Slides: bit.ly/lcl-slides
Preprint: bit.ly/lcl-paper
Code: bit.ly/lcl-code
Data: bit.ly/lcl-data

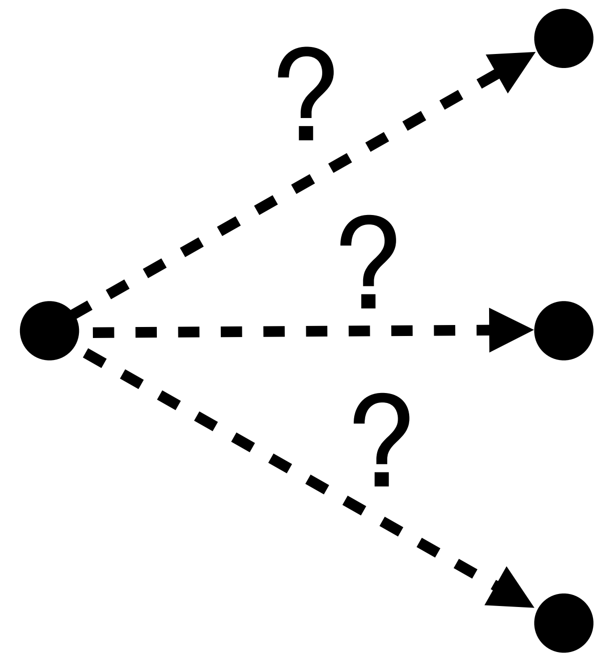
Learning Context Effects in Triadic Closure

Kiran Tomlinson

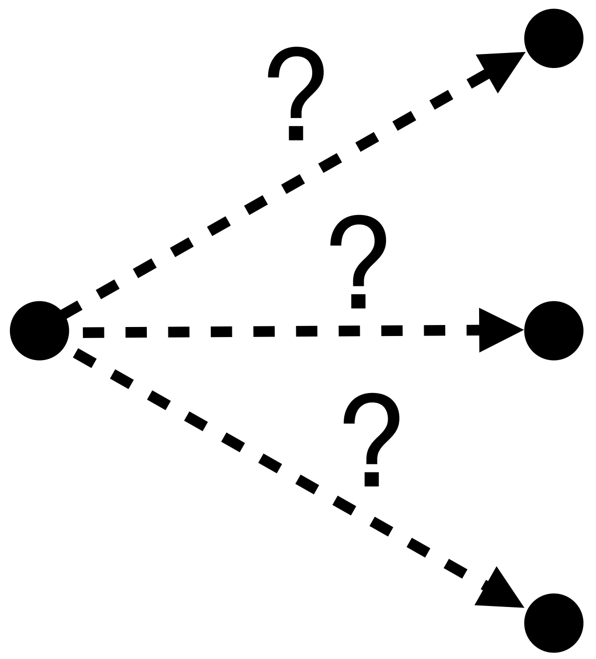
SINM 2020

research with Austin R. Benson

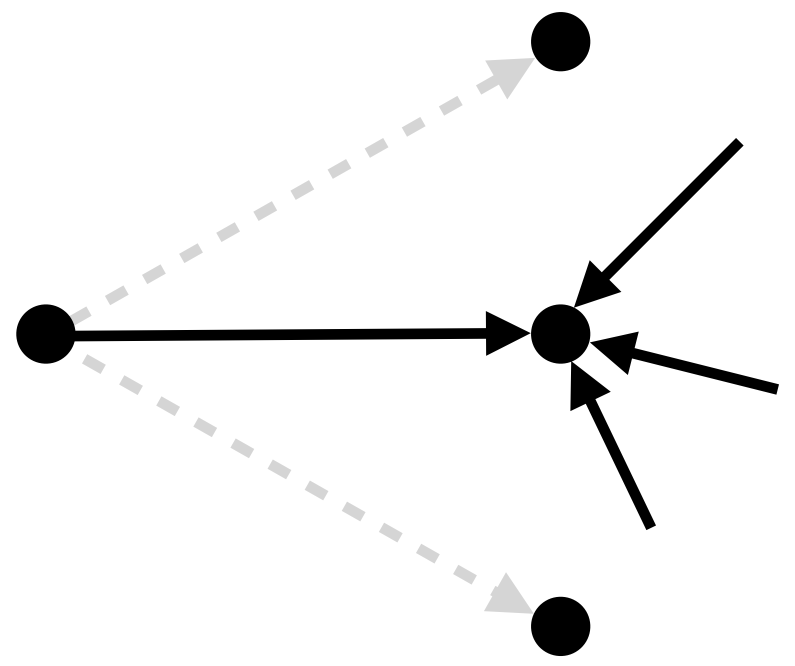
What factors drive edge formation?



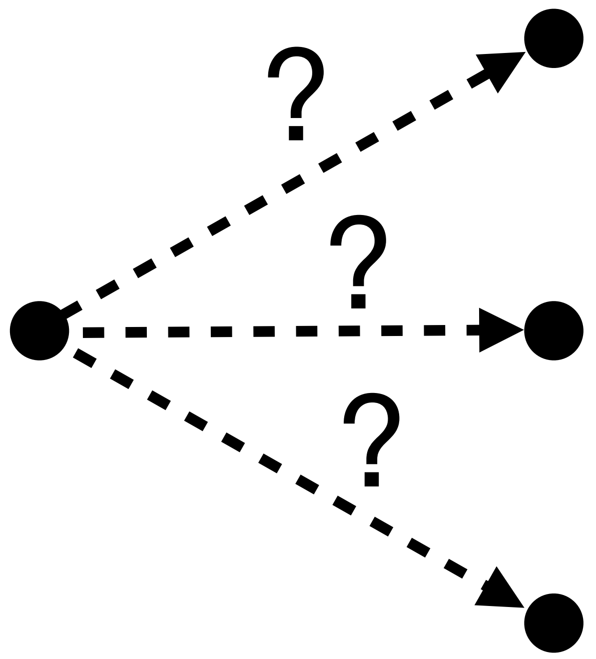
What factors drive edge formation?



Preferential attachment
(Barabási & Albert, *Science* 1999)

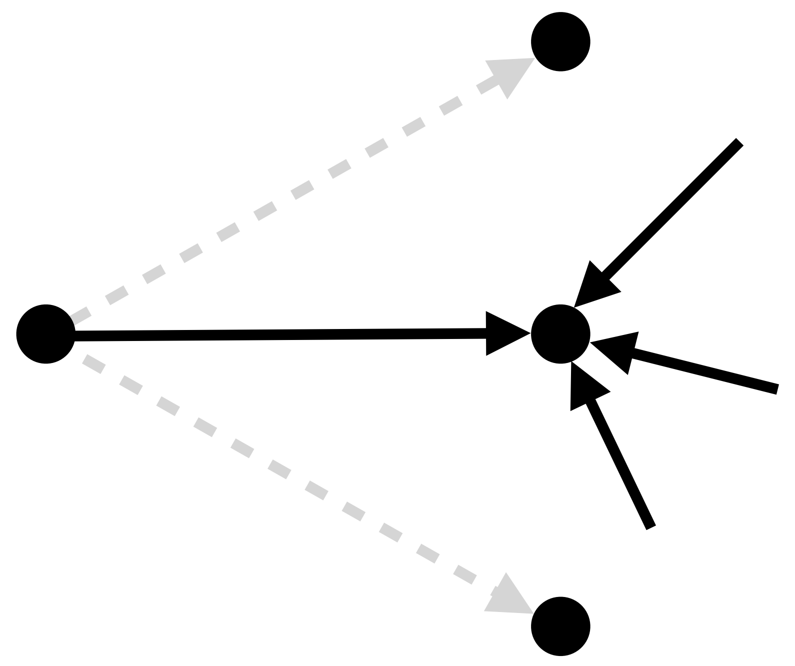


What factors drive edge formation?



