2017), or Parikh and Boyd, "Proximal Algorithms" (2014)

Dual-based proximal gradient methods

$$\underset{s}{\text{minimize}}\, \mathcal{L}(\boldsymbol{s}) + \alpha\|\boldsymbol{s}\|_2^2 + \lambda\Omega(L\boldsymbol{s})$$

Composite model: requires that f has a nice gradient and g has a easy-to-compute proximal operator

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}}\{F(\boldsymbol{x}) \equiv f(\boldsymbol{x}) + g(\boldsymbol{x})\}$$



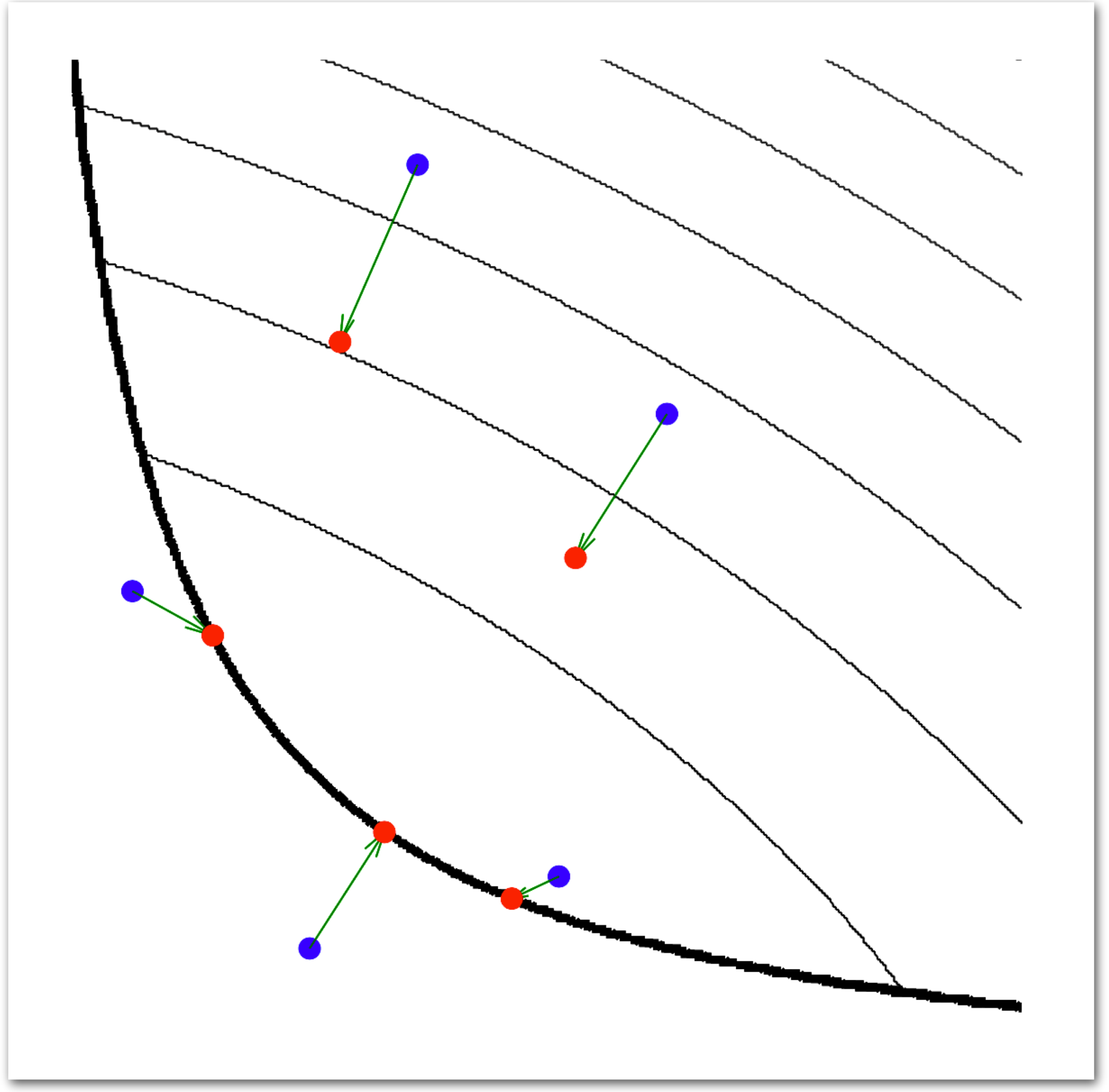
For details, see Beck, “First-Order Methods in Optimization” (2017), or Parikh and Boyd, “Proximal Algorithms” (2014)

Dual-based proximal gradient methods

$$\underset{s}{\text{minimize}}\, \mathcal{L}(\boldsymbol{s}) + \alpha\|\boldsymbol{s}\|_2^2 + \lambda\Omega(L\boldsymbol{s})$$

Composite model: requires that f has a nice gradient and g has a easy-to-compute proximal operator

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}}\{F(\boldsymbol{x}) \equiv f(\boldsymbol{x}) + g(\boldsymbol{x})\}$$



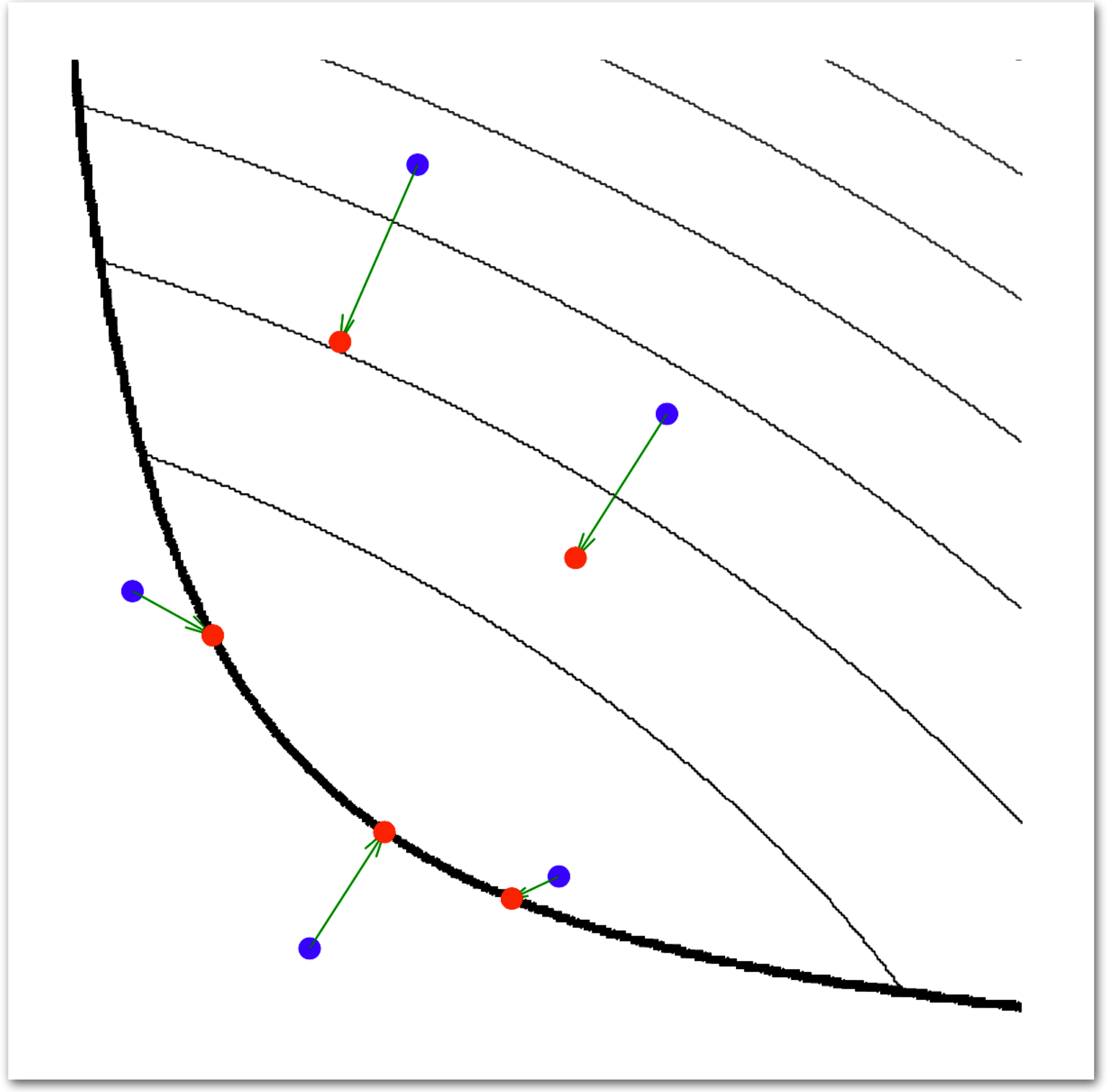
For details, see Beck, "First-Order Methods in Optimization" (2017), or Parikh and Boyd, "Proximal Algorithms" (2014)

Dual-based proximal gradient methods

$$\underset{s}{\text{minimize}}\, \mathcal{L}(\boldsymbol{s}) + \alpha\|\boldsymbol{s}\|_2^2 + \lambda\Omega(L\boldsymbol{s})$$

Composite model: requires that f has a nice gradient and g has a easy-to-compute proximal operator

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}}\{F(\boldsymbol{x}) \equiv f(\boldsymbol{x}) + g(\boldsymbol{x})\}$$



Dual problem: easier to solve, as the linear operator L goes to the part that relies on the gradient.

$$\text{(Primal)} \qquad \underset{\boldsymbol{s}\in\mathbb{R}^N}{\text{minimize}}\, f(\boldsymbol{s}) + g(L\boldsymbol{s})$$

$$\text{(Dual)} \qquad \underset{\boldsymbol{v}\in\mathcal{K}}{\text{minimize}}\, f^*(-L^*\boldsymbol{v}) + g^*(\boldsymbol{v})$$

$$f^*(s) = \max_x \langle s, x\rangle - f(x)$$

The Proximal Gradient Method

Initialization: pick $\mathbf{x}^0 \in \text{int}(\text{dom}(f))$.
General step: for any $k = 0, 1, 2, \ldots$ execute the following steps:

(a) pick $L_k > 0$;

(b) set $\mathbf{x}^{k+1} = \text{prox}_{\frac{1}{L_k}g}\left(\mathbf{x}^k - \frac{1}{L_k}\nabla f(\mathbf{x}^k)\right)$.

For details, see Beck, “First-Order Methods in Optimization” (2017), or Parikh and Boyd, “Proximal Algorithms” (2014)

Specific regularizations

Robust to outliers: substitute the quadratic loss with the Huber loss

$$\text{huber}(s) = \begin{cases} (1/2)s^2 & |s| \le 1 \\ |s| - (1/2) & |s| > 1 \,. \end{cases}$$

Regularize time dependency: apply $\ell 1$ or $\ell 2$ norm to the rank gradient in adjacent time window

$$\Omega_1(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_1$$

$$\Omega_2(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_2^2$$

Same node category, same mean ranking: encourage “no bias across categories” unless necessary

$$\Omega_{\text{group}}(\boldsymbol{s}) := \sum_{z(i) \neq z(j)} w_{ij} |\langle s_i \rangle - \langle s_j \rangle|$$

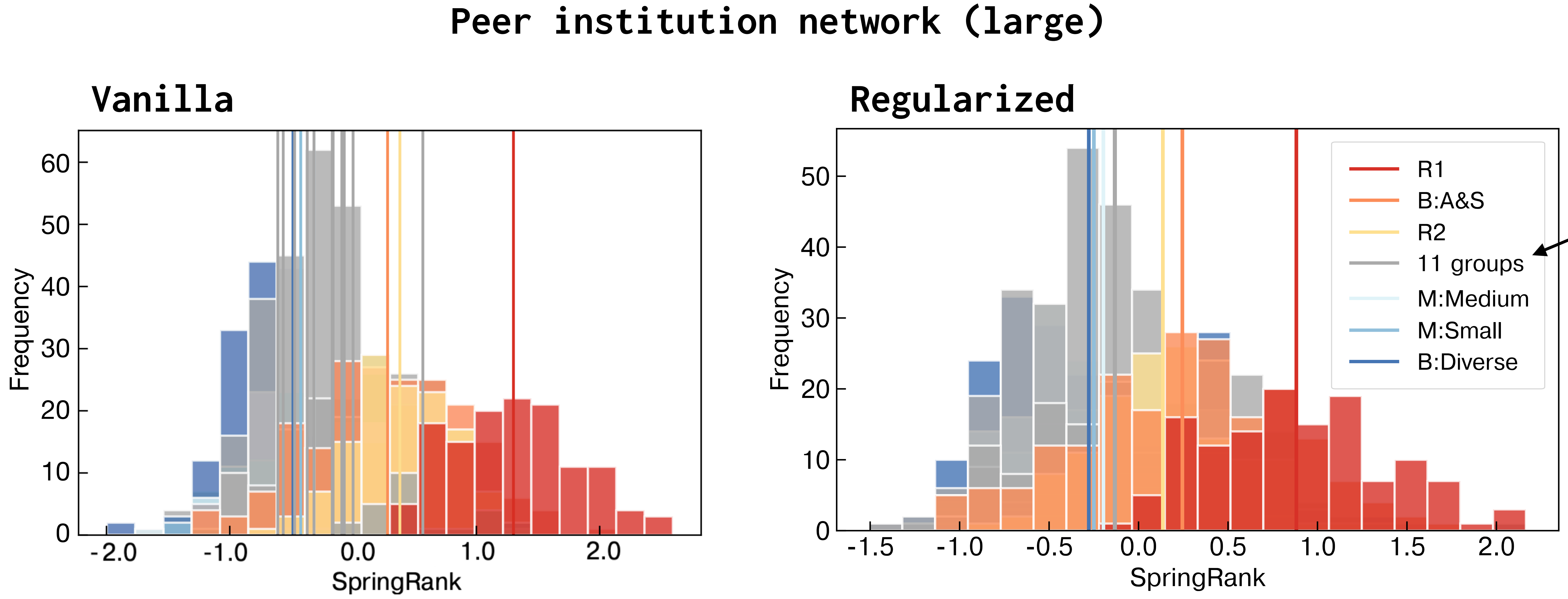
Result (3/3): Ranking to learn group-dependent phenomena

$$\underset{s}{\text{minimize}}\, \mathcal{L}(\boldsymbol{s}) + \alpha\|\boldsymbol{s}\|_2^2 + \lambda\Omega(L\boldsymbol{s})$$

We used α=0.01, λ=0 or λ=10 (regularized).

Result (3/3): Ranking to learn group-dependent phenomena

$$\underset{s}{\text{minimize}}\, \mathcal{L}(\boldsymbol{s}) + \alpha \|\boldsymbol{s}\|_2^2 + \lambda \Omega(L\boldsymbol{s})$$



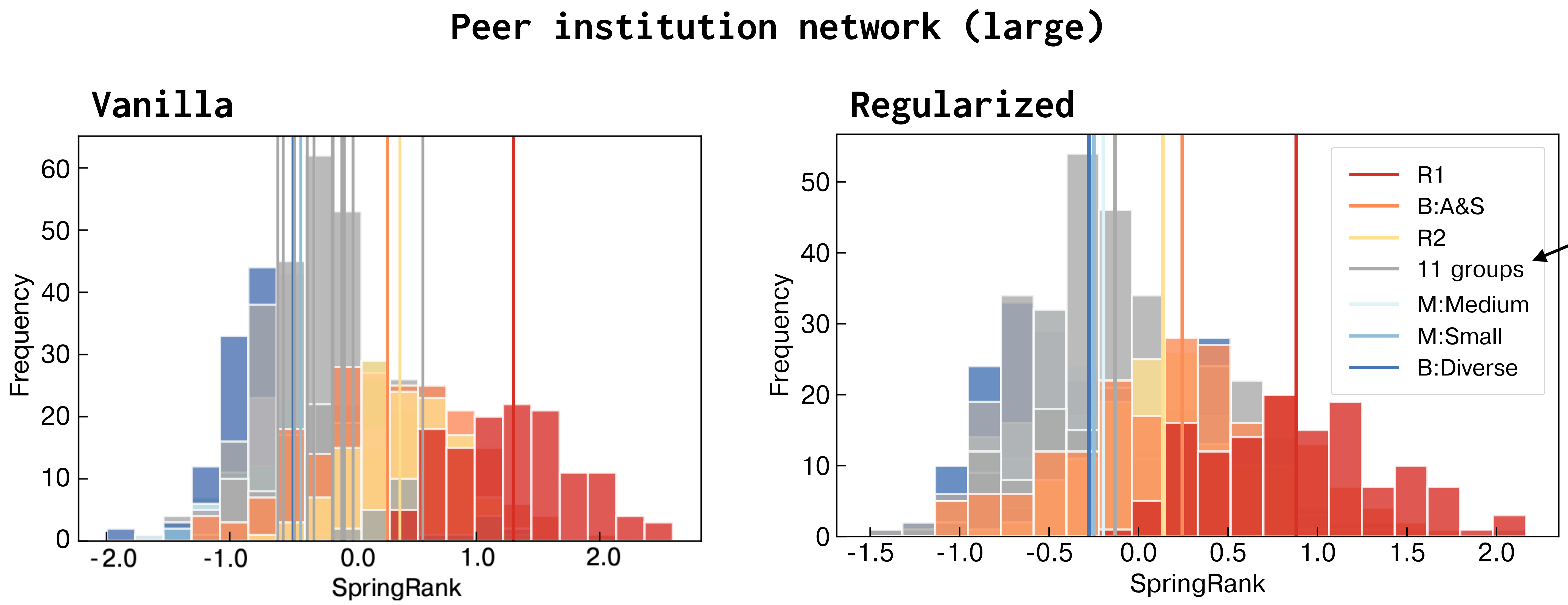
Containing *many* Carnegie Classifications.