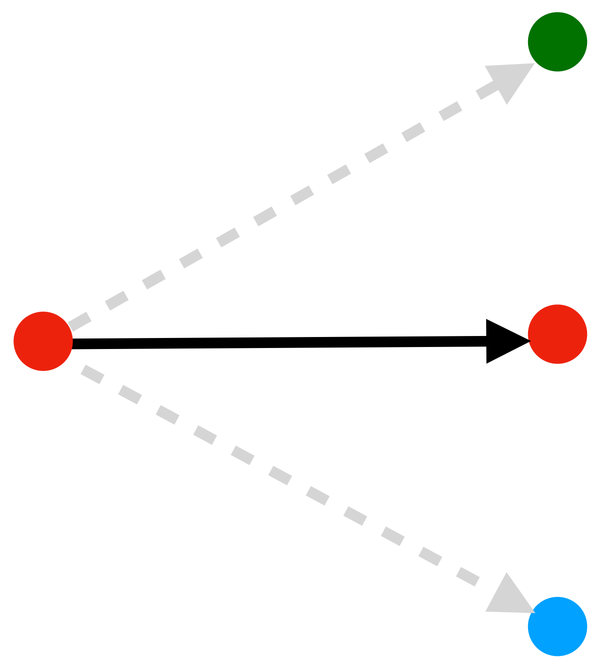
Preferential attachment

(Barabási & Albert, *Science* 1999)

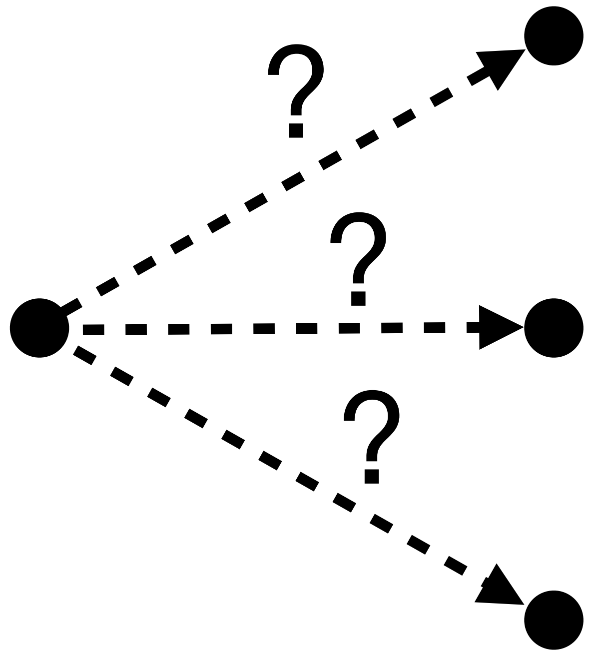


Homophily

(McPherson et al., *Annual Review of Sociology* 2001)
(Papadopoulos et al., *Nature* 2012)

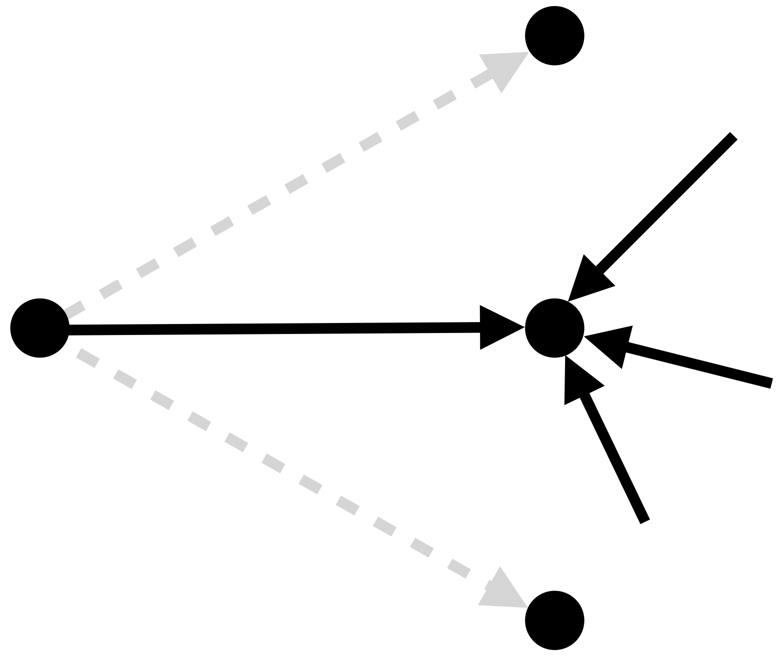


What factors drive edge formation?



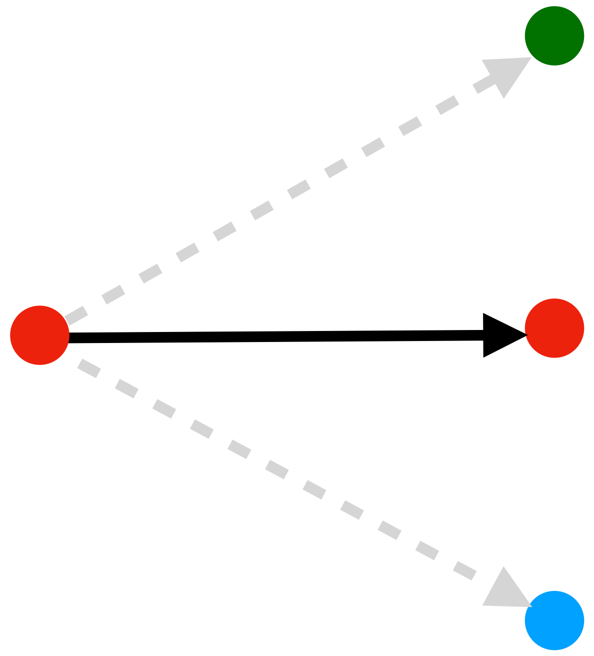
Preferential attachment

(Barabási & Albert, *Science* 1999)



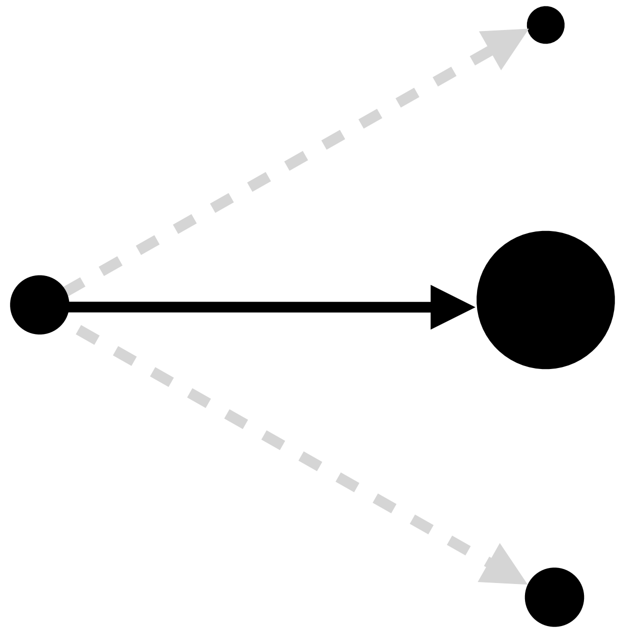
Homophily

(McPherson et al., *Annual Review of Sociology* 2001)
(Papadopoulos et al., *Nature* 2012)

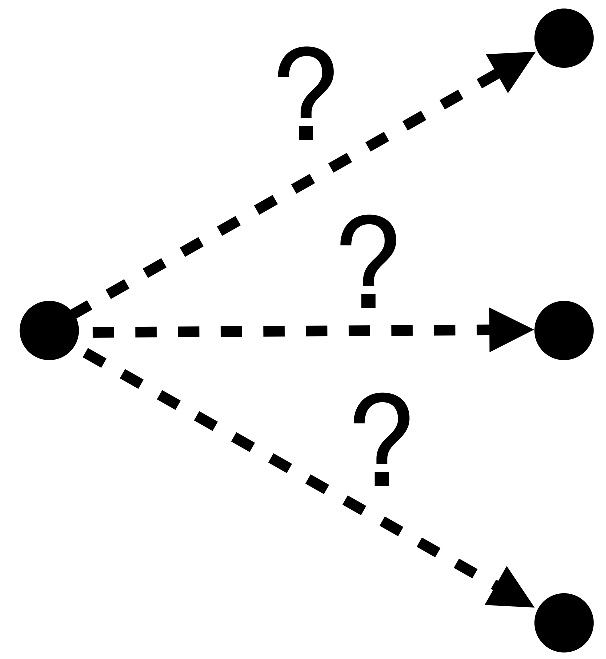


Fitness

(Bianconi & Barabási, *Europhysics Letters* 2001)
(Caldarelli et al., *Physical Review Letters* 2002)

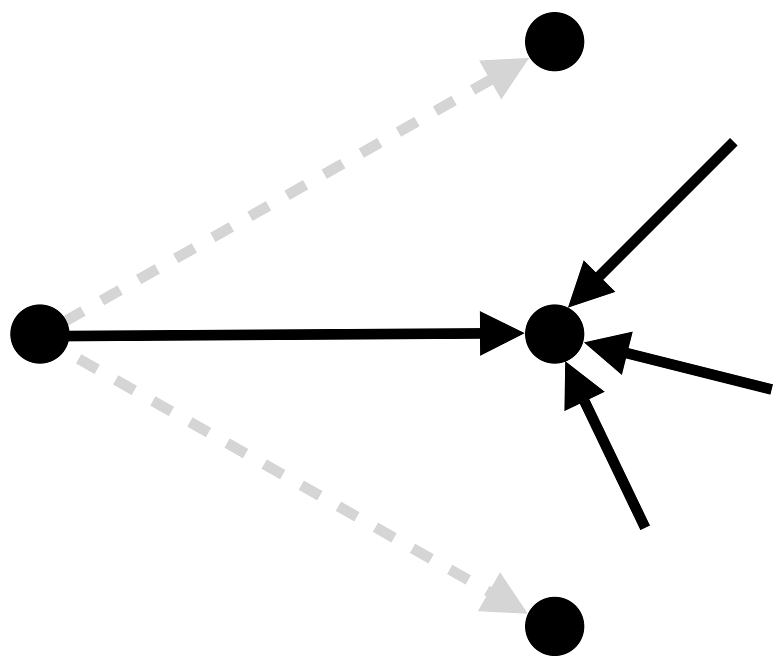


What factors drive edge formation?