Benefits: By encouraging nodes of different groups to have the same, mean ranking, we *select* the groups that are special, or biased.

We used α=0.01, λ=0 or λ=10 (regularized).

Result (3/3): Ranking to learn group-dependent phenomena

$$\underset{s}{\text{minimize}}\, \mathcal{L}(\boldsymbol{s}) + \alpha \|\boldsymbol{s}\|_2^2 + \lambda \Omega(L\boldsymbol{s})$$



Containing *many* Carnegie Classifications.

Benefits: By encouraging nodes of different groups to have the same, mean ranking, we *select* the groups that are special, or biased.

Caveats: *Is this a better model?* Not always. We use a cross-validated model to predict out-of-sample edge directions (worse) and compare their likelihoods (better) with the vanilla SpringRank. But "better" is subjective. We expect the model to outperform for smaller categories though.

We used α=0.01, λ=0 or λ=10 (regularized).

Implications

Implications

$$\underset{s\in\mathbb{R}^N}{\text{minimize}}\ \underbrace{\frac{1}{2}\sum_{ij} A_{ij}(s_i - s_j - 1)^2}_{\mathcal{L}_{\text{quad}}} + \alpha\sum_{i=1}^{N} s_i^2$$

Mix-and-match losses with regularizations allow new modeling capabilities.

$$\mathbf{huber}(s) = \begin{cases} (1/2)s^2 & |s| \le 1 \\ |s| - (1/2) & |s| > 1\ . \end{cases}$$

$$\Omega_1(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_1$$

$$\Omega_2(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_2^2$$

$$\Omega_{\text{group}}(\boldsymbol{s}) := \sum_{z(i)\neq z(j)} w_{ij}|(\langle s_i\rangle - \langle s_j\rangle)|$$

Implications

$$\underset{s \in \mathbb{R}^N}{\text{minimize}} \underbrace{\frac{1}{2} \sum_{ij} A_{ij}(s_i - s_j - 1)^2}_{\mathcal{L}_{\text{quad}}} + \alpha \sum_{i=1}^{N} s_i^2$$

Mix-and-match losses with regularizations allow new modeling capabilities.

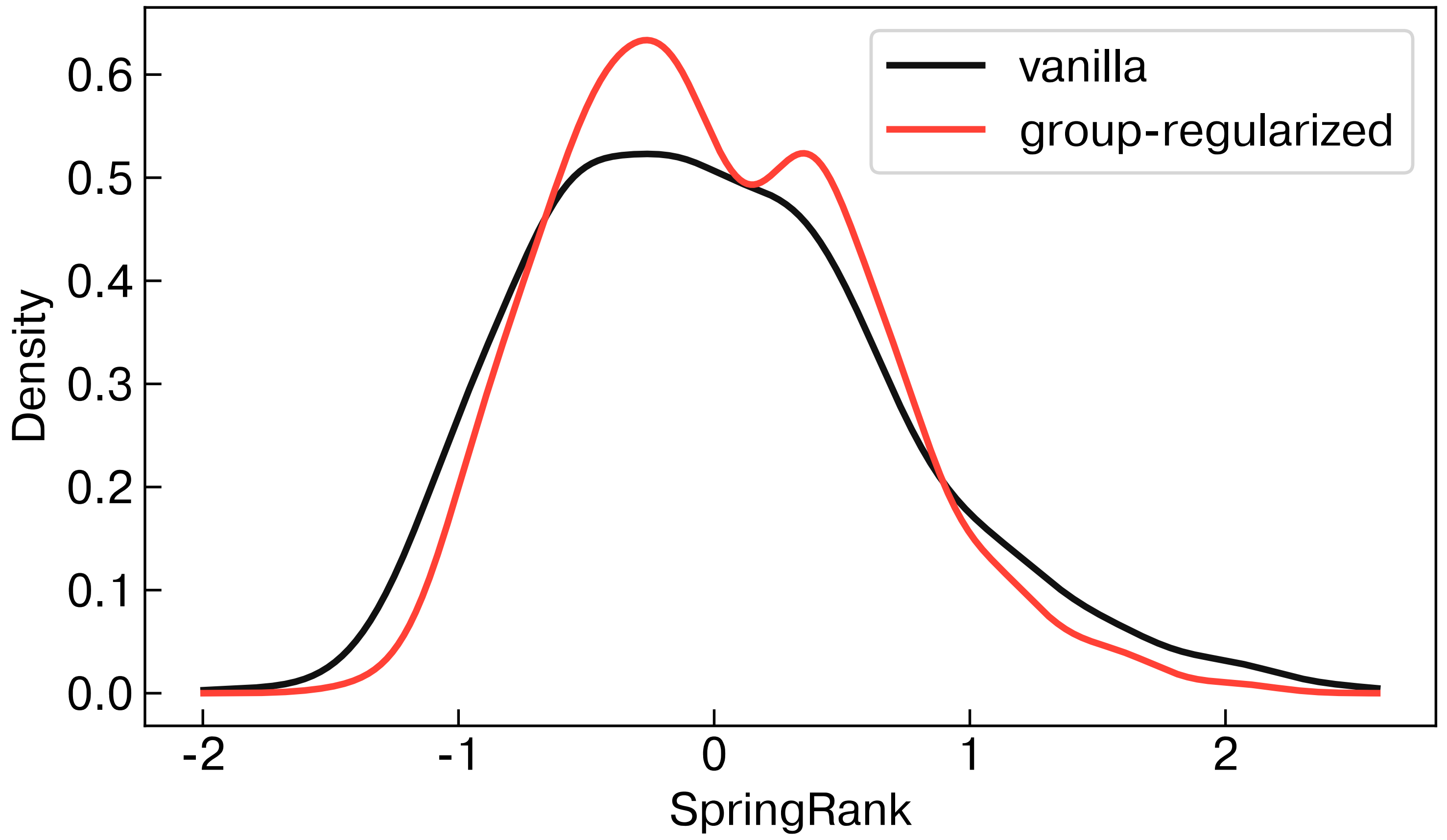
SpringRank can now select group-dependent (multi-modal) ranking distributions. Other fancier choice models do the same (e.g., contextual repeated selection model), but not conventional ones (e.g., Bradley-Terry-Luce model & vanilla SpringRank).

$$\textbf{huber}(s) = \begin{cases} (1/2)s^2 & |s| \leq 1 \\ |s| - (1/2) & |s| > 1 \;. \end{cases}$$

$$\Omega_1(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_1$$

$$\Omega_2(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_2^2$$

$$\Omega_{\text{group}}(\boldsymbol{s}) := \sum_{z(i) \neq z(j)} w_{ij} |(\langle s_i \rangle - \langle s_j \rangle)|$$



Implications

$$\underset{s\in\mathbb{R}^N}{\text{minimize}} \underbrace{\frac{1}{2}\sum_{ij} A_{ij}(s_i - s_j - 1)^2}_{\mathcal{L}_{\text{quad}}} + \alpha \sum_{i=1}^{N} s_i^2$$

Mix-and-match losses with regularizations allow new modeling capabilities.

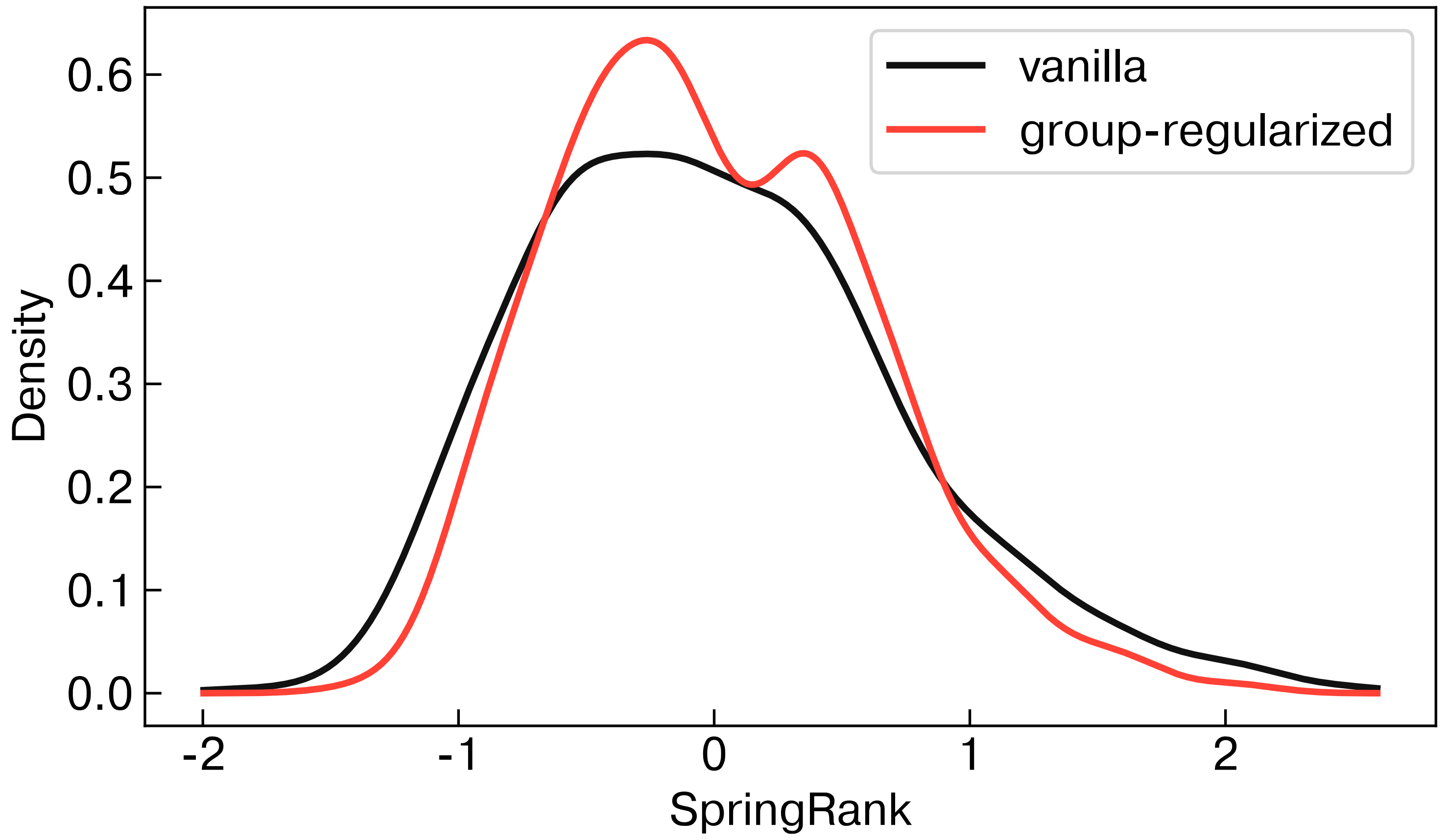
$$\mathbf{huber}(s) = \begin{cases} (1/2)s^2 & |s| \leq 1 \\ |s| - (1/2) & |s| > 1\ . \end{cases}$$

$$\Omega_1(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_1$$

$$\Omega_2(\boldsymbol{s}) := \sum_t \|\boldsymbol{s}_t - \boldsymbol{s}_{t-1}\|_2^2$$

$$\Omega_{\text{group}}(\boldsymbol{s}) := \sum_{z(i)\neq z(j)} w_{ij} |(\langle s_i \rangle - \langle s_j \rangle)|$$

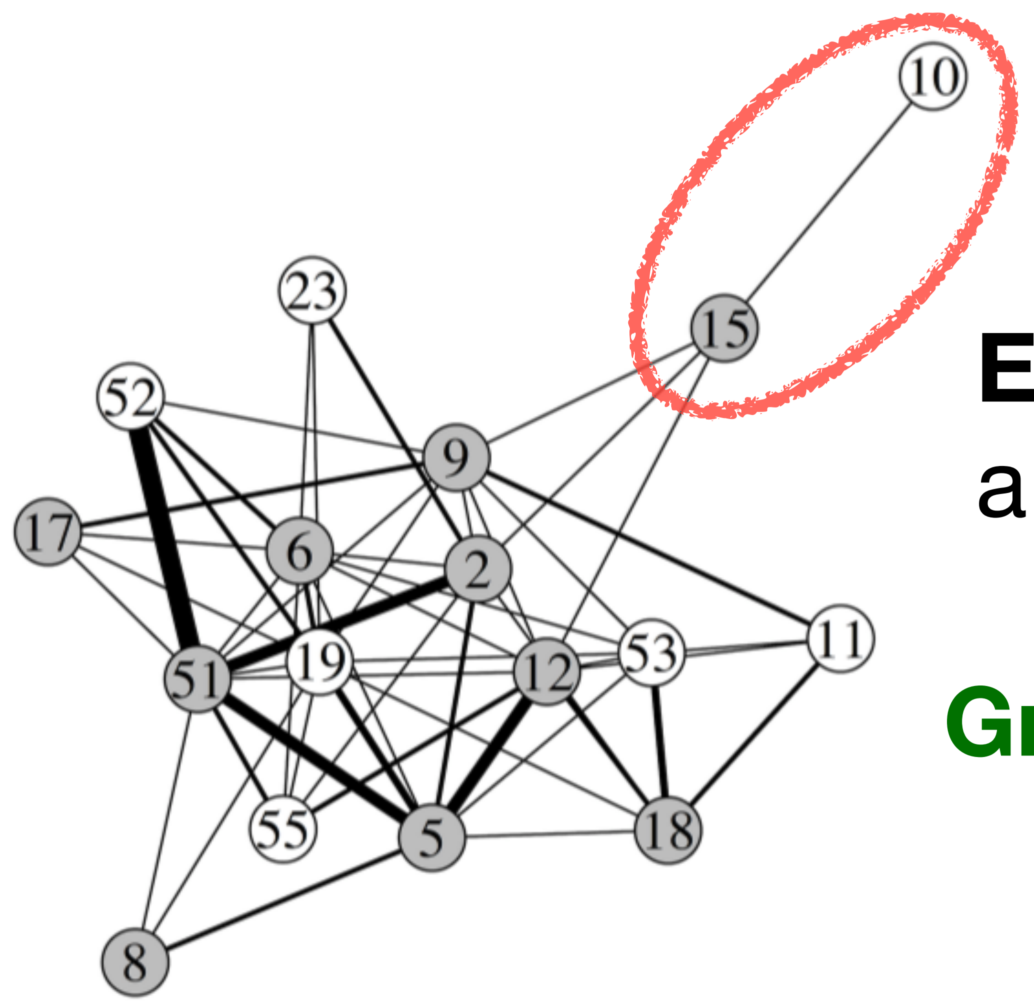
SpringRank can now select group-dependent (multi-modal) ranking distributions. Other fancier choice models do the same (e.g., contextual repeated selection model), but not conventional ones (e.g., Bradley-Terry-Luce model & vanilla SpringRank).



Limitations. No natural way to apply to numeric groups (think: how does admission rate relate to the hierarchy?); Cross-validating >2 hyperparameters is slow.

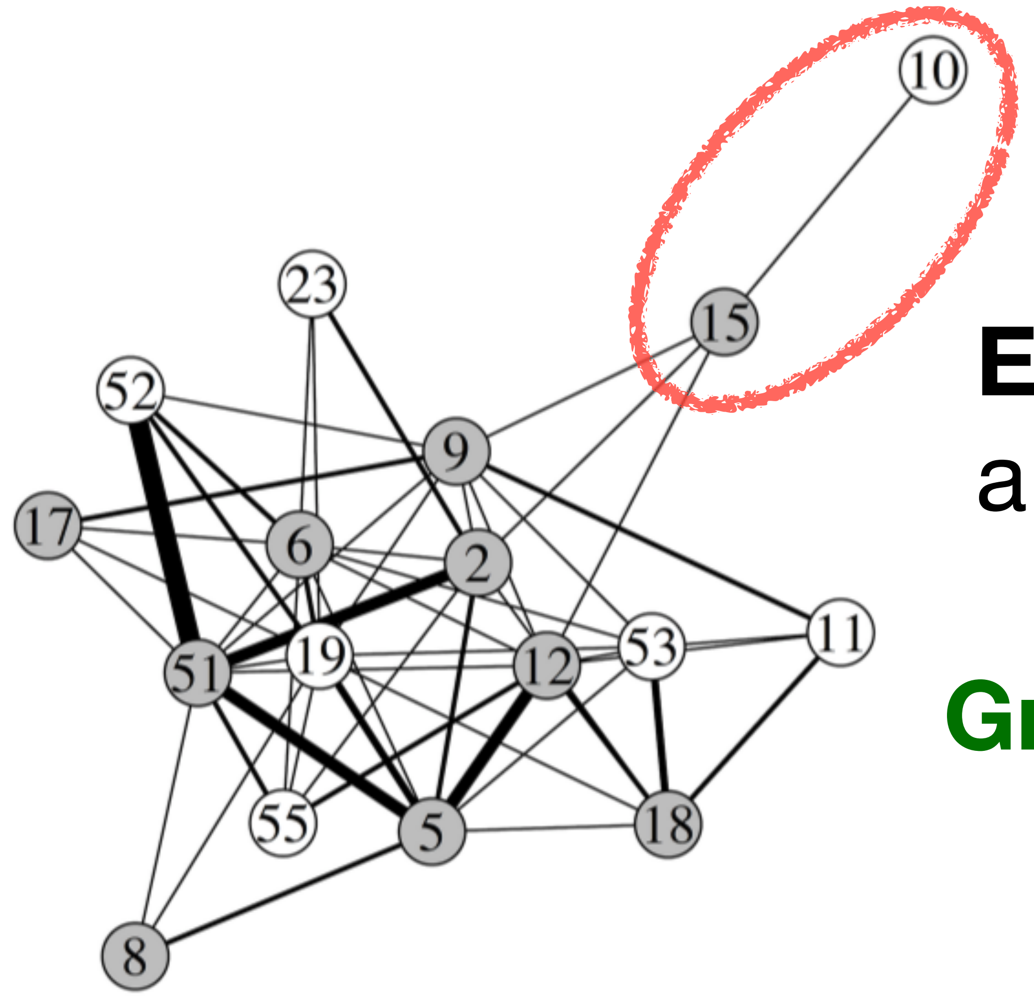
Outline

1. Context & motivation (chapter 1)
2. Aspiration of prestige in institutional peer selection (chapter 5)
3. Community detection in bipartite networks (chapter 2)
4. Regularized methods for efficient ranking (chapter 3)
5. A detour in higher-order structures (chapter 4)
6. Conclusions & future research (chapter 6)



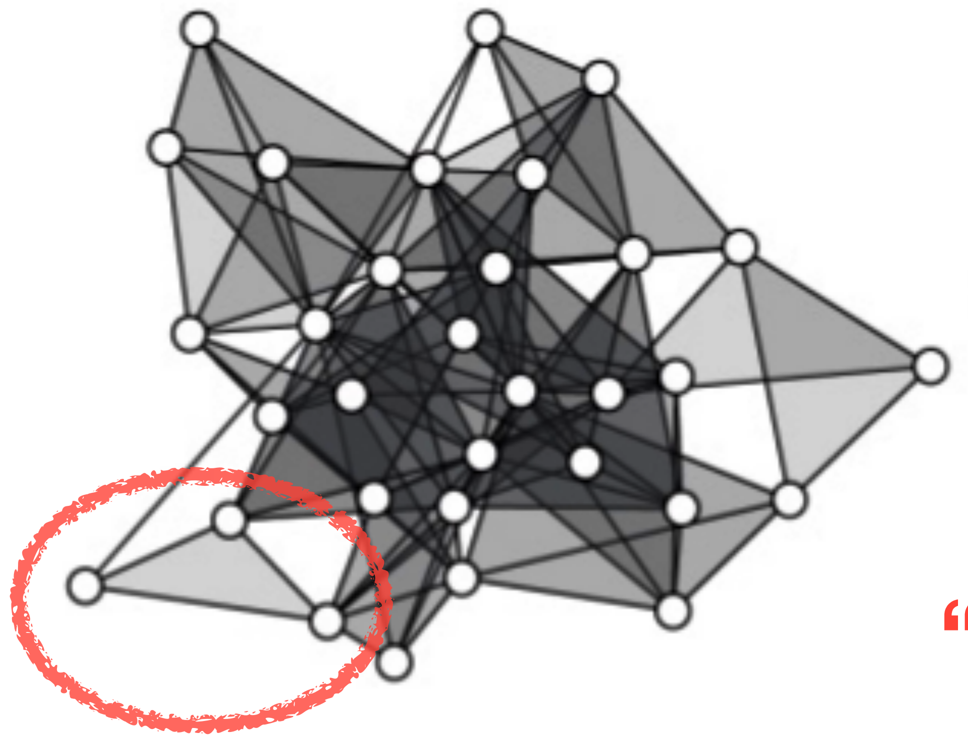
Edges. These 2 authors have written a paper together.

Graph



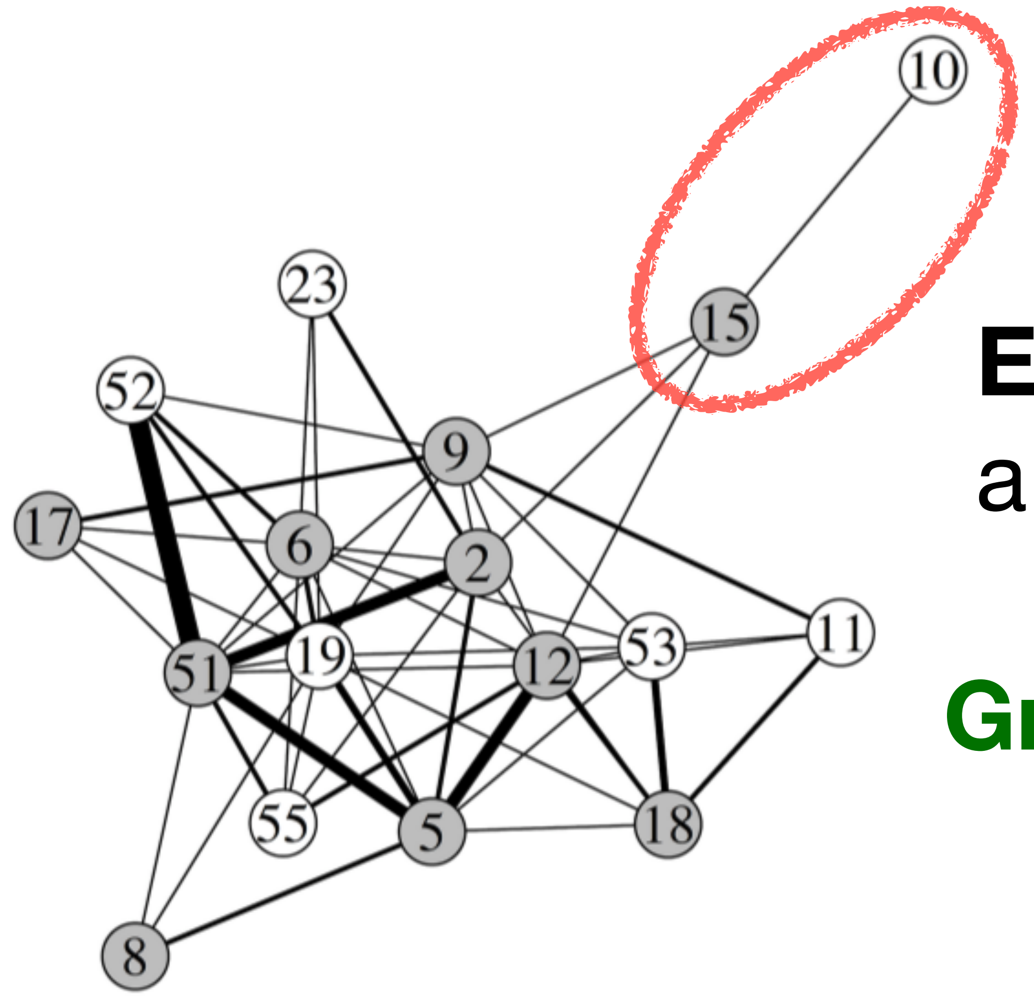
Edges. These 2 authors have written a paper together.

Graph



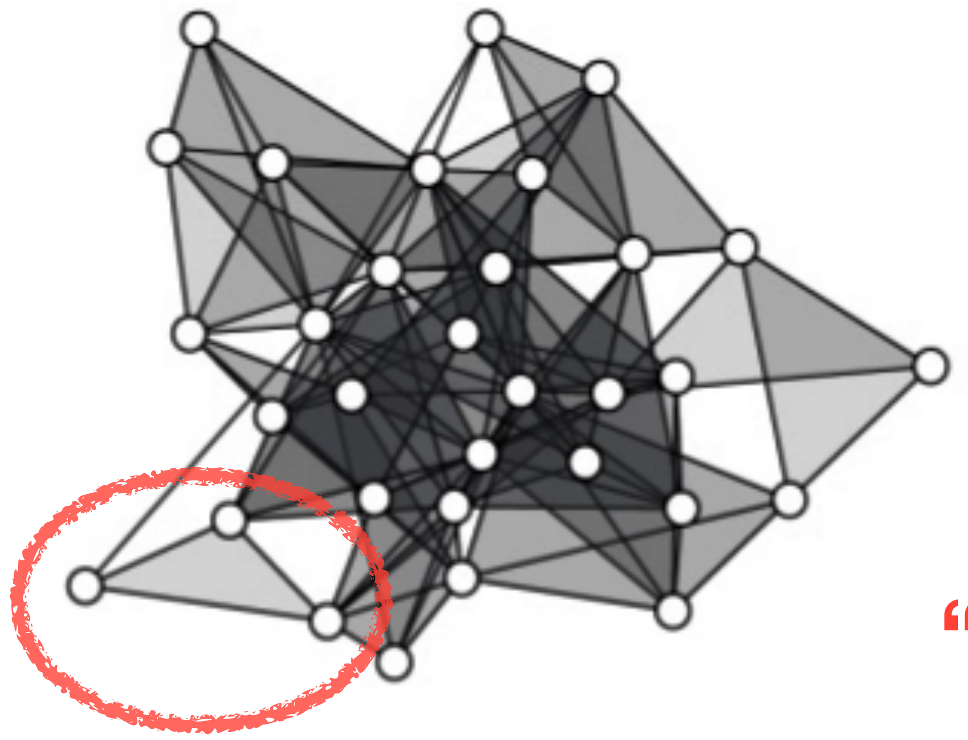
“Higher-order” graph

Edges. These 3 authors have worked together (and wrote >1 papers).



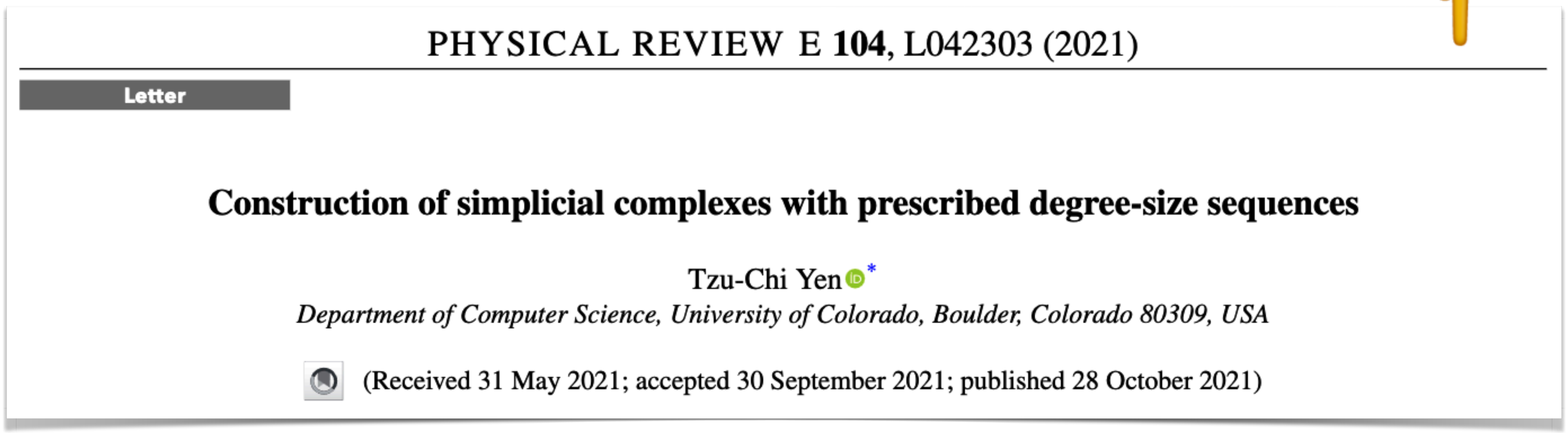
Edges. These 2 authors have written a paper together.

Graph



"Higher-order" graph

Edges. These 3 authors have worked together (and wrote >1 papers).



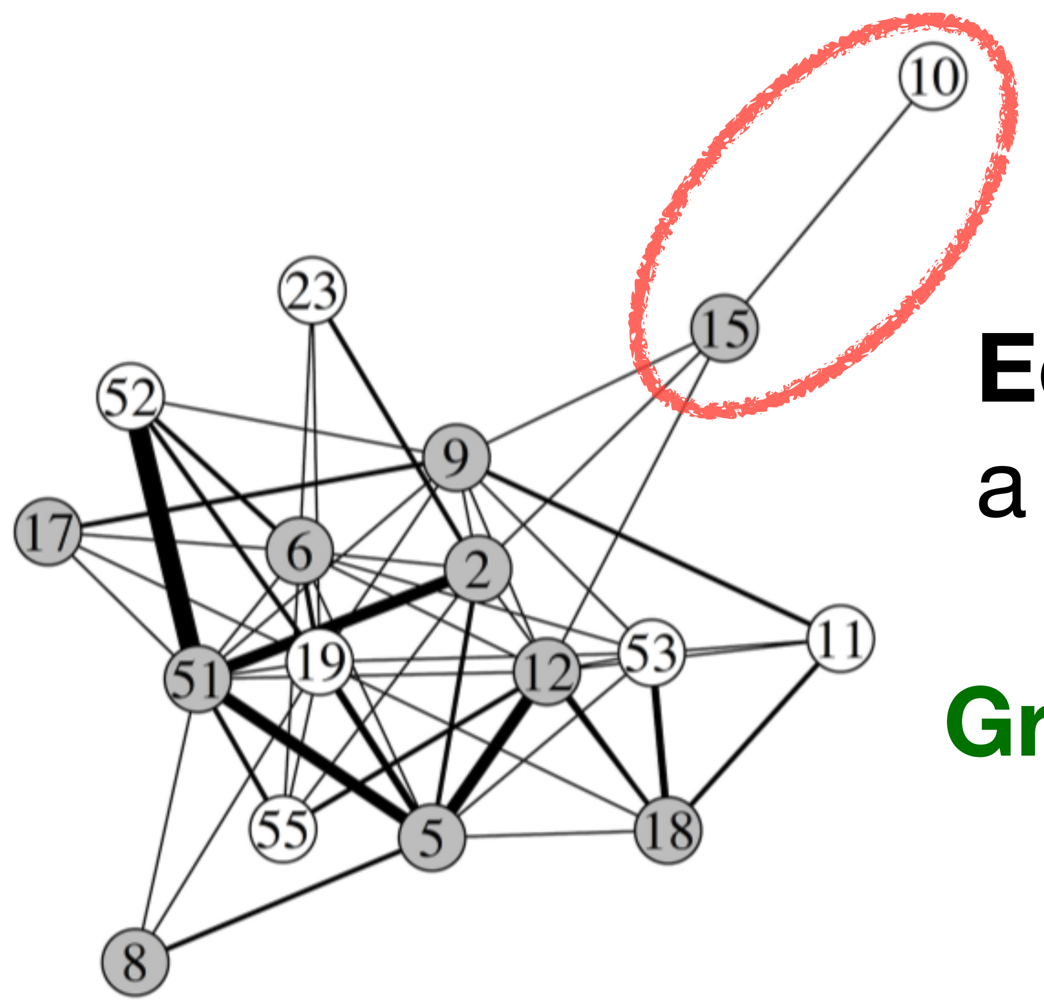
PHYSICAL REVIEW E **104**, L042303 (2021)

Letter

Construction of simplicial complexes with prescribed degree-size sequences

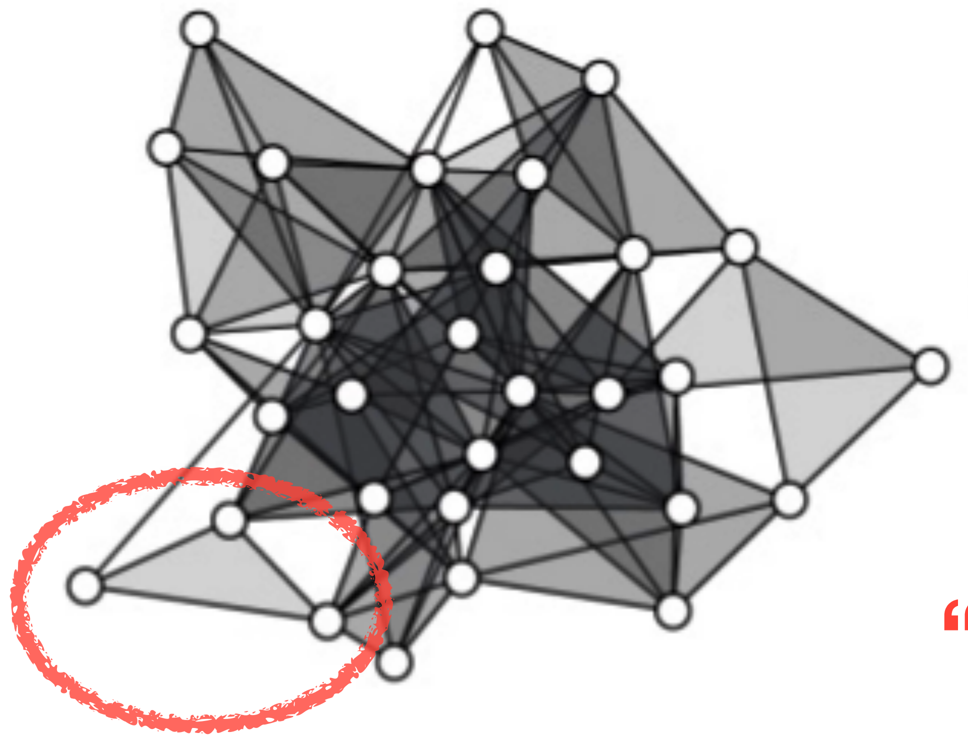
Tzu-Chi Yen *
Department of Computer Science, University of Colorado, Boulder, Colorado 80309, USA

(Received 31 May 2021; accepted 30 September 2021; published 28 October 2021)



Edges. These 2 authors have written a paper together.