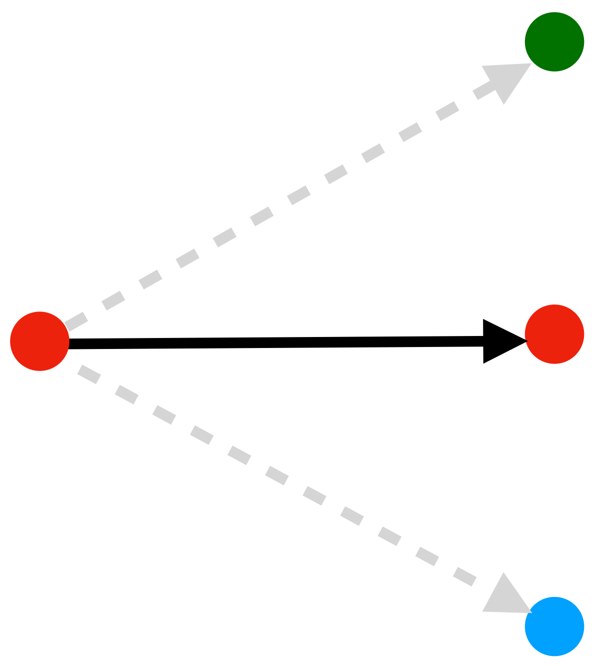
Preferential attachment

(Barabási & Albert, *Science* 1999)



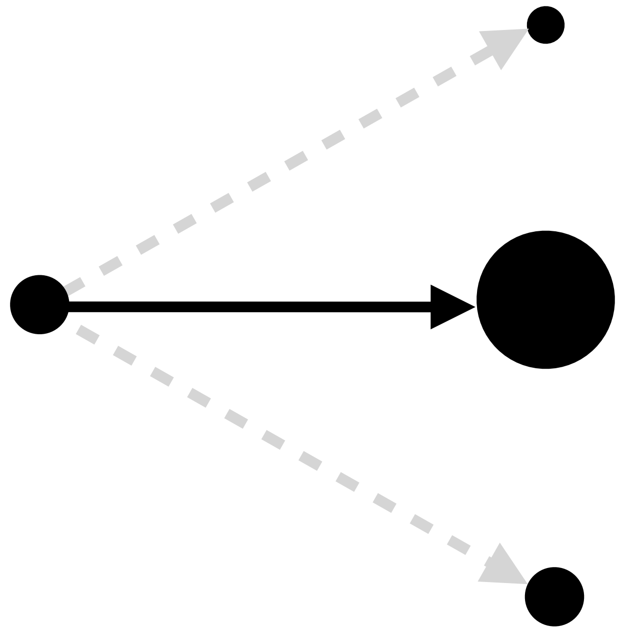
Homophily

(McPherson et al., *Annual Review of Sociology* 2001)
(Papadopoulos et al., *Nature* 2012)



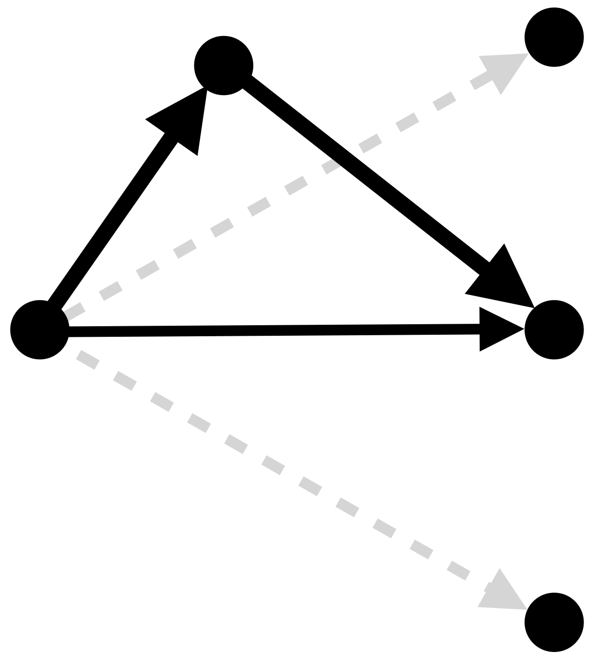
Fitness

(Bianconi & Barabási, *Europhysics Letters* 2001)
(Caldarelli et al., *Physical Review Letters* 2002)



Triadic closure

(Rapoport, *Bulletin of Mathematical Biophysics* 1953)
(Jin et al., *Physical Review E* 2001)



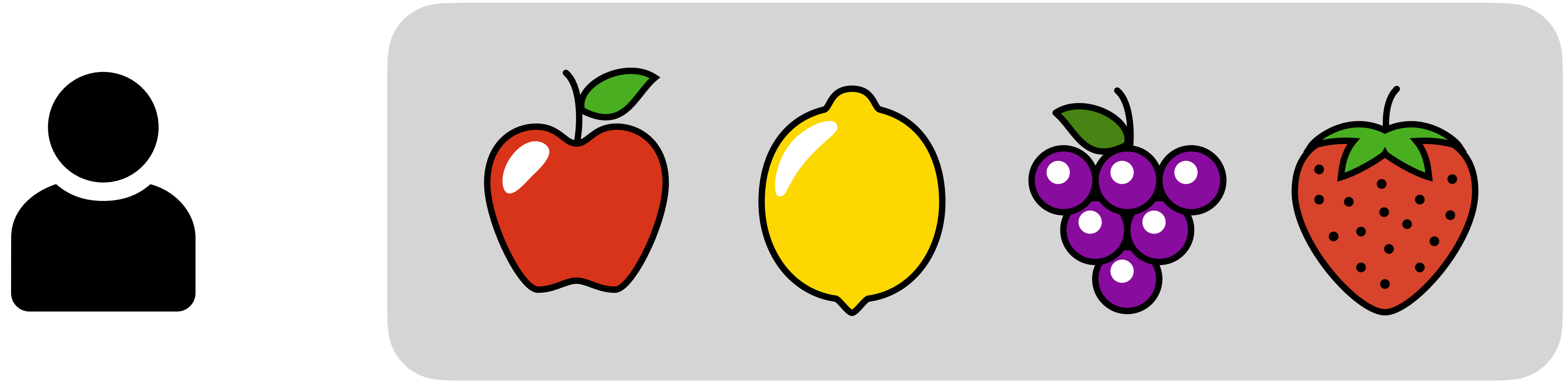
“Choosing to grow a graph”

(Overgoor et al., *SINM* ’19 & *WWW* ’19)
(Gupta & Porter, *arXiv* 2020)

"Choosing to grow a graph"

(Overgoor et al., *SINM* '19 & *WWW* '19)
(Gupta & Porter, *arXiv* 2020)

Traditional discrete choice:

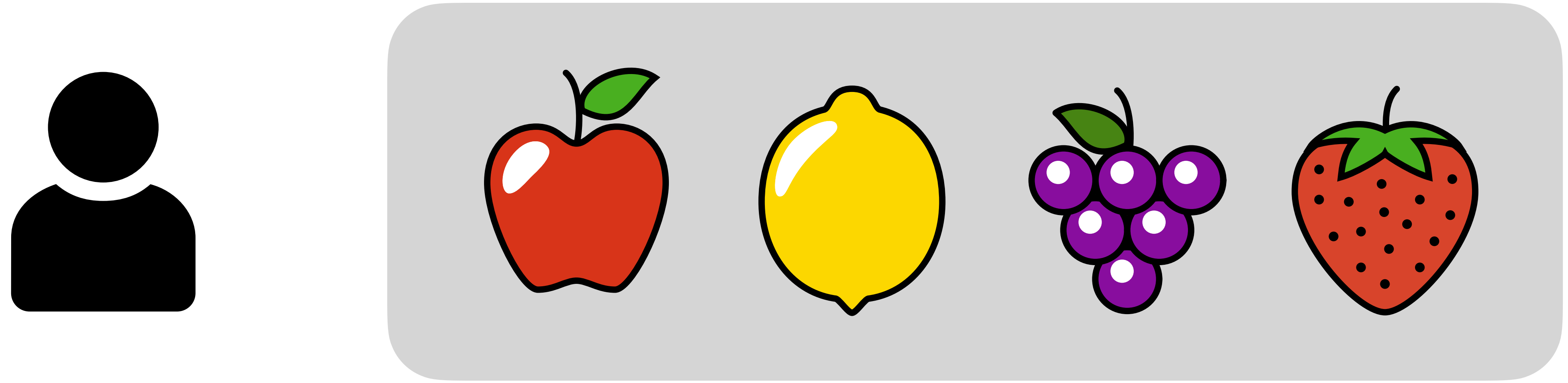


chooser *choice set*

“Choosing to grow a graph”

(Overgoor et al., *SINM* ’19 & *WWW* ’19)
(Gupta & Porter, *arXiv* 2020)

Traditional discrete choice:



chooser *choice set*

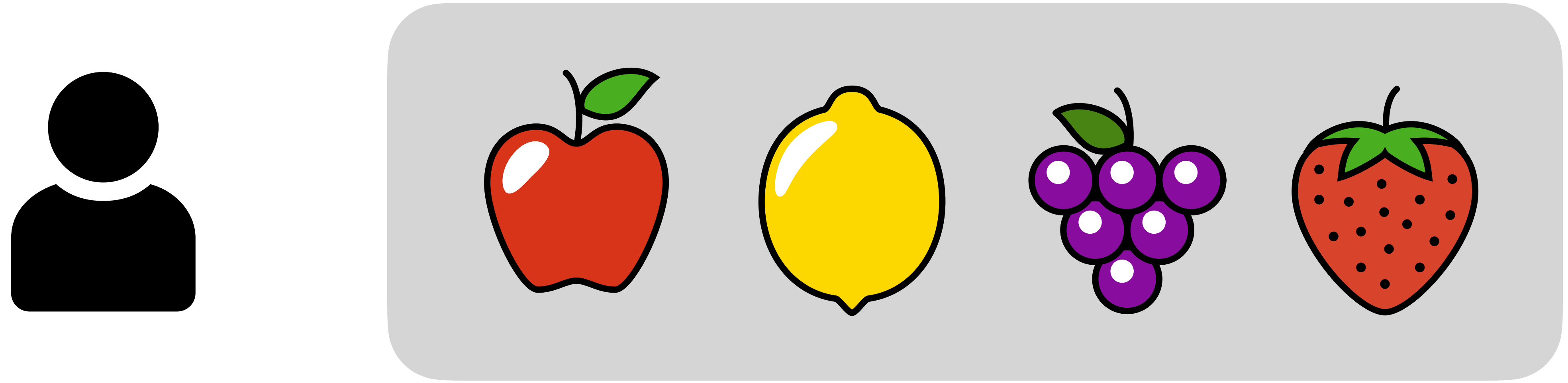
(under-explored in sociology)

(Bruch & Feinberg, *Annual Review of Sociology* 2017)

“Choosing to grow a graph”

(Overgoor et al., *SINM* ’19 & *WWW* ’19)
(Gupta & Porter, *arXiv* 2020)

Traditional discrete choice:

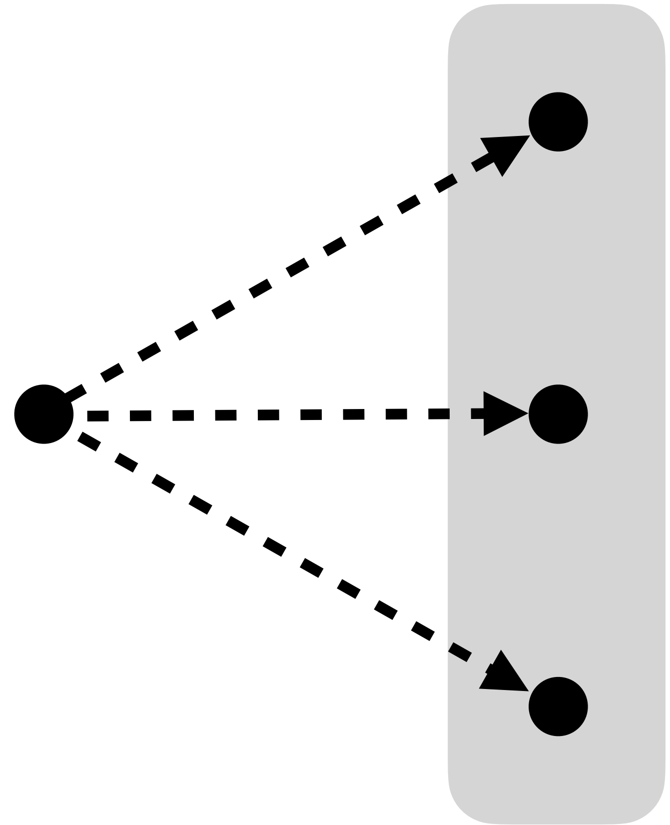


chooser *choice set*

(under-explored in sociology)
(Bruch & Feinberg, *Annual Review of Sociology* 2017)

in network growth

chooser

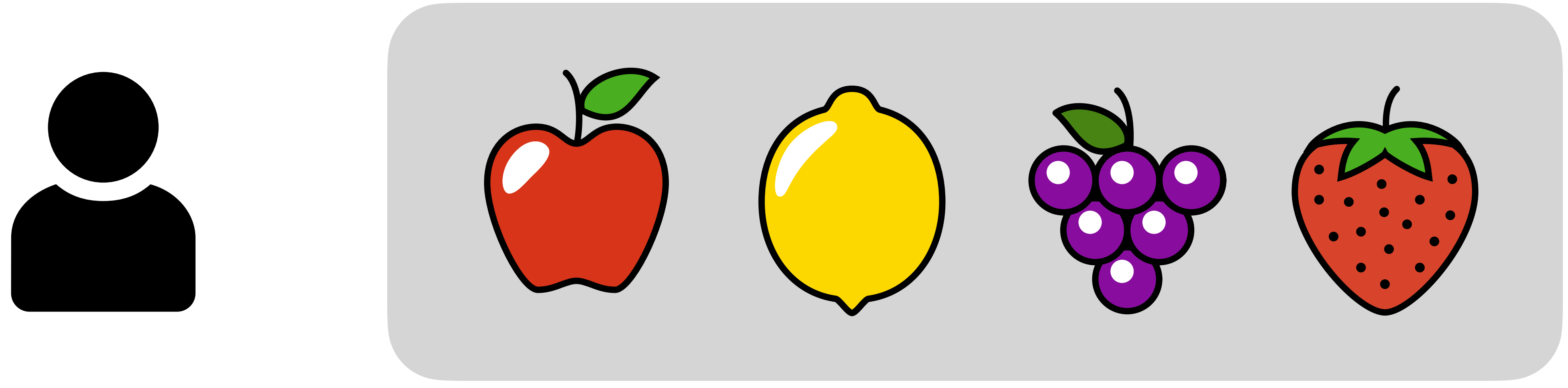


choice set

“Choosing to grow a graph”

(Overgoor et al., *SINM* ’19 & *WWW* ’19)
(Gupta & Porter, *arXiv* 2020)

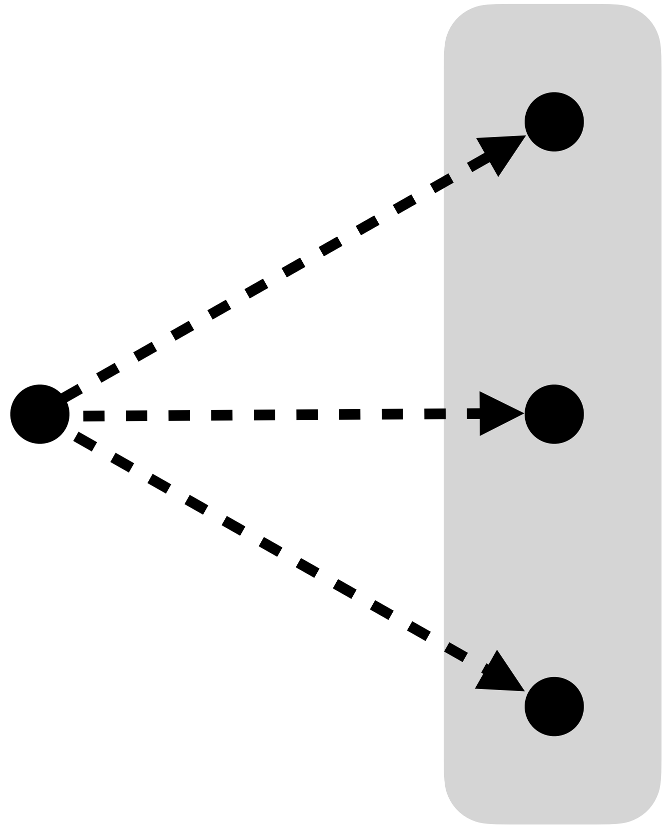
Traditional discrete choice:



chooser *choice set*

→ in network growth

chooser



choice set

(under-explored in sociology)
(Bruch & Feinberg, *Annual Review of Sociology* 2017)