Graph



"Higher-order" graph

Edges. These 3 authors have worked together (and wrote >1 papers).

NP-completeness. Do not give up if a problem is NP-hard, sometimes a simple algorithm exists!

PHYSICAL REVIEW E **104**, L042303 (2021)

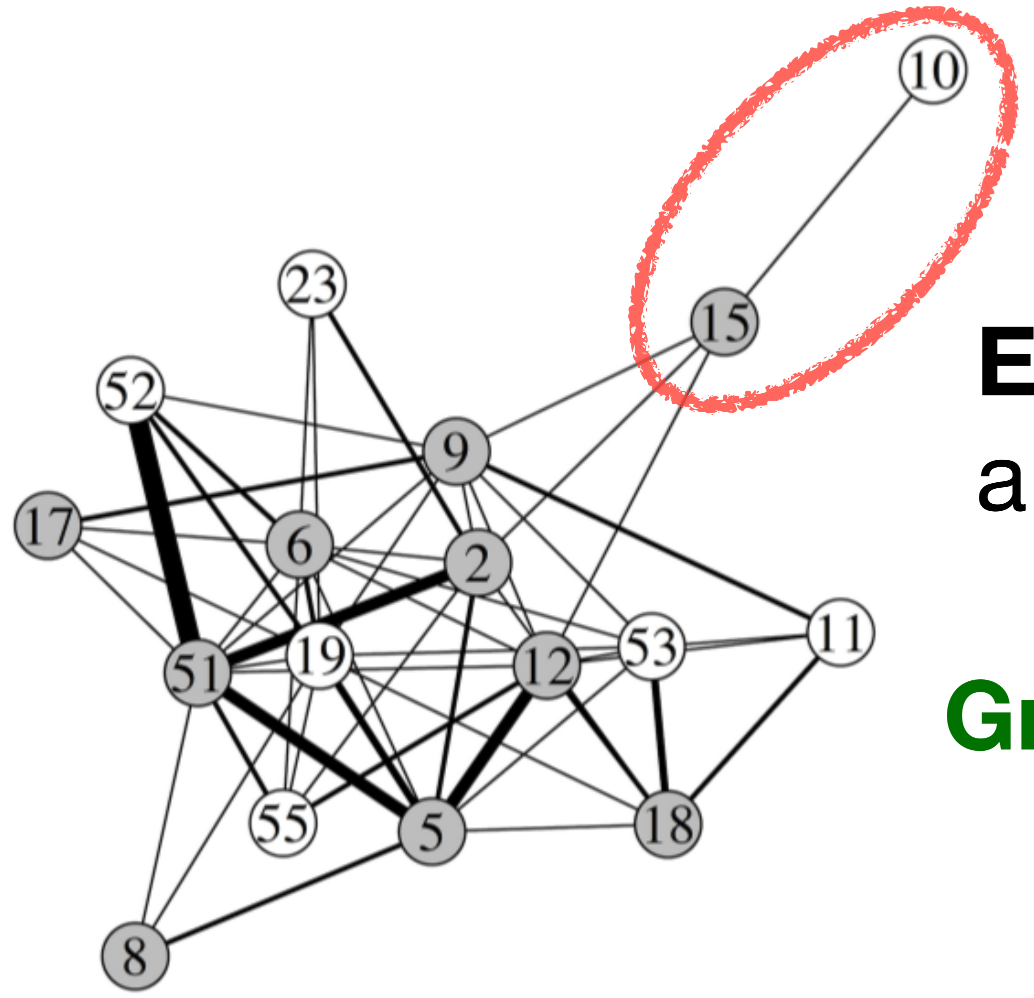
Letter

Construction of simplicial complexes with prescribed degree-size sequences

Tzu-Chi Yen*

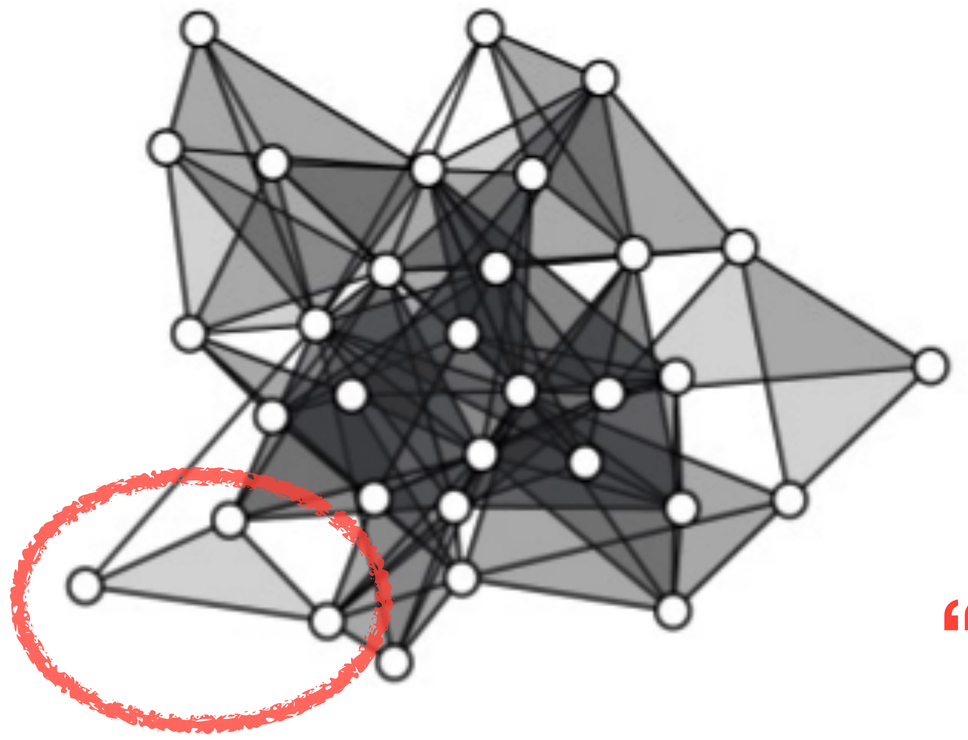
Department of Computer Science, University of Colorado, Boulder, Colorado 80309, USA

(Received 31 May 2021; accepted 30 September 2021; published 28 October 2021)



Edges. These 2 authors have written a paper together.

Graph



“Higher-order” graph

Edges. These 3 authors have worked together (and wrote >1 papers).

NP-completeness. Do not give up if a problem is NP-hard, sometimes a simple algorithm exists!

Null models. Configuration model for graphs vs. for higher-order graphs


PHYSICAL REVIEW E **104**, L042303 (2021)

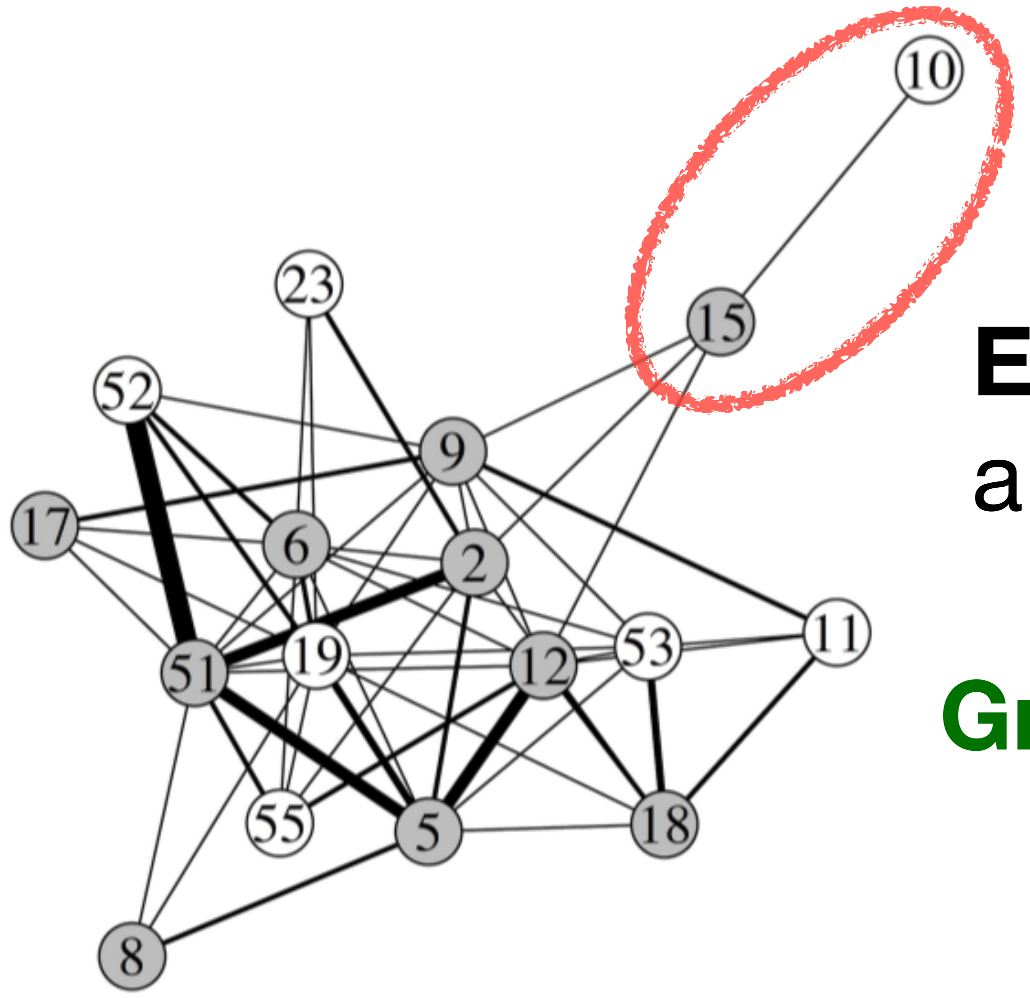
Letter

Construction of simplicial complexes with prescribed degree-size sequences

Tzu-Chi Yen*

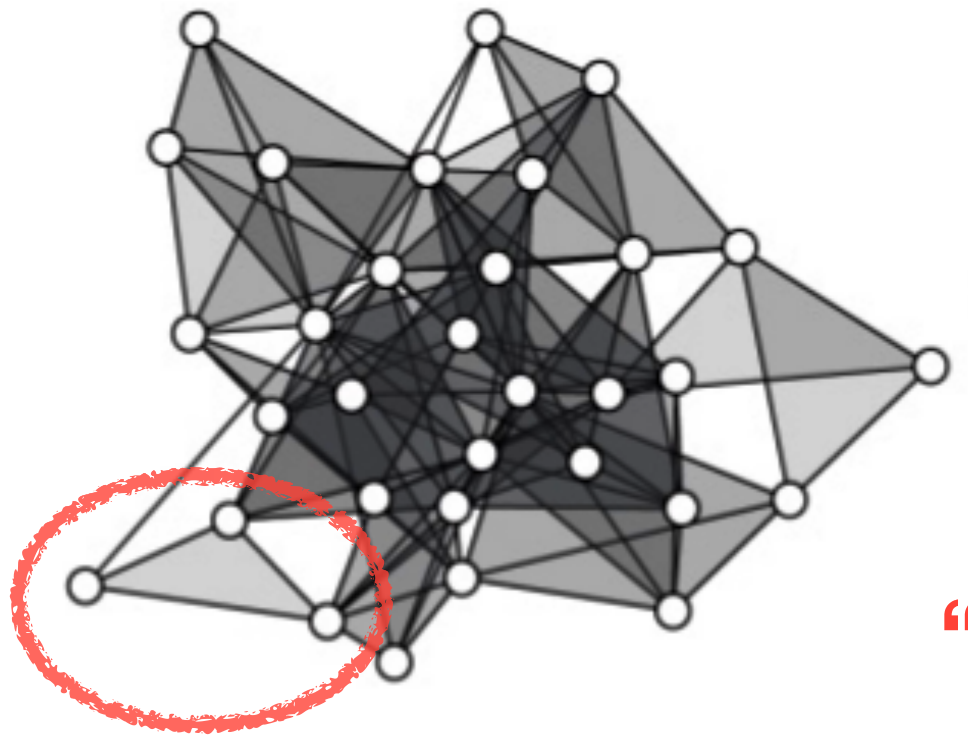
Department of Computer Science, University of Colorado, Boulder, Colorado 80309, USA

(Received 31 May 2021; accepted 30 September 2021; published 28 October 2021)



Edges. These 2 authors have written a paper together.

Graph



"Higher-order" graph

Edges. These 3 authors have worked together (and wrote >1 papers).

NP-completeness. Do not give up if a problem is NP-hard, sometimes a simple algorithm exists!

Null models. Configuration model for graphs vs. for higher-order graphs

Applied topology. How does #holes in a higher-order graph scale?

PHYSICAL REVIEW E **104**, L042303 (2021)

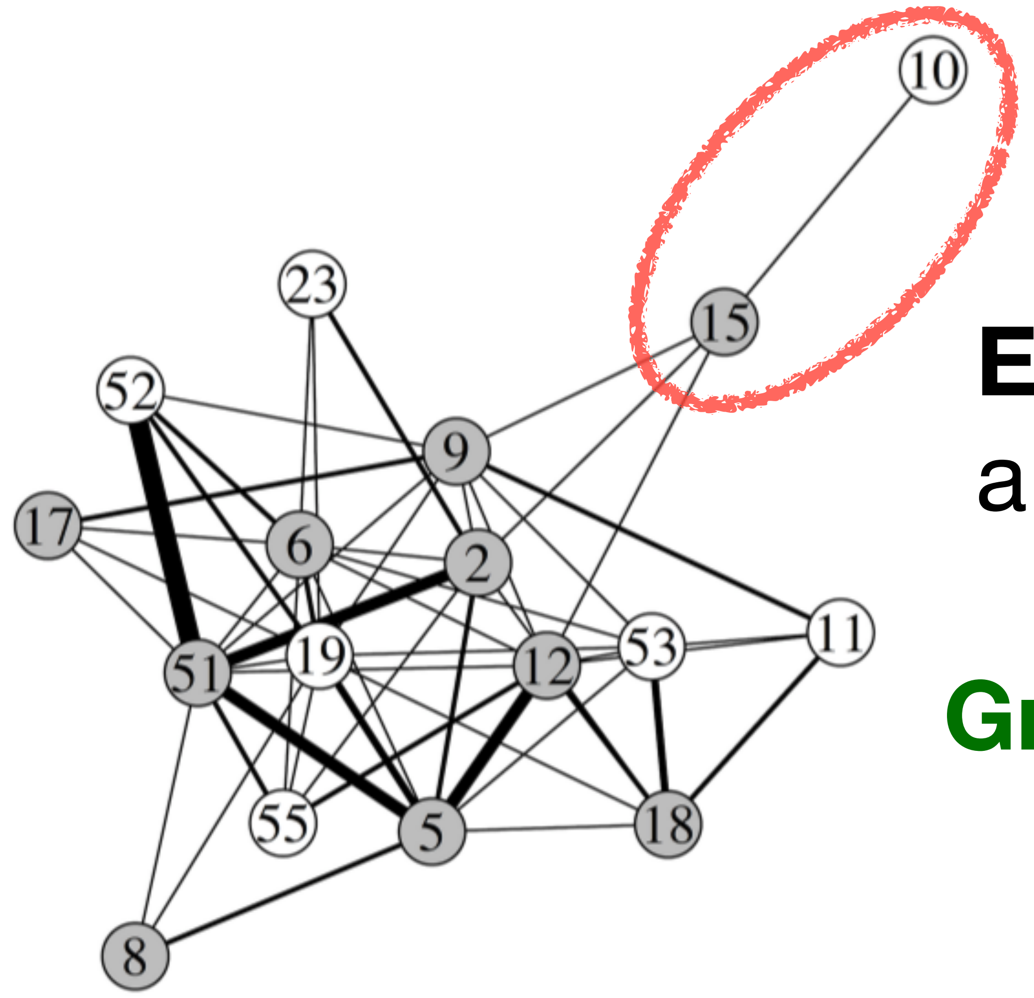
Letter

Construction of simplicial complexes with prescribed degree-size sequences

Tzu-Chi Yen*

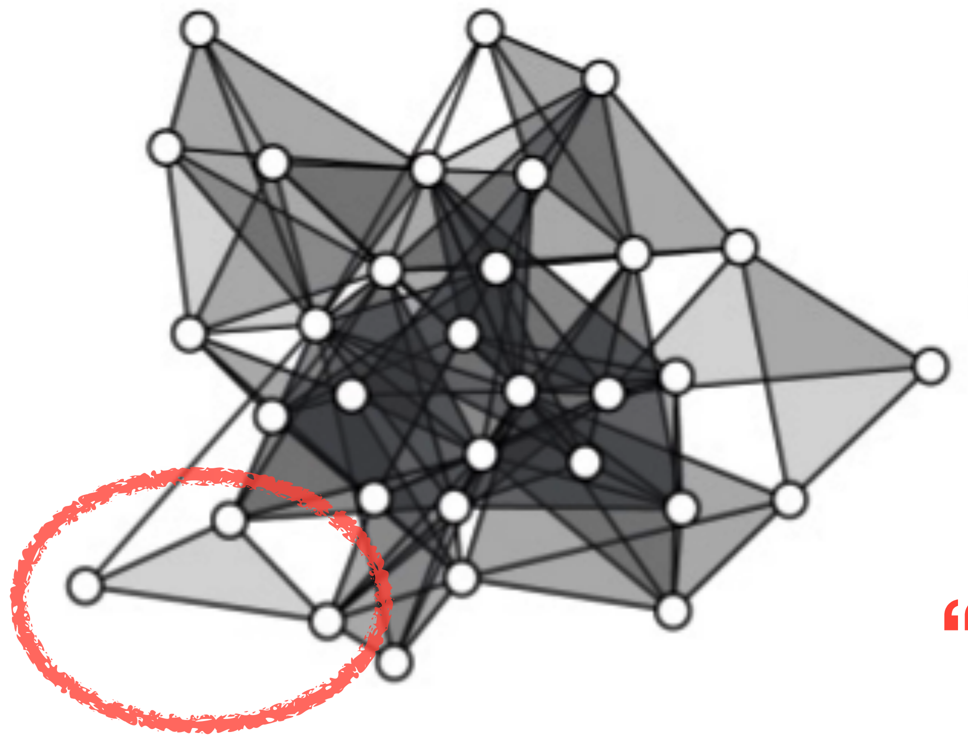
Department of Computer Science, University of Colorado, Boulder, Colorado 80309, USA

(Received 31 May 2021; accepted 30 September 2021; published 28 October 2021)



Edges. These 2 authors have written a paper together.