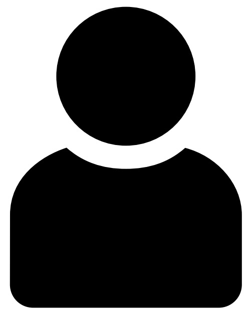
Key usage

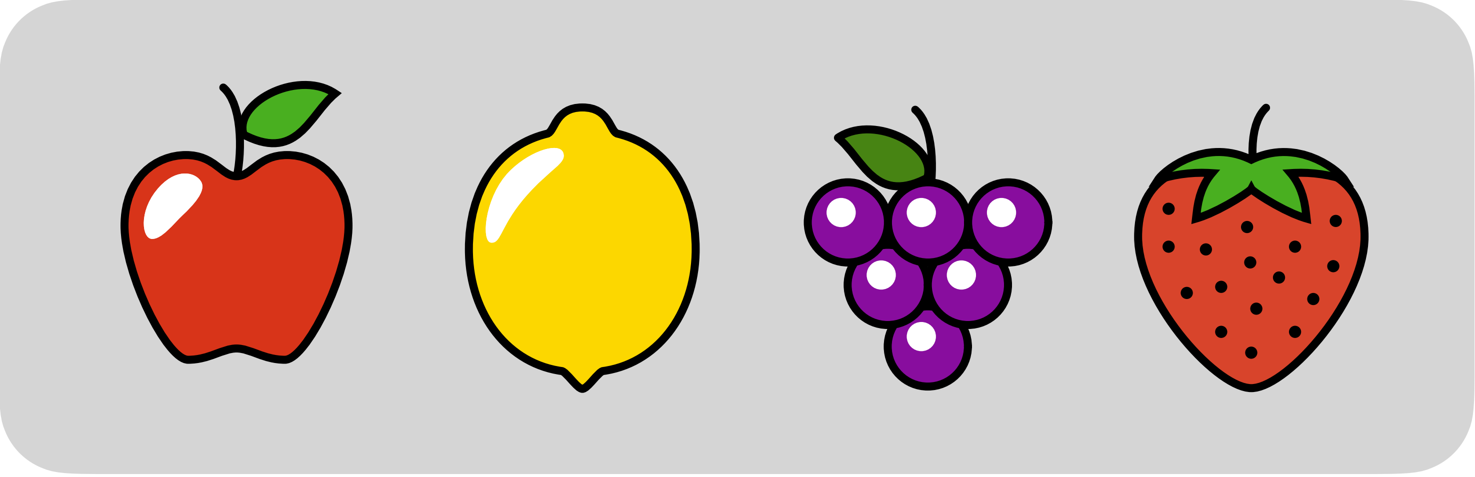
Timestamped edges
→ meaningful choice sets

“Choosing to grow a graph”

(Overgoor et al., *SINM* ’19 & *WWW* ’19)
(Gupta & Porter, *arXiv* 2020)

Traditional discrete choice:



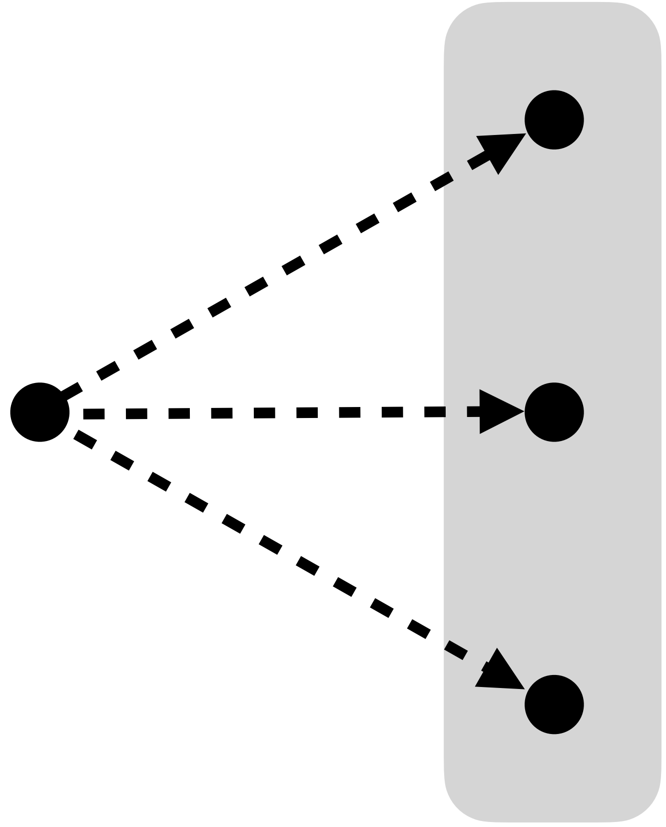


chooser *choice set*

(under-explored in sociology)
(Bruch & Feinberg, *Annual Review of Sociology* 2017)

in network growth

chooser



choice set

Key usage

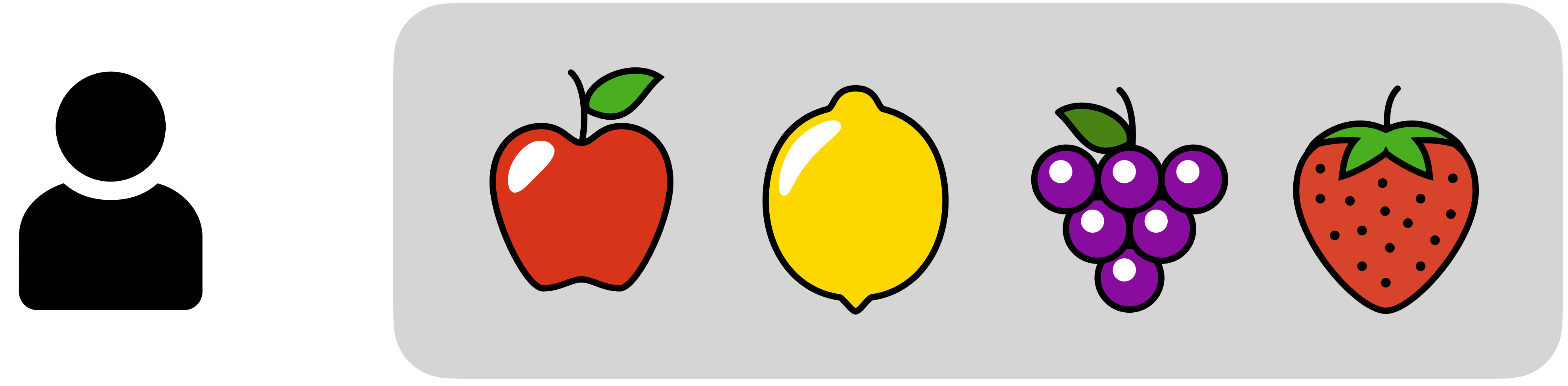
Timestamped edges
→ meaningful choice sets

Infer relative importance of edge formation mechanisms from data

“Choosing to grow a graph”

(Overgoor et al., *SINM* ’19 & *WWW* ’19)
(Gupta & Porter, *arXiv* 2020)

Traditional discrete choice:

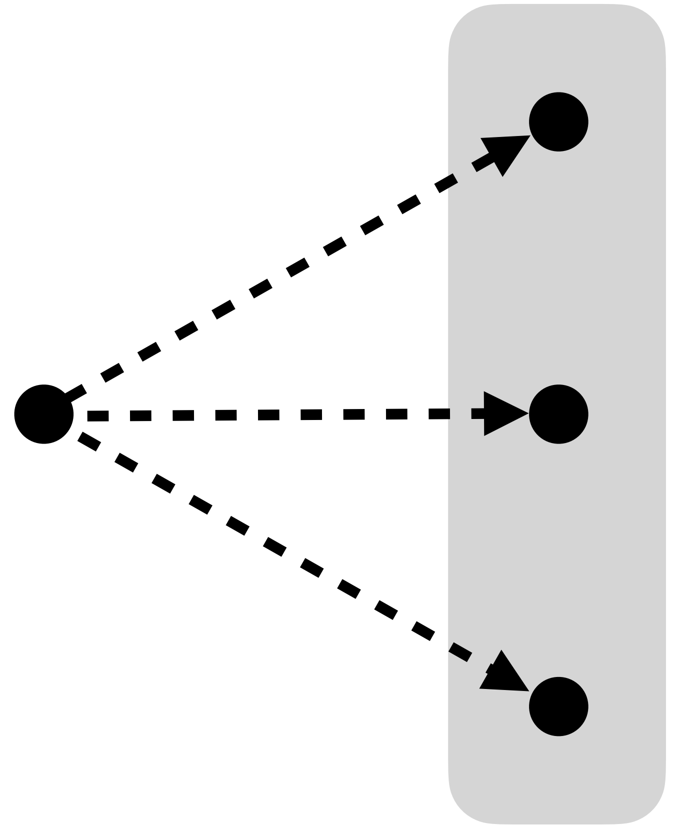


chooser *choice set*

(under-explored in sociology)
(Bruch & Feinberg, *Annual Review of Sociology* 2017)

in network growth

chooser



choice set

Key usage

Timestamped edges
→ meaningful choice sets

Infer relative importance of edge formation mechanisms from data

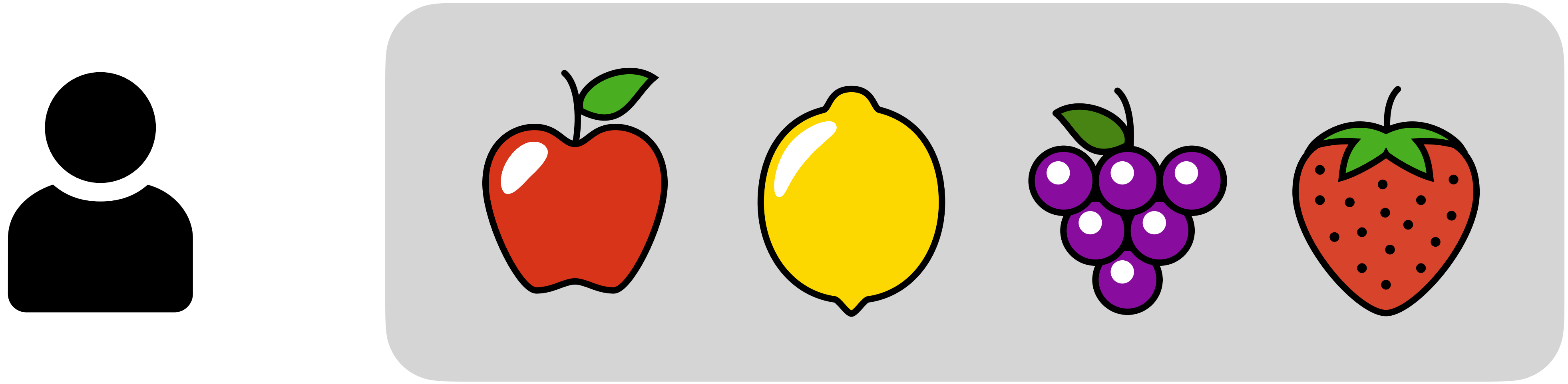
$$\Pr(i, C) = \frac{\exp(\theta^T x_i)}{\sum_{j \in C} \exp(\theta^T x_j)}$$

Multinomial logit (MNL) (McFadden, 1973)

"Choosing to grow a graph"

(Overgoor et al., *SINM* '19 & *WWW* '19)
(Gupta & Porter, *arXiv* 2020)

Traditional discrete choice:

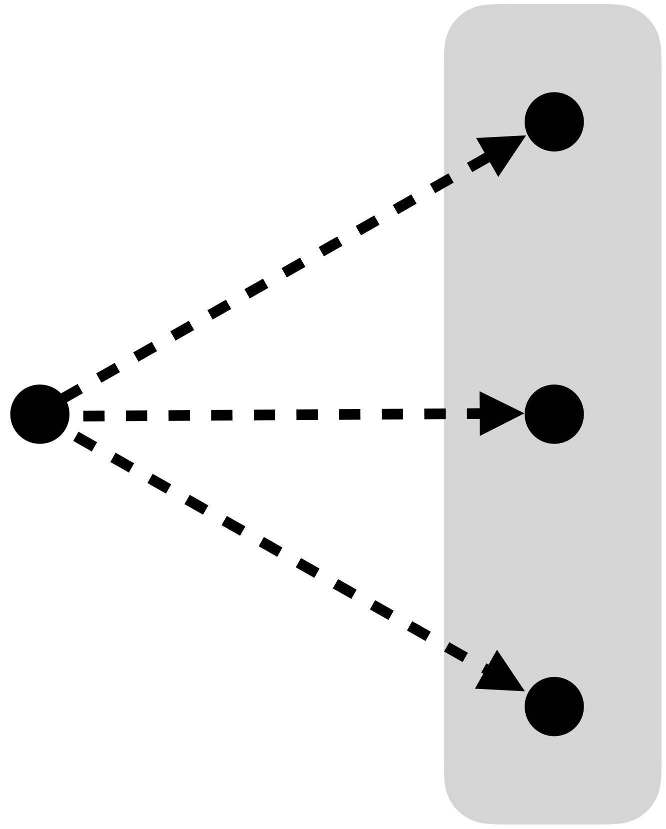


chooser *choice set*

(under-explored in sociology)
(Bruch & Feinberg, *Annual Review of Sociology* 2017)

in network growth

chooser



choice set

Key usage

Timestamped edges
→ meaningful choice sets

Infer relative importance of edge formation mechanisms from data

$$\Pr(i, C) = \frac{\exp(\theta^T x_i)}{\sum_{j \in C} \exp(\theta^T x_j)}$$

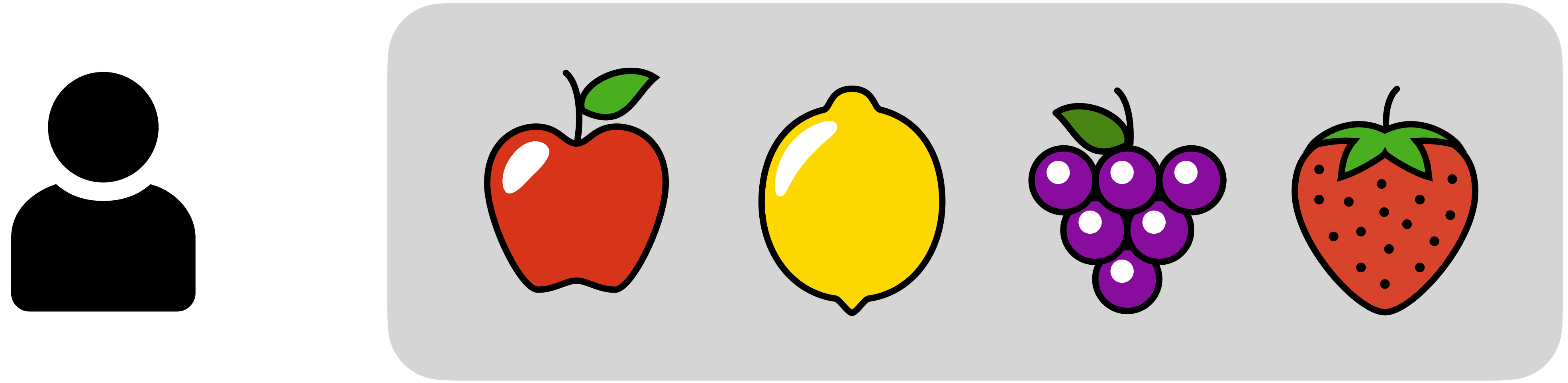
node

Multinomial logit (MNL) (McFadden, 1973)

“Choosing to grow a graph”

(Overgoor et al., *SINM* ’19 & *WWW* ’19)
(Gupta & Porter, *arXiv* 2020)

Traditional discrete choice:

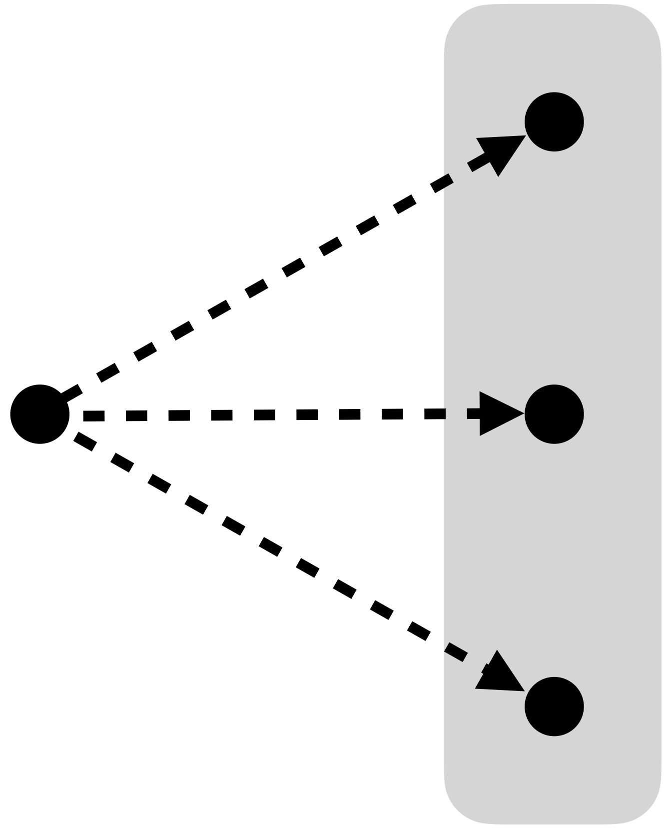


chooser *choice set*

(under-explored in sociology)
(Bruch & Feinberg, *Annual Review of Sociology* 2017)

in network growth

chooser



choice set

Key usage

Timestamped edges
→ meaningful choice sets

Infer relative importance of edge formation mechanisms from data

$$\Pr(i, C) = \frac{\exp(\theta^T x_i)}{\sum_{j \in C} \exp(\theta^T x_j)}$$

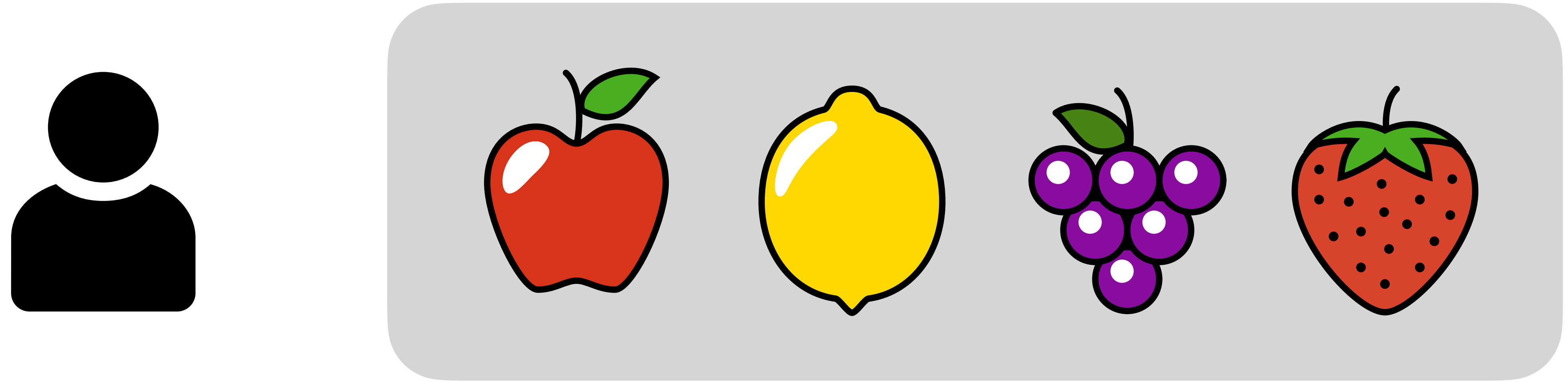
node choice set

Multinomial logit (MNL) (McFadden, 1973)