Graph



"Higher-order" graph

Edges. These 3 authors have worked together (and wrote >1 papers).

NP-completeness. Do not give up if a problem is NP-hard, sometimes a simple algorithm exists!

Null models. Configuration model for graphs vs. for higher-order graphs

Applied topology. How does #holes in a higher-order graph scale?

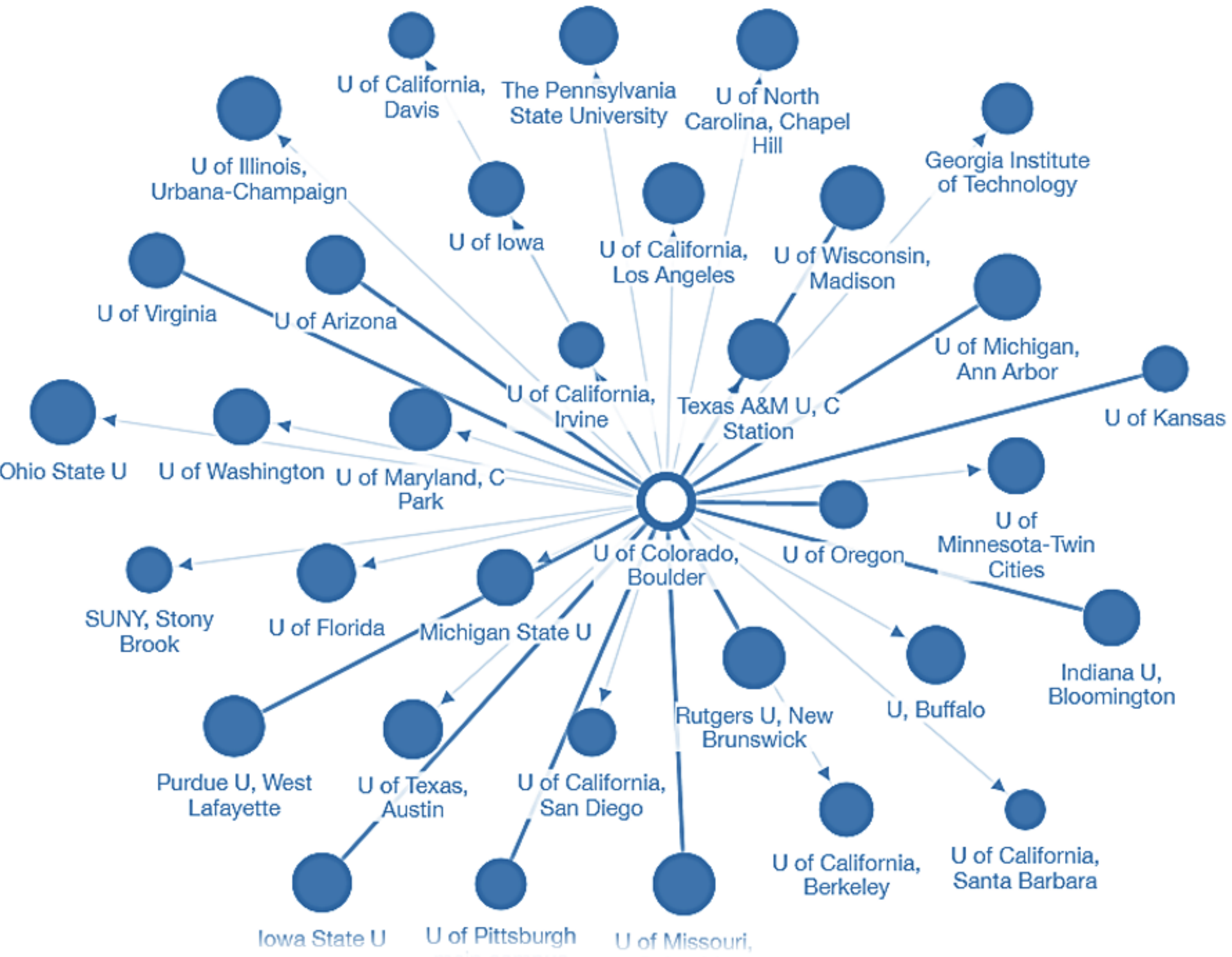
PHYSICAL REVIEW E **104**, L042303 (2021)

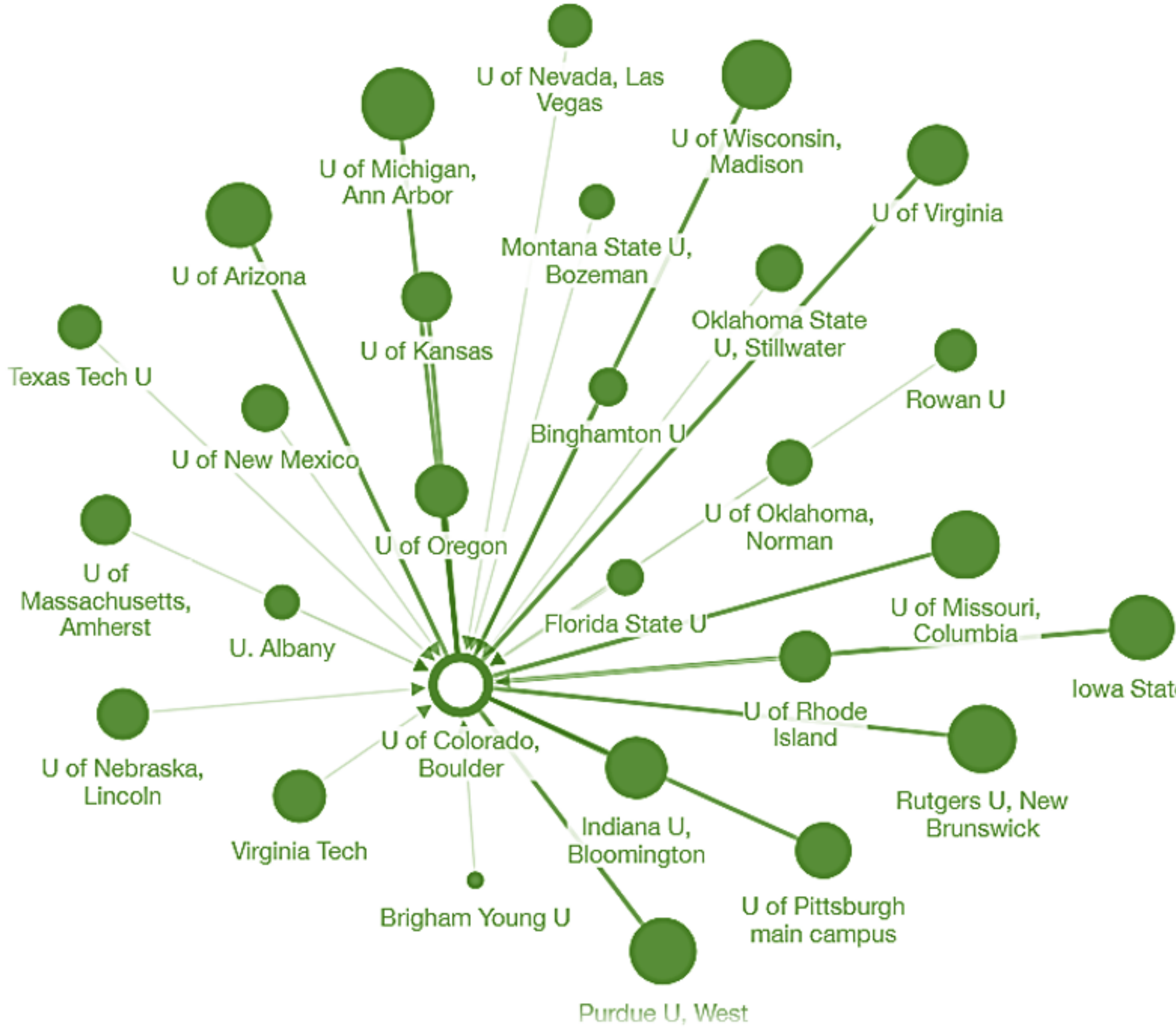
Letter

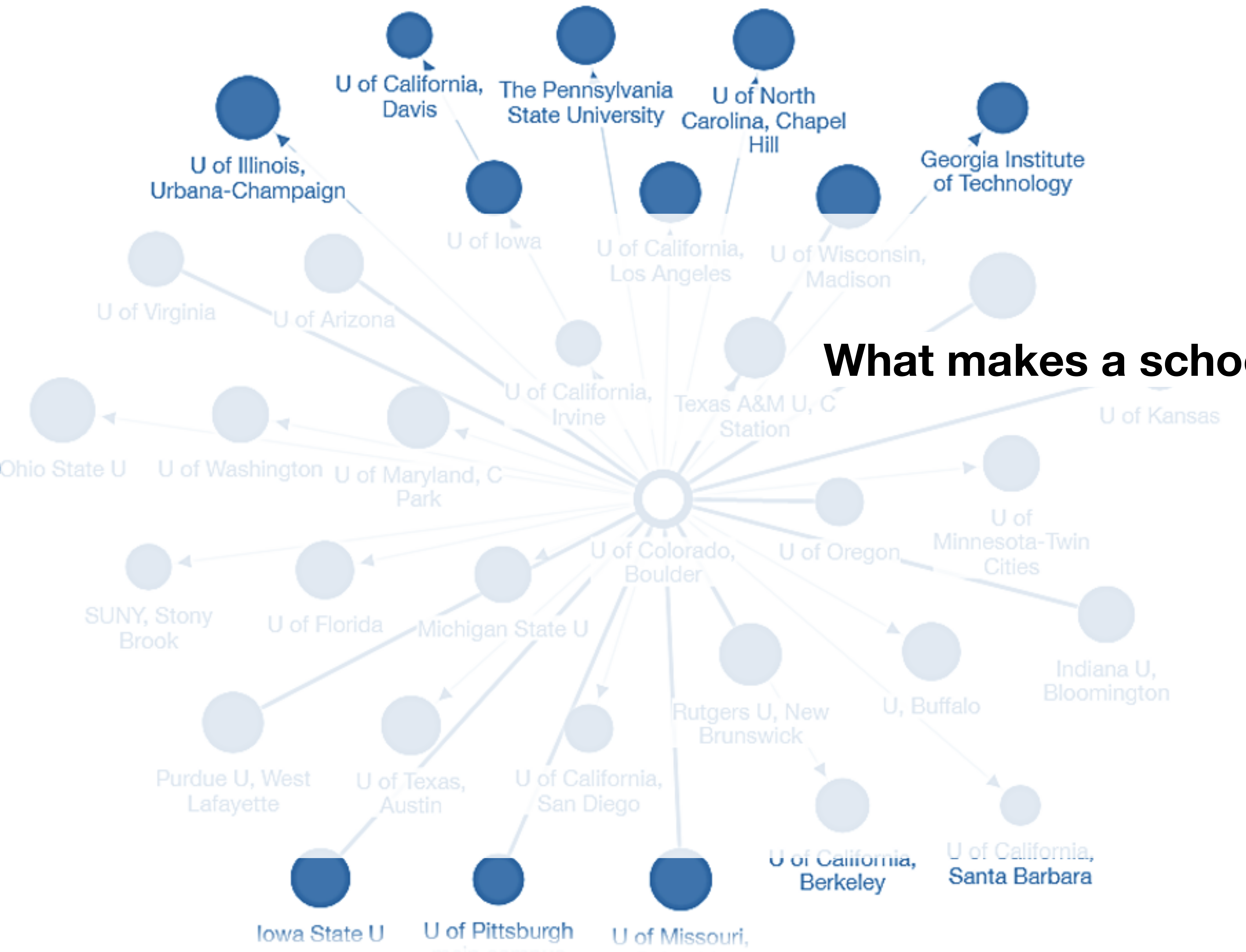
Construction of simplicial complexes with prescribed degree-size sequences

Tzu-Chi Yen *
Department of Computer Science, University of Colorado, Boulder, Colorado 80309, USA

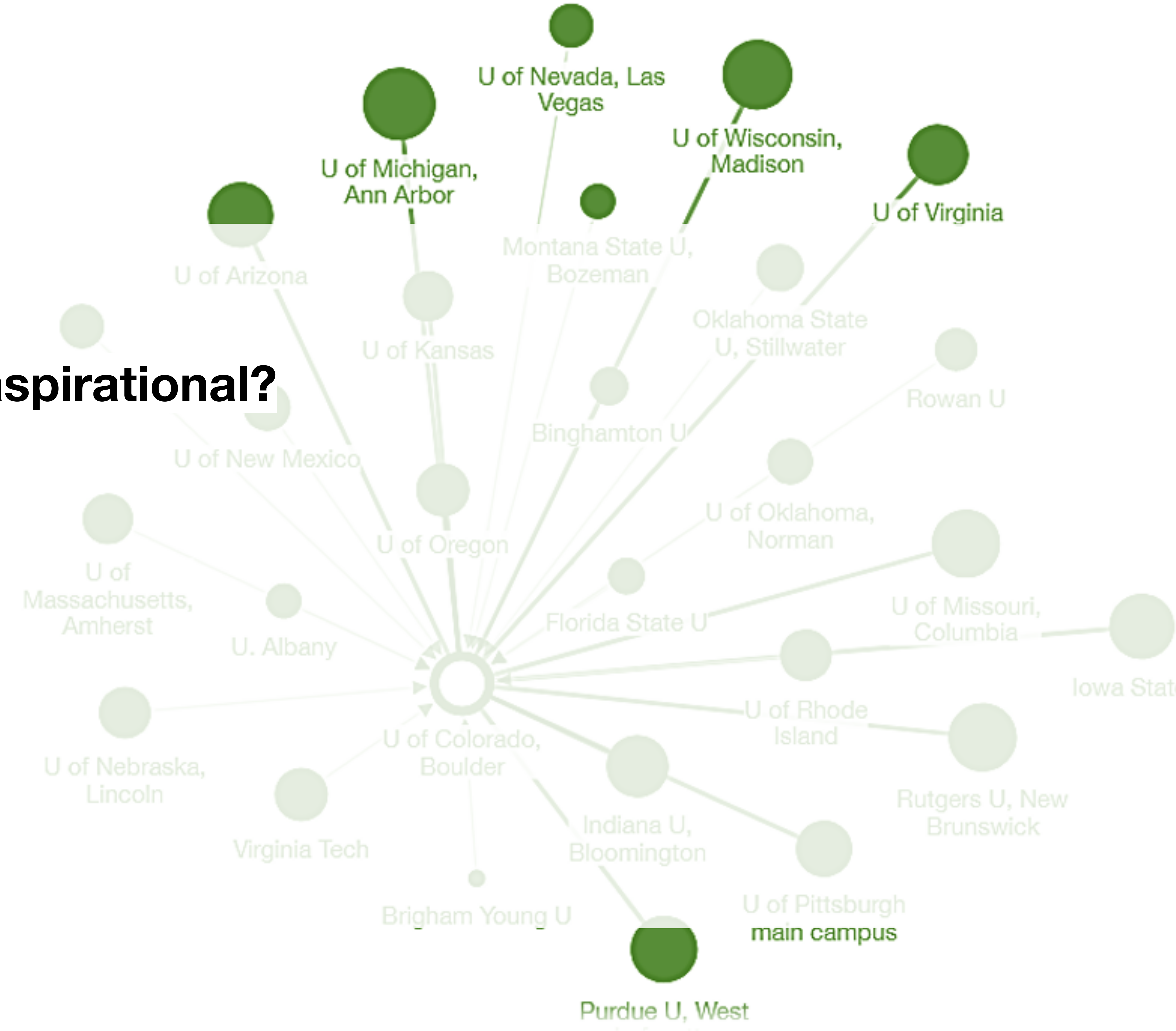
(Received 31 May 2021; accepted 30 September 2021; published 28 October 2021)