“Choosing to grow a graph”

(Overgoor et al., *SINM* ’19 & *WWW* ’19)
(Gupta & Porter, *arXiv* 2020)

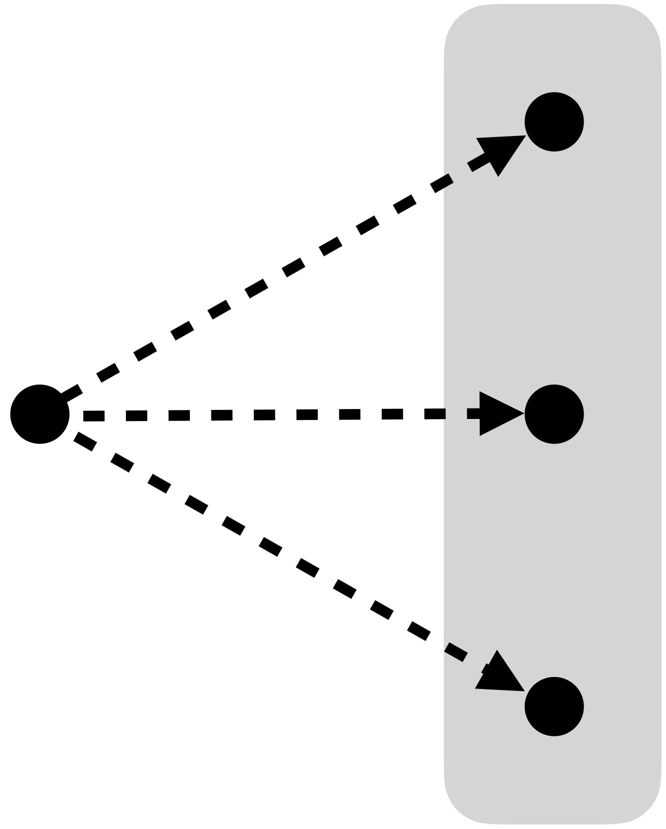
Traditional discrete choice:



chooser *choice set*

in network growth

chooser



choice set

(under-explored in sociology)
(Bruch & Feinberg, *Annual Review of Sociology* 2017)

Key usage

Timestamped edges
→ meaningful choice sets

Infer relative importance of edge formation mechanisms from data

preferences

$$\Pr(i, C) = \frac{\exp(\theta^T x_i)}{\sum_{j \in C} \exp(\theta^T x_j)}$$

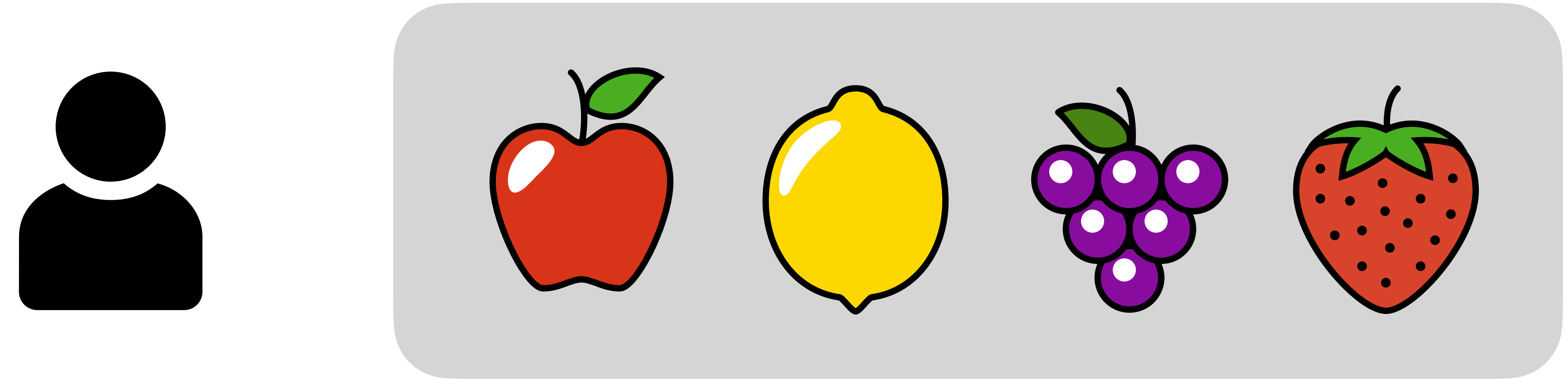
node choice set

Multinomial logit (MNL) (McFadden, 1973)

“Choosing to grow a graph”

(Overgoor et al., *SINM* ’19 & *WWW* ’19)
(Gupta & Porter, *arXiv* 2020)

Traditional discrete choice:

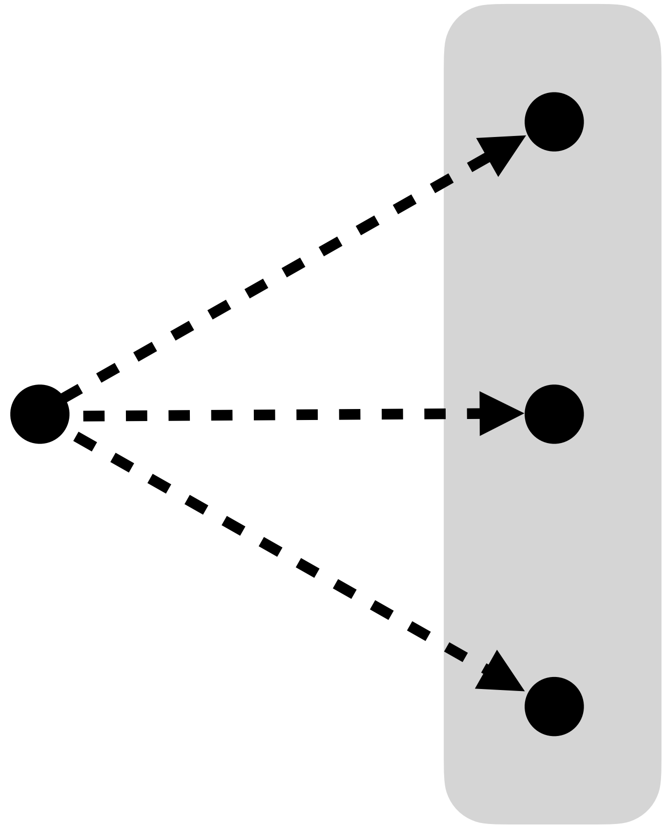


chooser *choice set*

(under-explored in sociology)
(Bruch & Feinberg, *Annual Review of Sociology* 2017)

→ in network growth

chooser



choice set

Key usage

Timestamped edges
→ meaningful choice sets

Infer relative importance of edge formation mechanisms from data

preferences node features

$$\Pr(i, C) = \frac{\exp(\theta^T x_i)}{\sum_{j \in C} \exp(\theta^T x_j)}$$

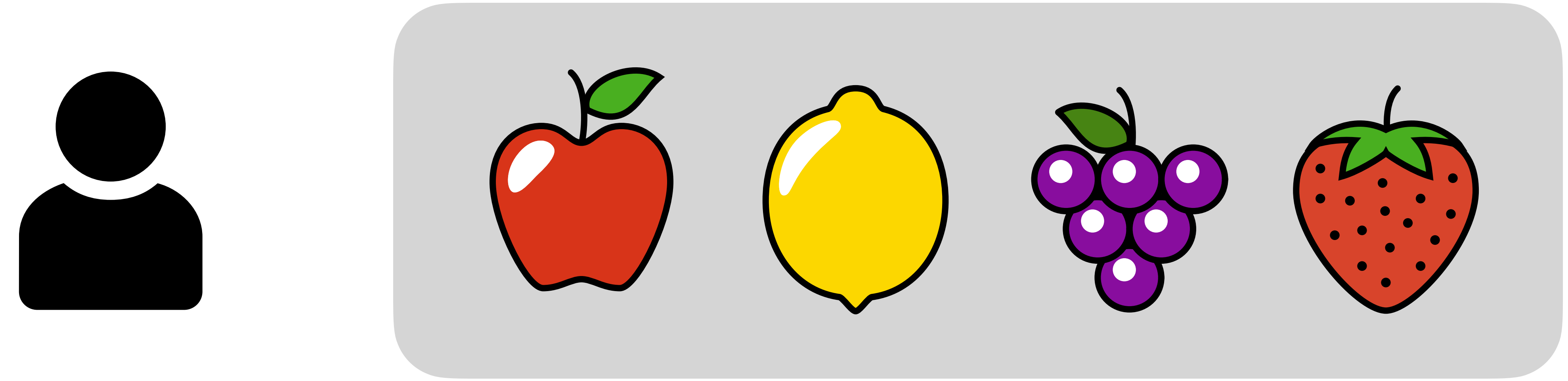
node choice set

Multinomial logit (MNL) (McFadden, 1973)