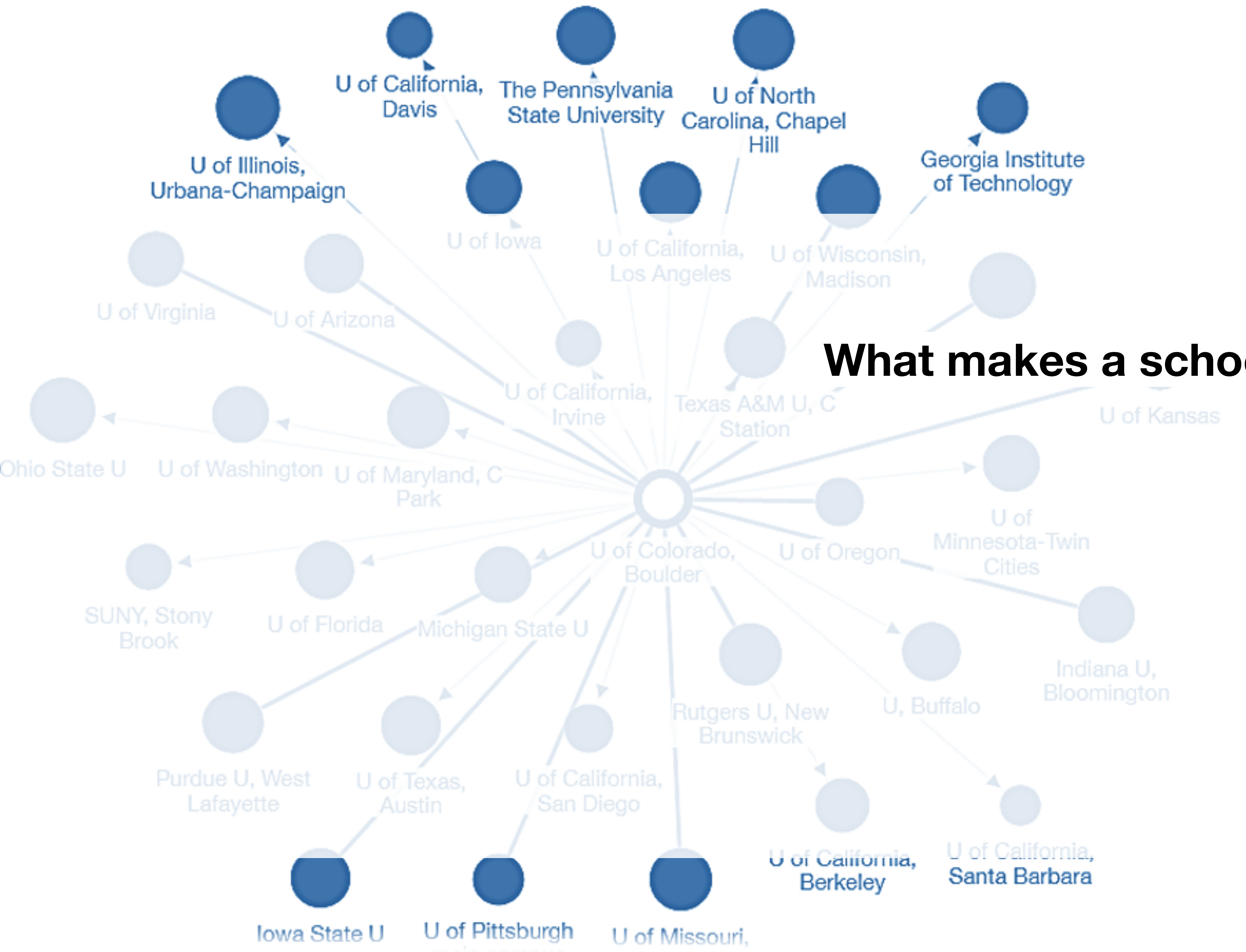

U of California, Davis
The Pennsylvania State University
U of North Carolina, Chapel Hill
U of Illinois, Urbana-Champaign
Georgia Institute of Technology
U of Iowa
U of California, Los Angeles
U of Wisconsin, Madison
U of Virginia
U of Arizona
U of Michigan, Ann Arbor
U of California, Irvine
Texas A&M U, C Station
U of Kansas
Ohio State U
U of Washington
U of Maryland, C Park
U of Minnesota-Twin Cities
U of Colorado, Boulder
U of Oregon
SUNY, Stony Brook
U of Florida
Michigan State U
Indiana U, Bloomington
Rutgers U, New Brunswick
U, Buffalo
Purdue U, West Lafayette
U of Texas, Austin
U of California, San Diego
U of California, Berkeley
U of California, Santa Barbara
Iowa State U
U of Pittsburgh
U of Missouri,


U of Nevada, Las Vegas
U of Michigan, Ann Arbor
U of Wisconsin, Madison
U of Virginia
U of Arizona
Montana State U, Bozeman
U of Kansas
Oklahoma State U, Stillwater
Texas Tech U
Rowan U
Binghamton U
U of New Mexico
U of Oregon
U of Oklahoma, Norman
U of Massachusetts, Amherst
Florida State U
U of Missouri, Columbia
U. Albany
Iowa Stat
U of Rhode Island
U of Colorado, Boulder
U of Nebraska, Lincoln
Rutgers U, New Brunswick
Virginia Tech
Indiana U, Bloomington
U of Pittsburgh main campus
Brigham Young U
Purdue U, West



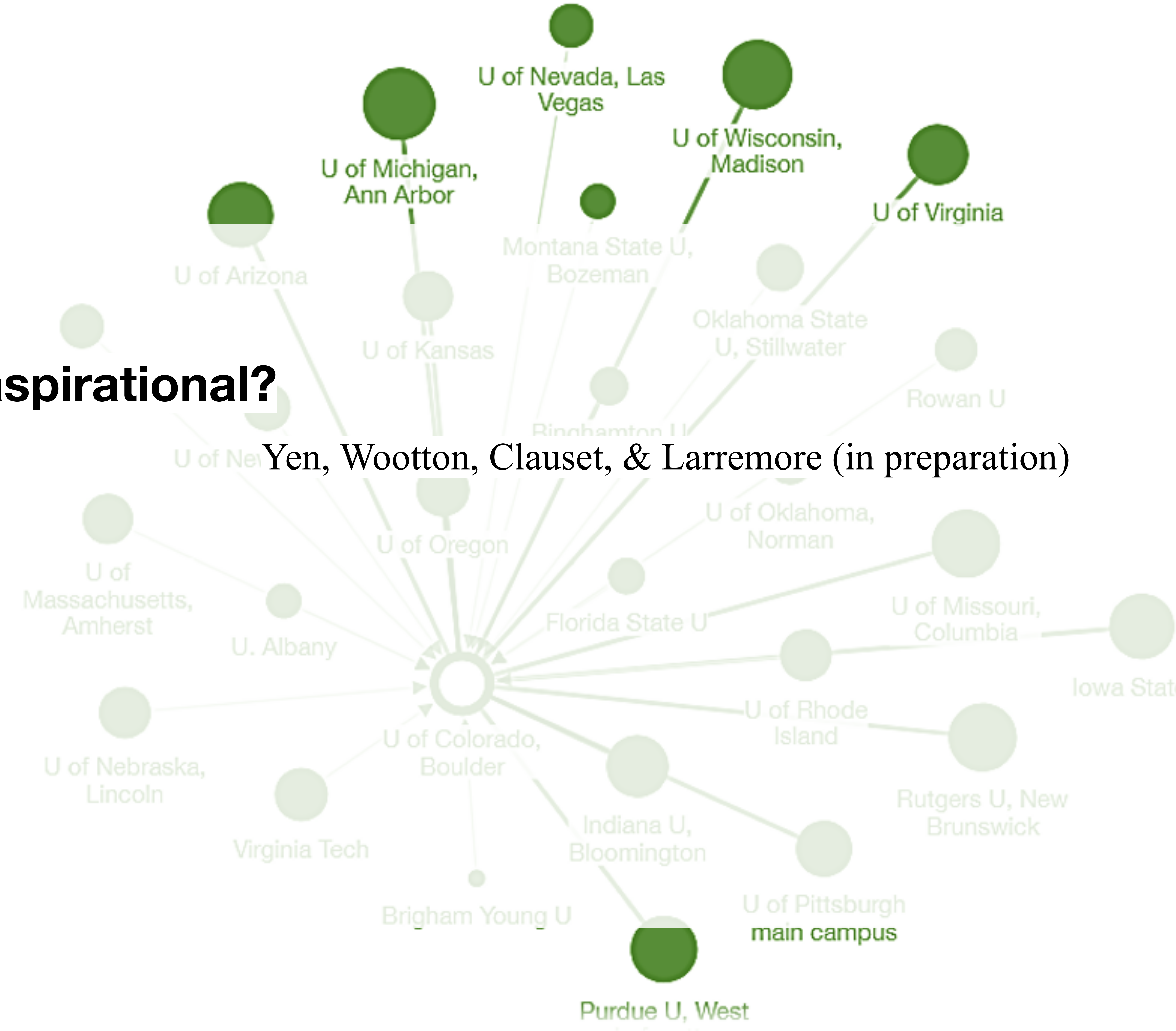
What makes a school aspirational?

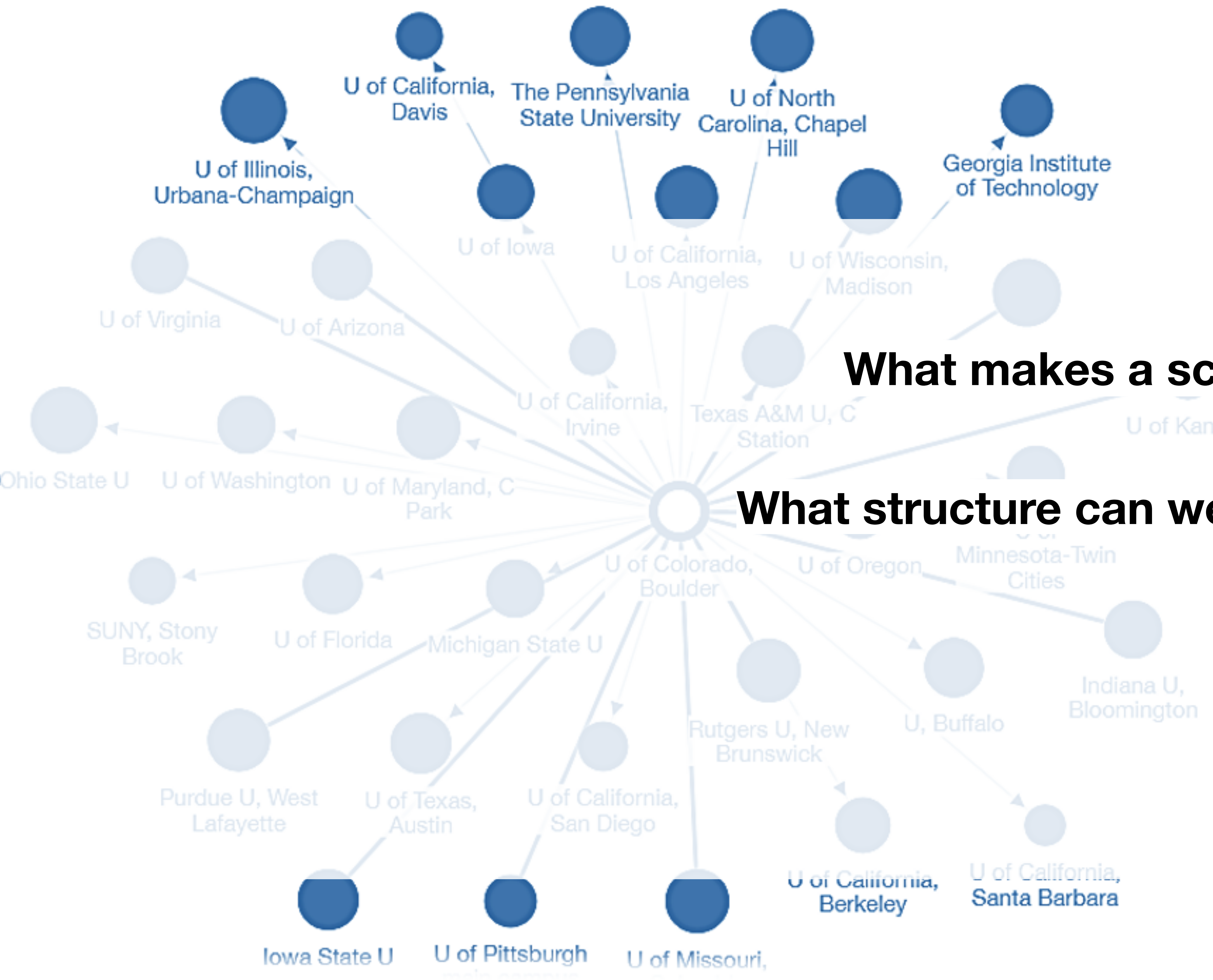




What makes a school aspirational?

Yen, Wootton, Clauset, & Larremore (in preparation)

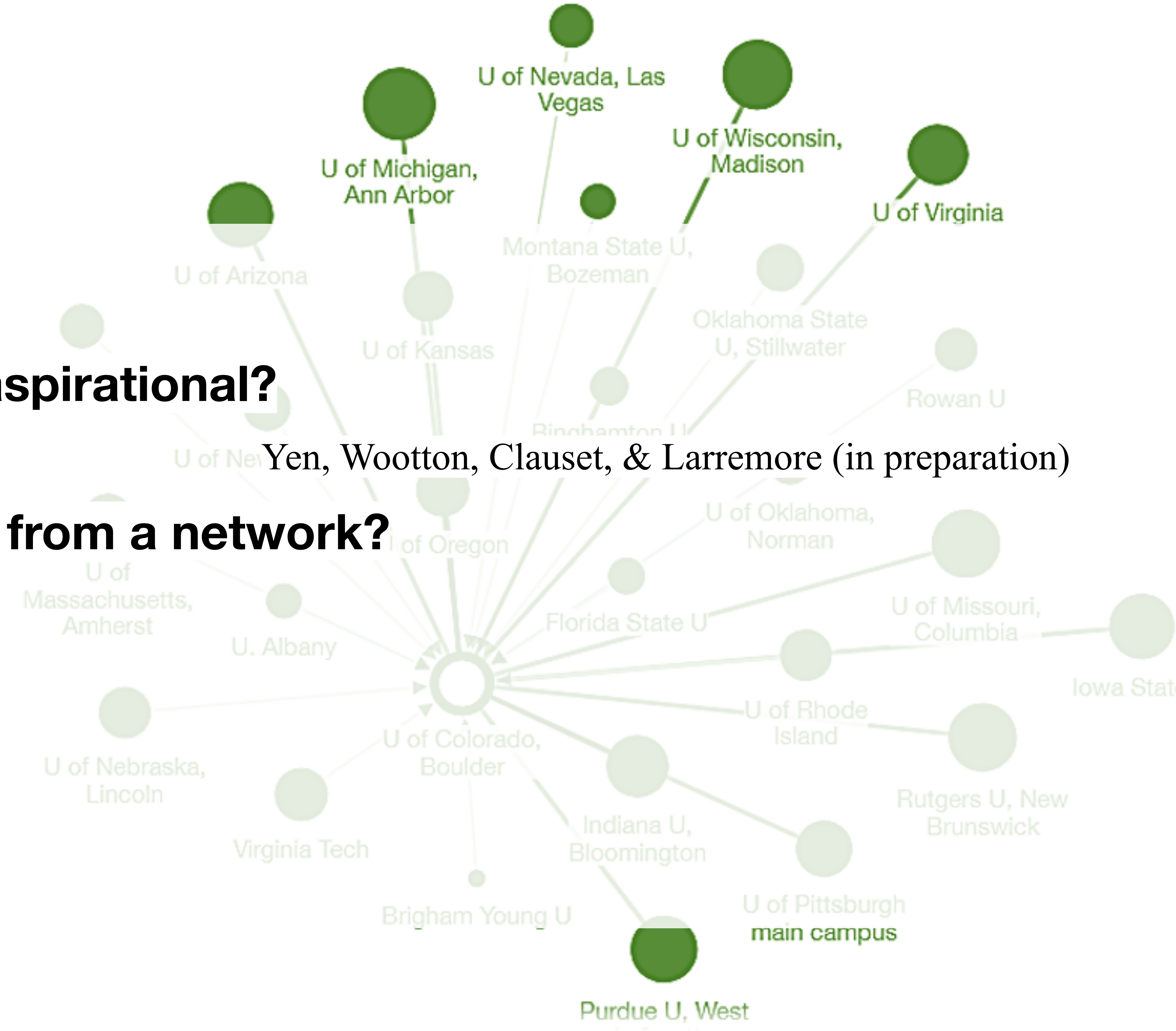


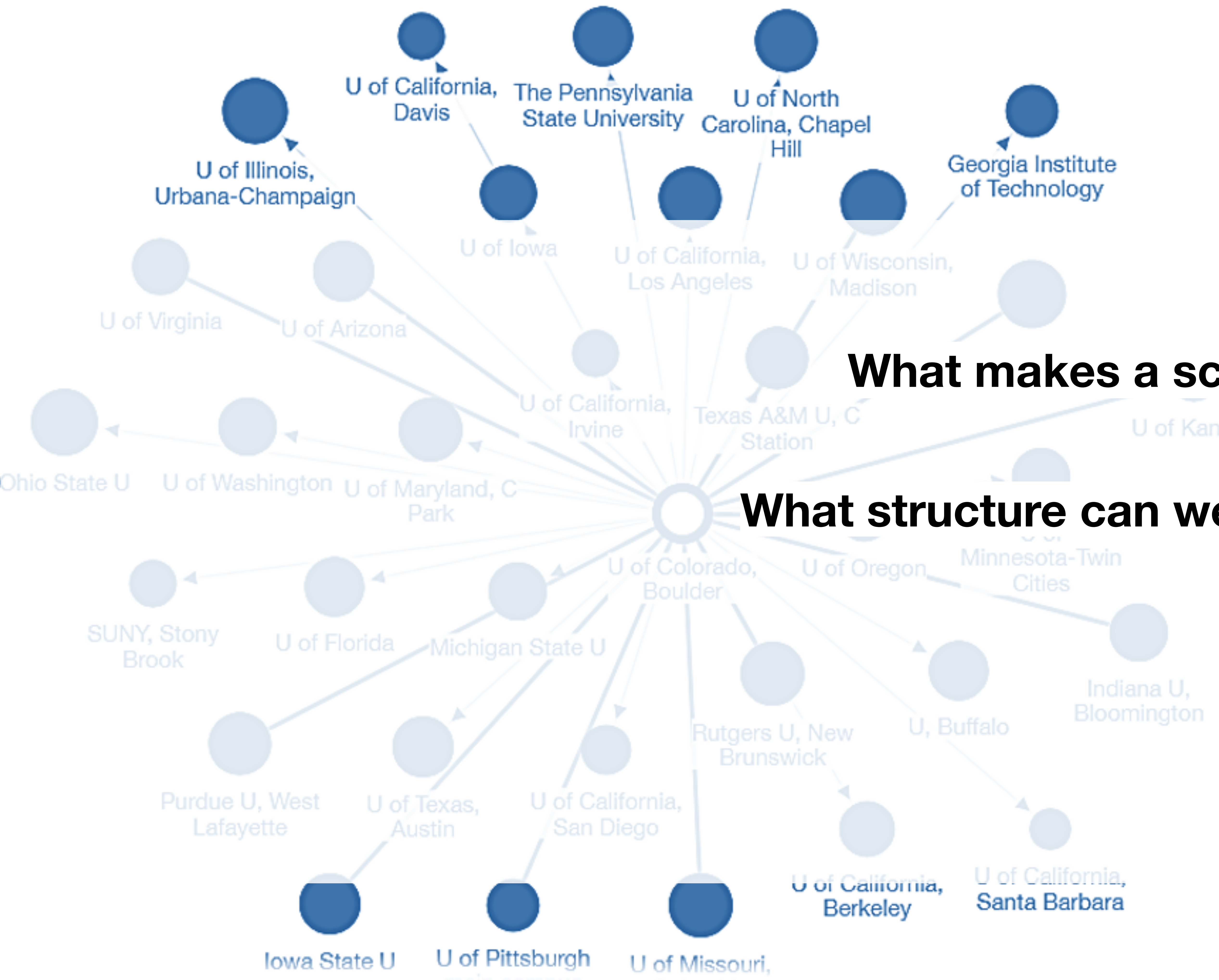


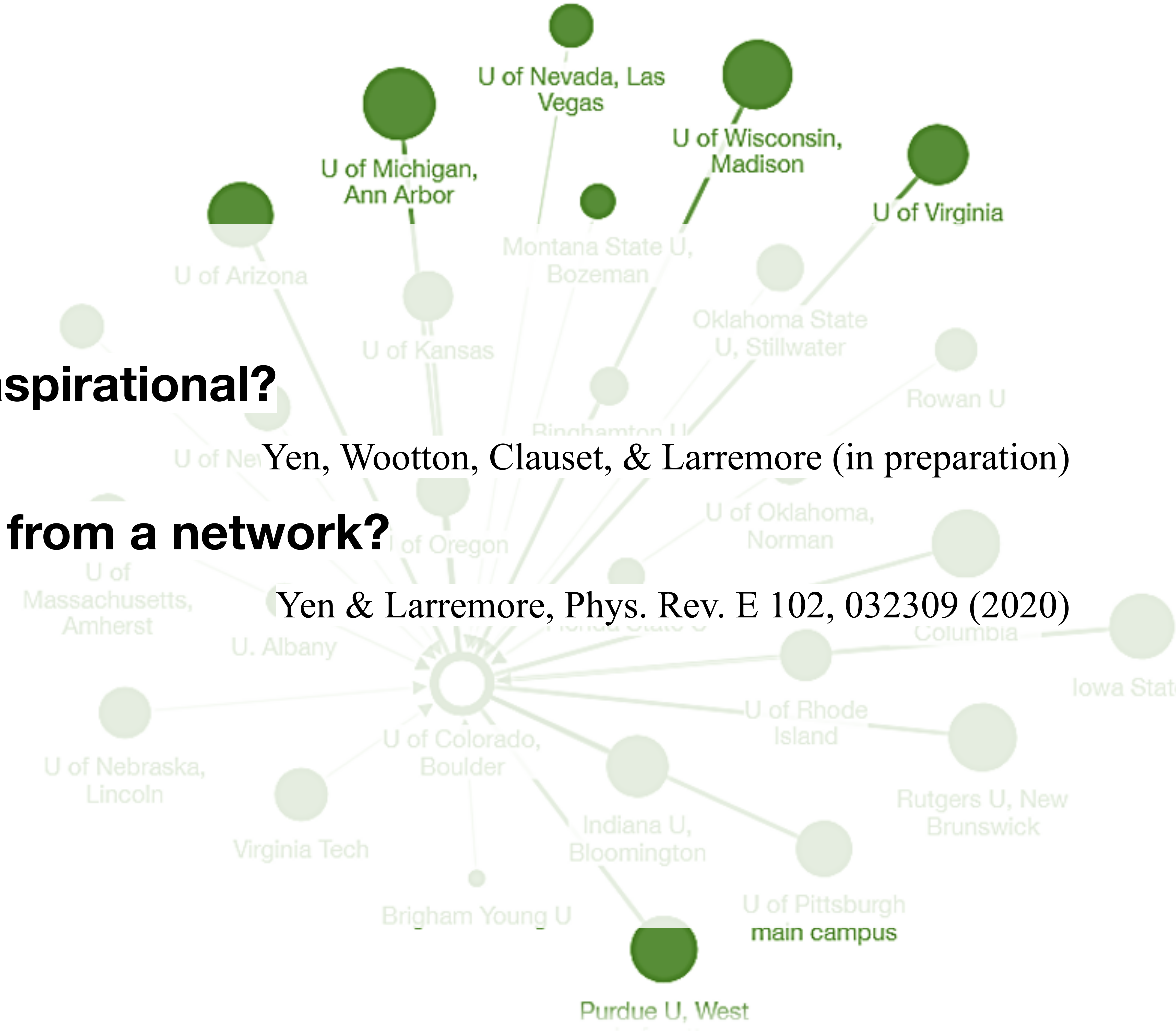
What makes a school aspirational?

Yen, Wootton, Clauset, & Larremore (in preparation)

What structure can we infer from a network?





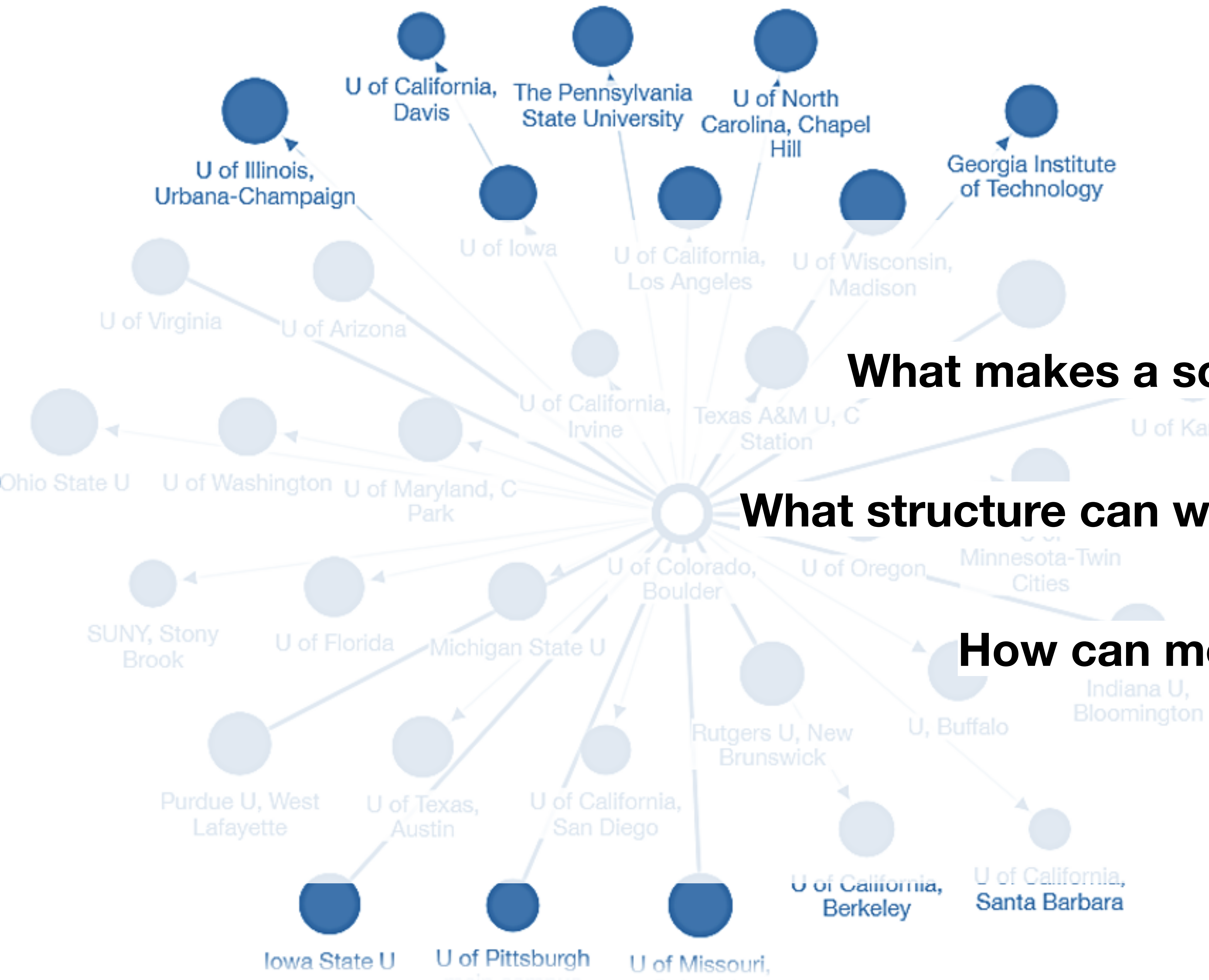


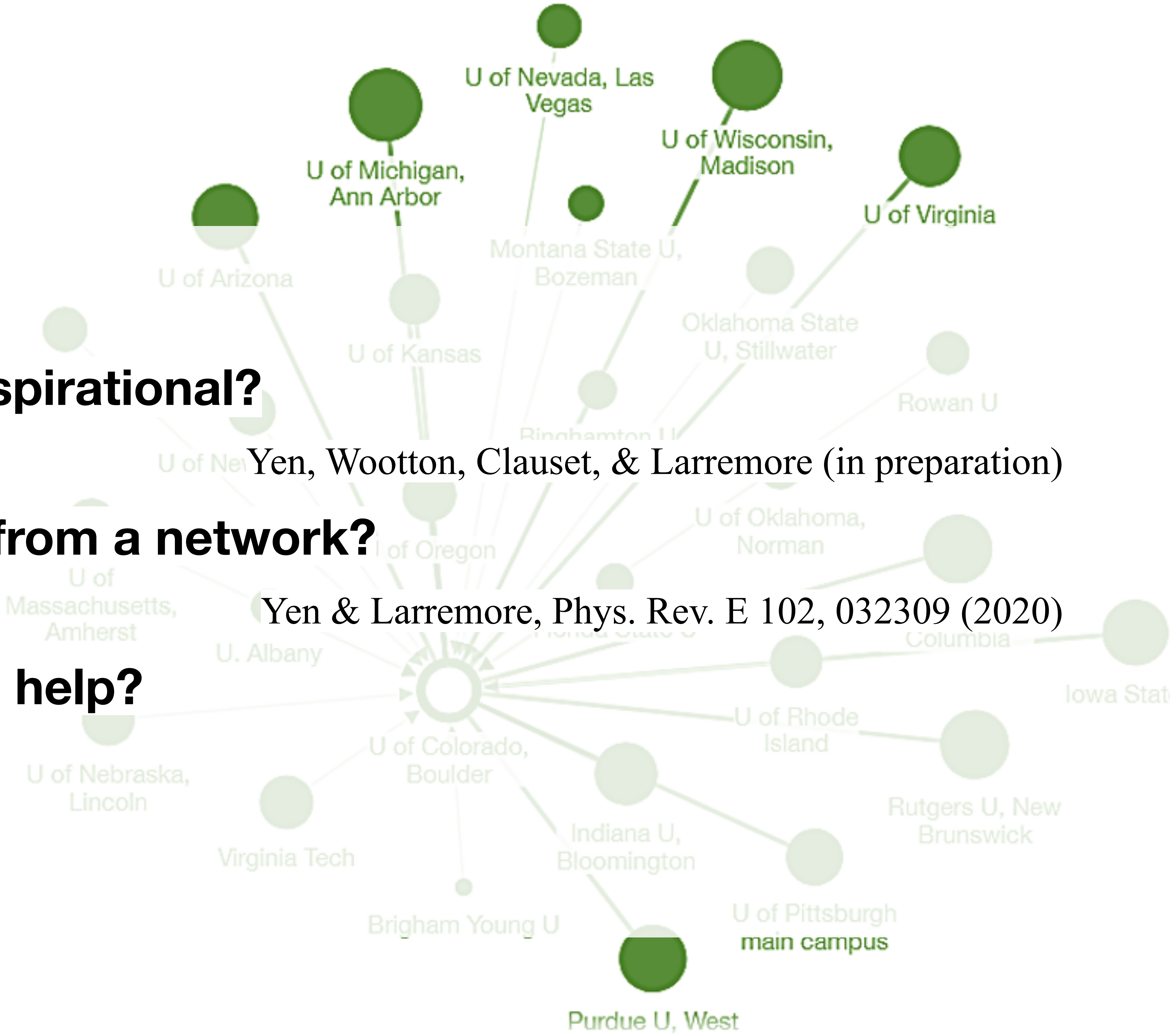
What makes a school aspirational?

Yen, Wootton, Clauset, & Larremore (in preparation)

What structure can we infer from a network?

Yen & Larremore, Phys. Rev. E 102, 032309 (2020)





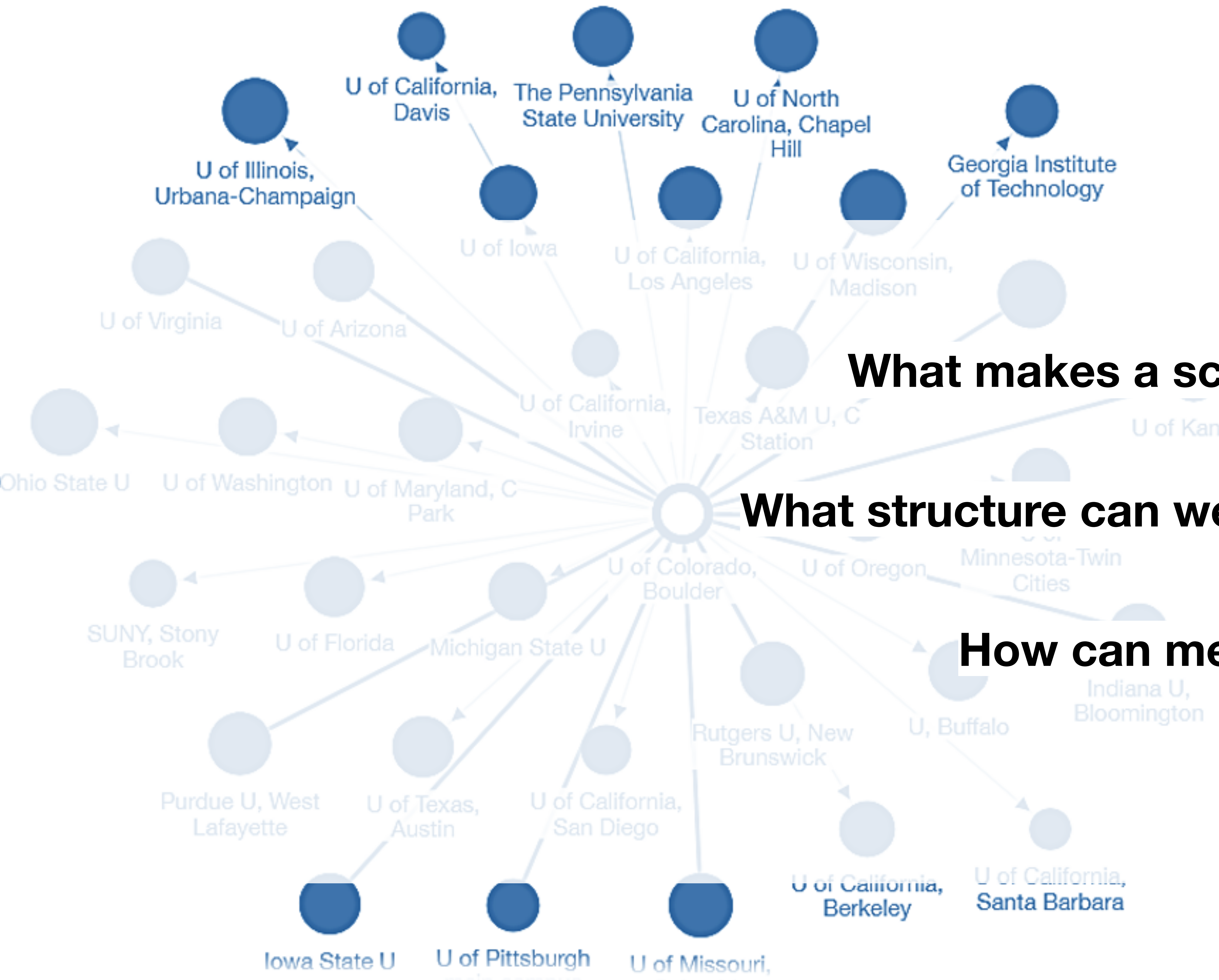
What makes a school aspirational?

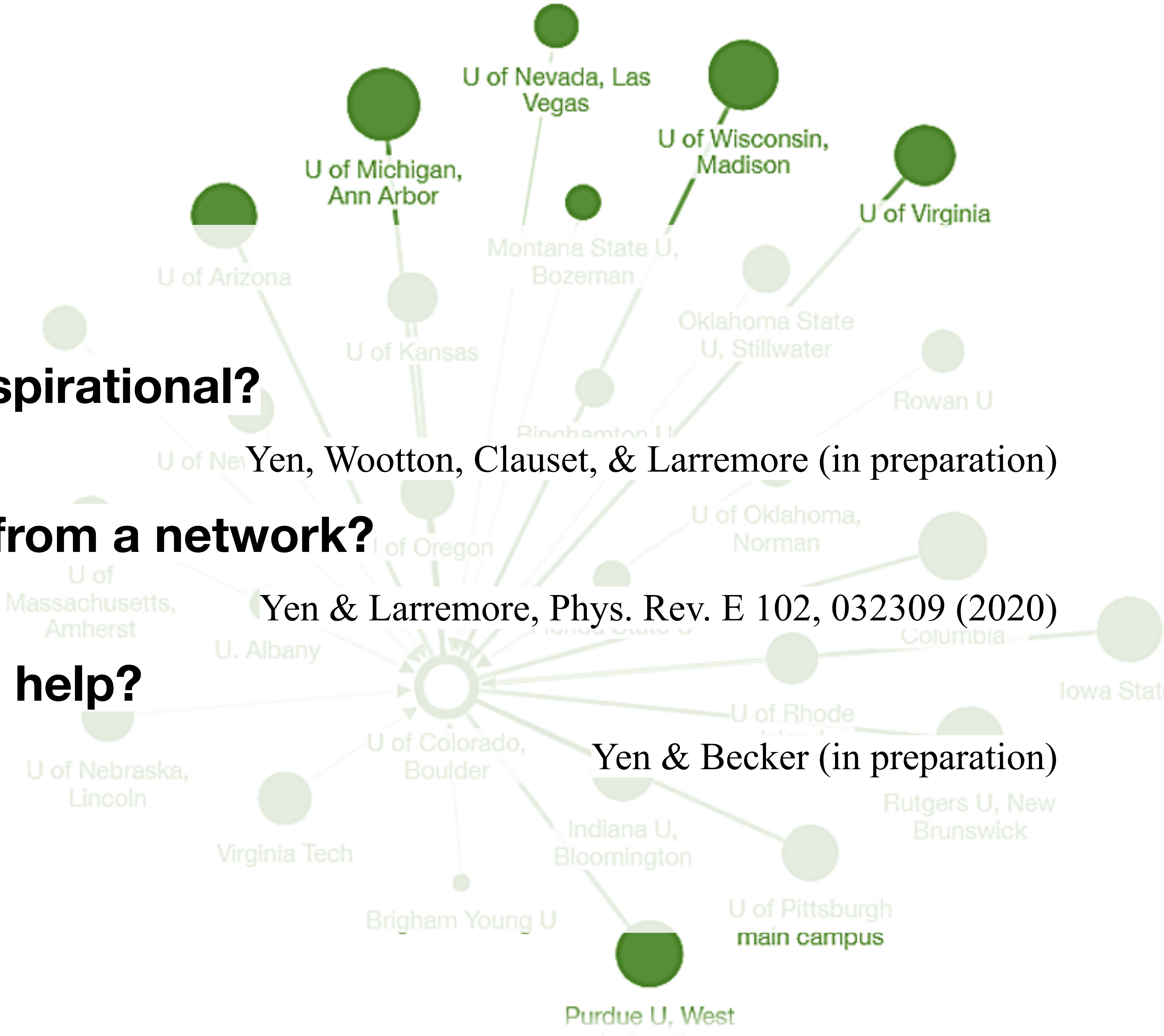
Yen, Wootton, Clauset, & Larremore (in preparation)

What structure can we infer from a network?

Yen & Larremore, Phys. Rev. E 102, 032309 (2020)

How can metadata help?





What makes a school aspirational?

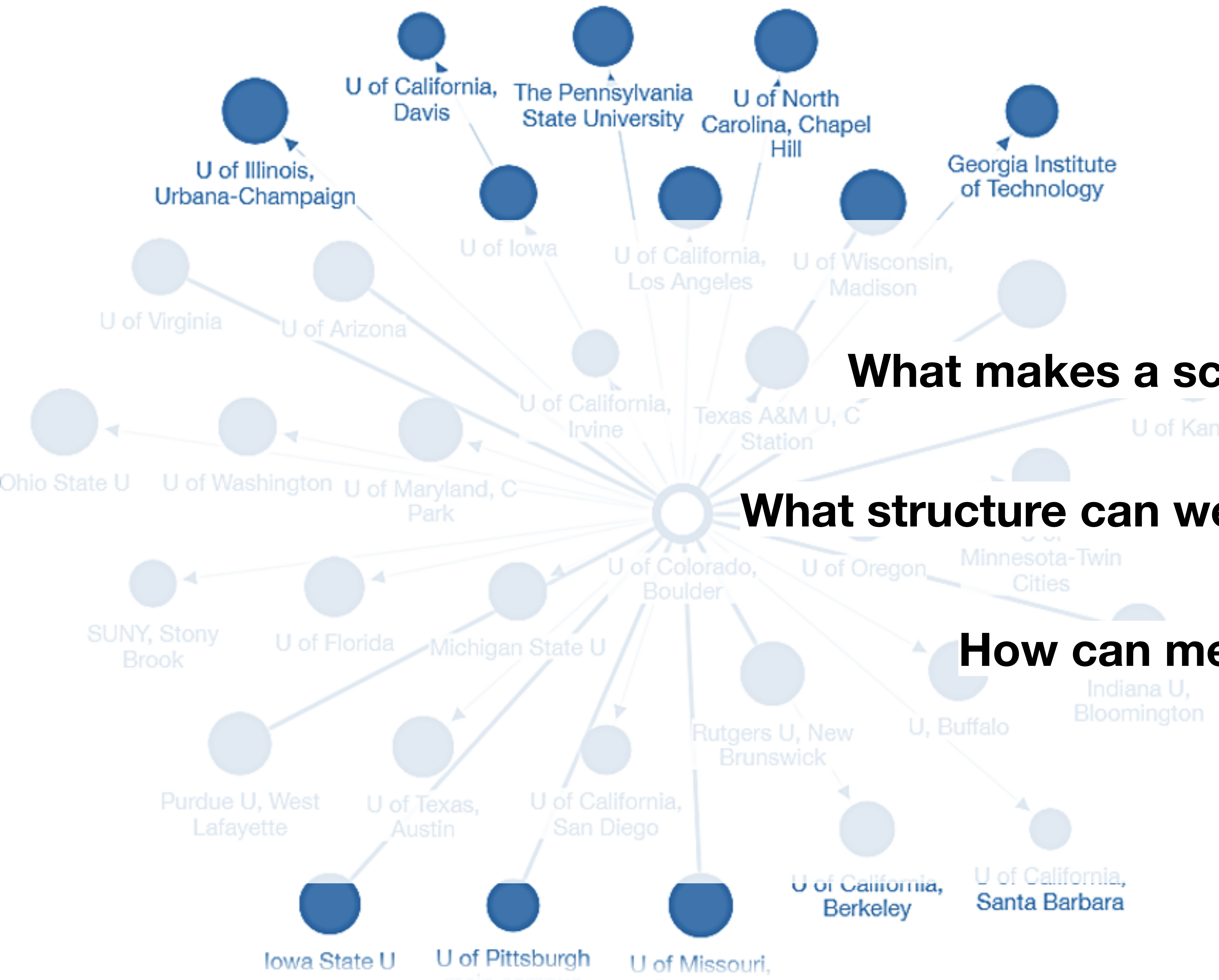
Yen, Wootton, Clauset, & Larremore (in preparation)

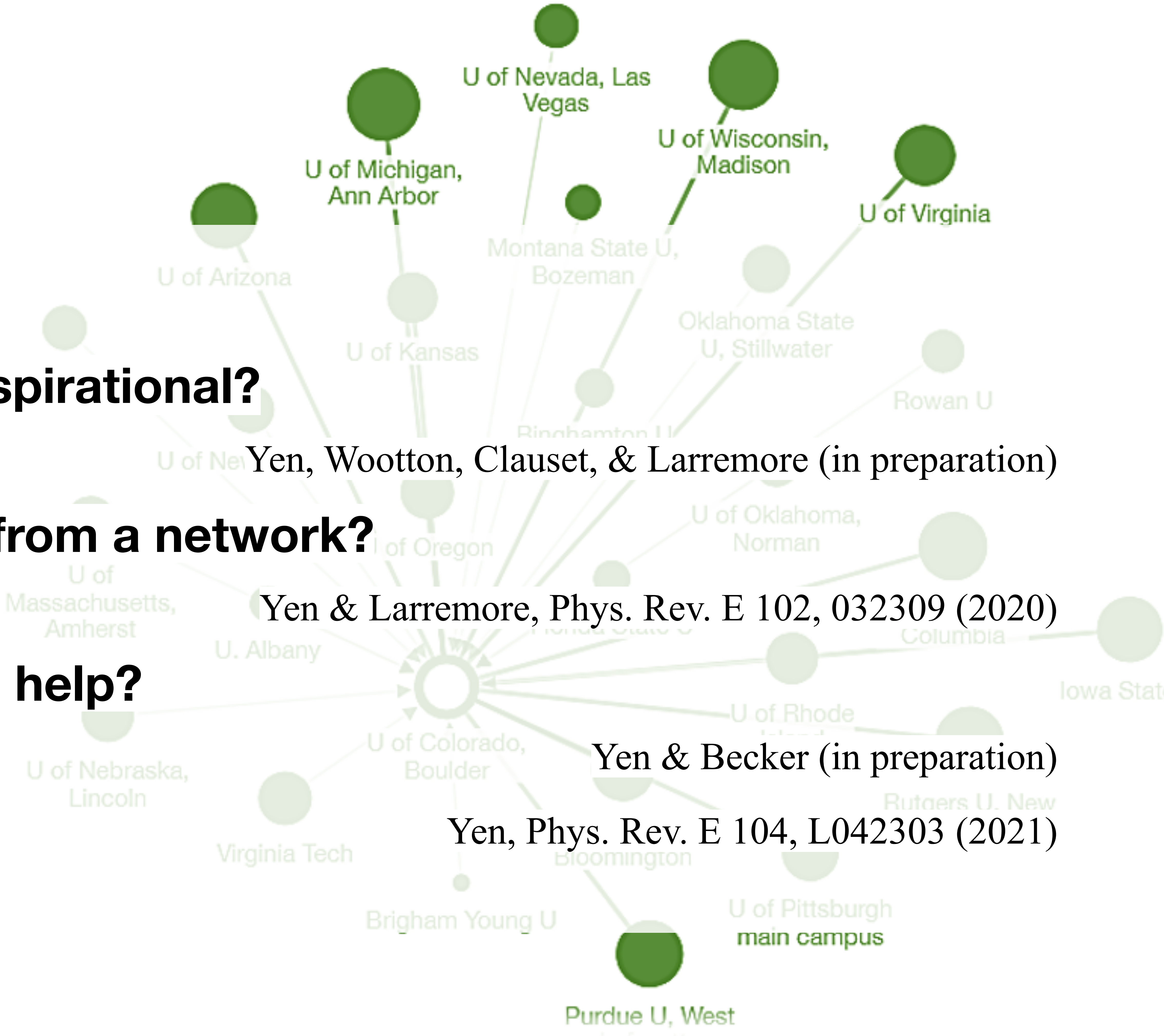
What structure can we infer from a network?

Yen & Larremore, Phys. Rev. E 102, 032309 (2020)

How can metadata help?

Yen & Becker (in preparation)





What makes a school aspirational?

Yen, Wootton, Clauset, & Larremore (in preparation)

What structure can we infer from a network?

Yen & Larremore, Phys. Rev. E 102, 032309 (2020)

How can metadata help?

Yen & Becker (in preparation)

Yen, Phys. Rev. E 104, L042303 (2021)

Future research

Future research

Mix-and-match. Network science + □□□□□□.

Future research

Mix-and-match. Network science + □□□□□□.

Structure is mundane, psychology is alive.

- Directed edge. Action verbs.
- Metadata. Show intent.

Future research

Mix-and-match. Network science + □□□□□□.

Structure is mundane, psychology is alive.

- Directed edge. Action verbs.
- Metadata. Show intent.

Reproducible results. Statistically grounded methodologies.

Statistical inference links data and theory in network science

Leto Peel, Tiago P. Peixoto & Manlio De Domenico

Nature Communications **13**, Article number: 6794 (2022) | Cite this article

Future research

Mix-and-match. Network science + □□□□□□.

Structure is mundane, psychology is alive.

- Directed edge. Action verbs.
- Metadata. Show intent.

Reproducible results. Statistically grounded methodologies.

Exploiting sparsity. Combinatorial problems often admits a natural (convex) relaxation.

Statistical inference links data and theory in network science

Leto Peel, Tiago P. Peixoto & Manlio De Domenico

Nature Communications **13**, Article number: 6794 (2022) | Cite this article

A unified perspective on convex structured sparsity: Hierarchical, symmetric, submodular norms and beyond

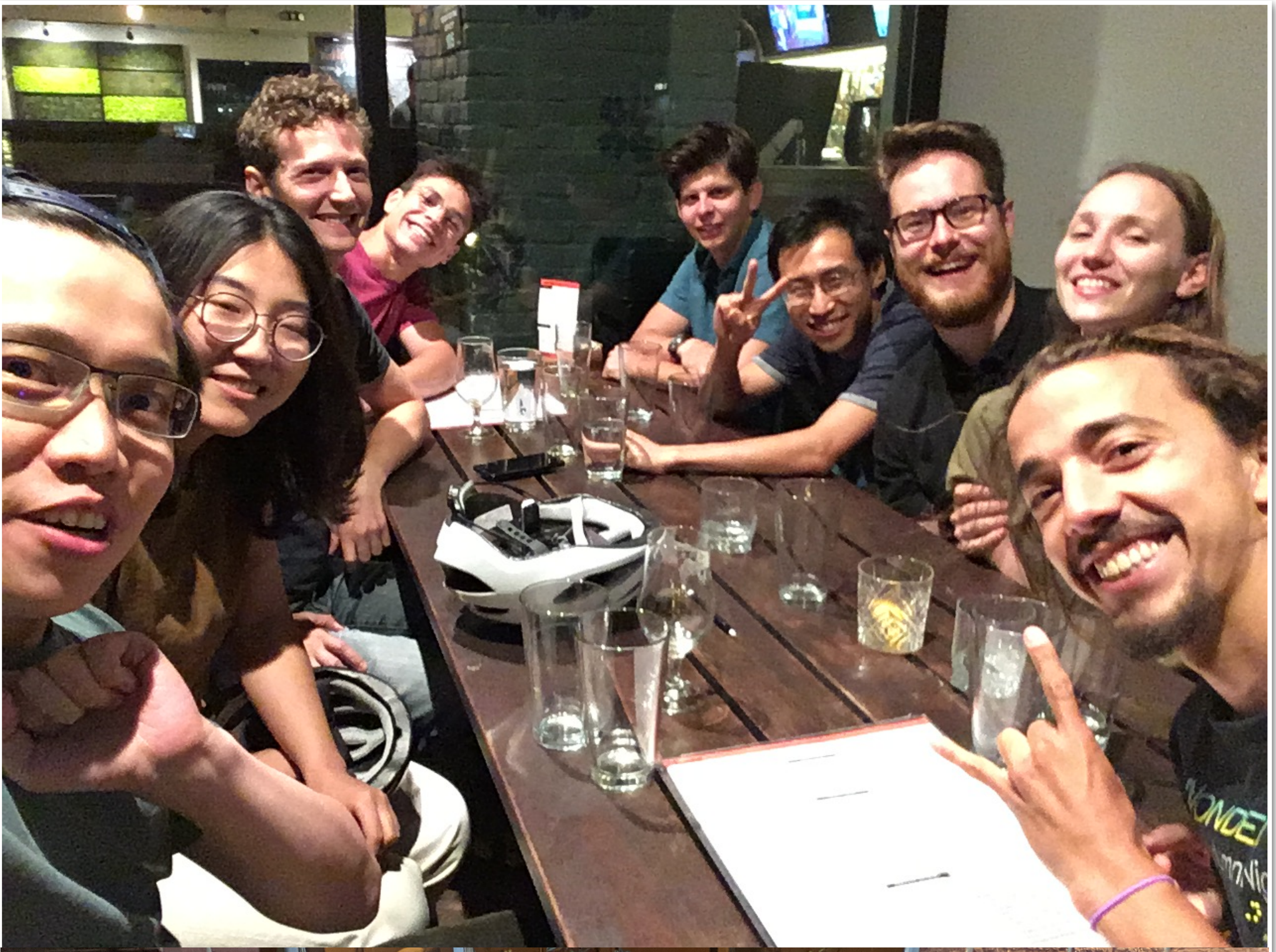
Guillaume Obozinski
Université Paris-Est
Laboratoire d'Informatique Gaspard Monge
Groupe Imagine, Ecole des Ponts - ParisTech
Marne-la-Vallée, France
guillaume.obozinski@enpc.fr

Francis Bach
INRIA - Sierra project-team
Département d'Informatique
de l'Ecole Normale Supérieure
Paris, France
francis.bach@ens.fr

December 8, 2016





Thanks!



Jose Cuervo
BEST MARGARITAS
EFRAIN'S II
MEXICAN
RESTAURANT
BEST
WEST