"Choosing to grow a graph"

(Overgoor et al., *SINM* '19 & *WWW* '19)
(Gupta & Porter, *arXiv* 2020)

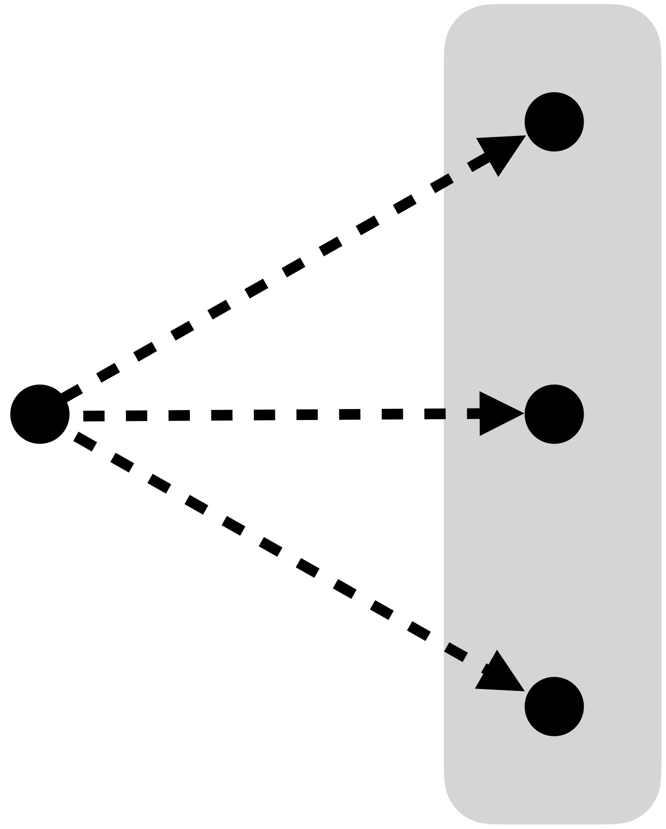
Traditional discrete choice:



chooser *choice set*

(under-explored in sociology)
(Bruch & Feinberg, *Annual Review of Sociology* 2017)

→ in network growth



Key usage

Timestamped edges
→ meaningful choice sets

Infer relative importance of edge formation mechanisms from data

$$\Pr(i, C) = \frac{\exp(\theta^T x_i)}{\sum_{j \in C} \exp(\theta^T x_j)}$$

preferences — node features (similarity, in-degree, fitness…)

node — choice set

Multinomial logit (MNL) (McFadden, 1973)

The choice set affects preferences

The choice set affects preferences

Context effects

(Huber et al., *Journal of Consumer Research* 1982)
(Simonson & Tversky, *Journal of Marketing Research* 1992)

The choice set affects preferences

Context effects

(Huber et al., *Journal of Consumer Research* 1982)
(Simonson & Tversky, *Journal of Marketing Research* 1992)

e.g., *compromise effect*:

(Simonson, *Journal of Consumer Research* 1989)

The choice set affects preferences

Context effects

(Huber et al., *Journal of Consumer Research* 1982)
(Simonson & Tversky, *Journal of Marketing Research* 1992)

e.g., *compromise effect*:

(Simonson, *Journal of Consumer Research* 1989)

The choice set affects preferences

Context effects

(Huber et al., *Journal of Consumer Research* 1982)
(Simonson & Tversky, *Journal of Marketing Research* 1992)

e.g., *compromise effect*:

(Simonson, *Journal of Consumer Research* 1989)

The choice set affects preferences

Context effects

(Huber et al., *Journal of Consumer Research* 1982)
(Simonson & Tversky, *Journal of Marketing Research* 1992)

e.g., *compromise effect*:

(Simonson, *Journal of Consumer Research* 1989)

★

The choice set affects preferences

Context effects

(Huber et al., *Journal of Consumer Research* 1982)
(Simonson & Tversky, *Journal of Marketing Research* 1992)

e.g., *compromise effect*:

(Simonson, *Journal of Consumer Research* 1989)

The choice set affects preferences

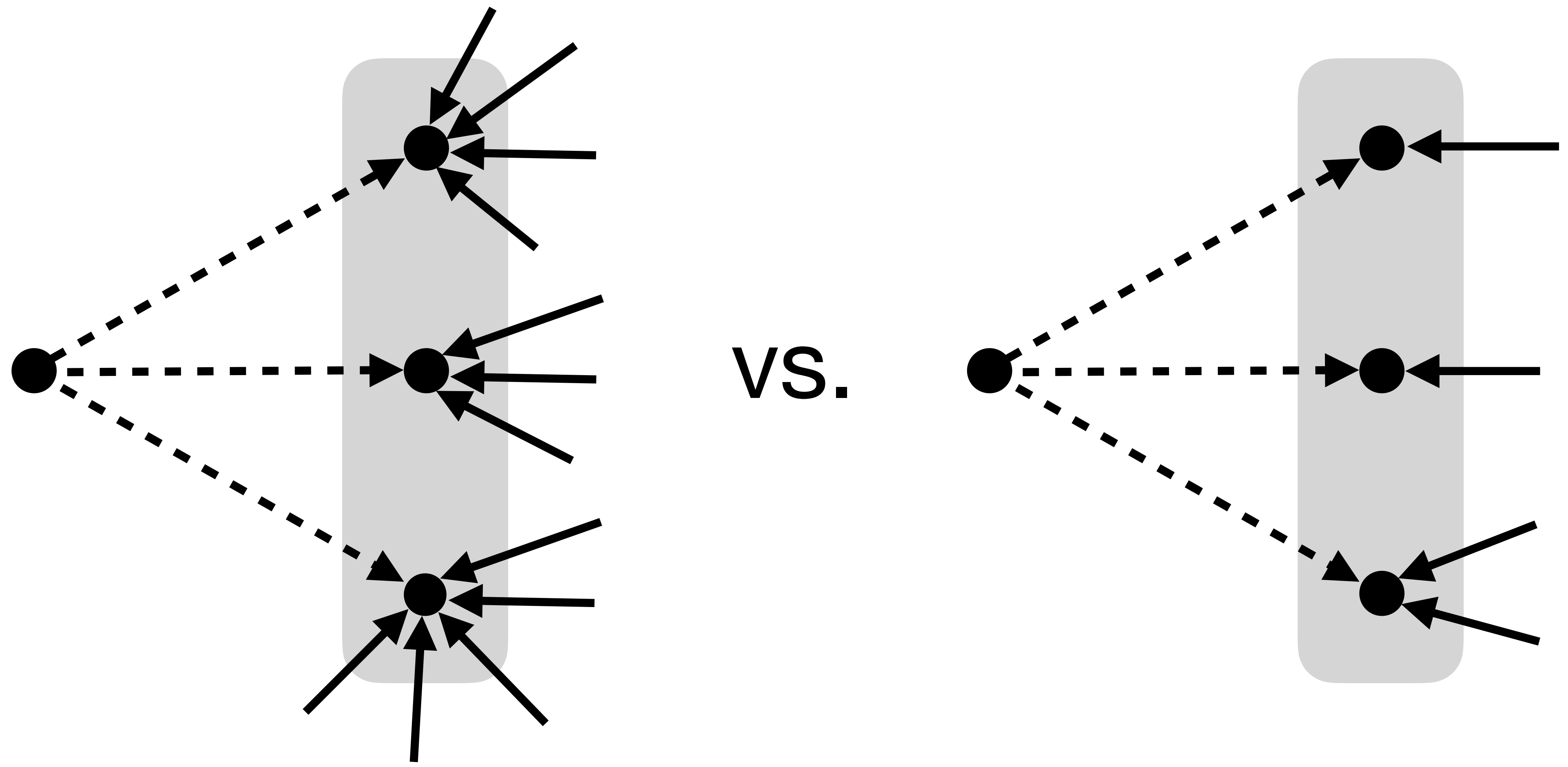
Context effects

(Huber et al., *Journal of Consumer Research* 1982)
(Simonson & Tversky, *Journal of Marketing Research* 1992)

e.g., *compromise effect*:

(Simonson, *Journal of Consumer Research* 1989)

In networks
e.g., how do preferences change
when choosing from a popular group?



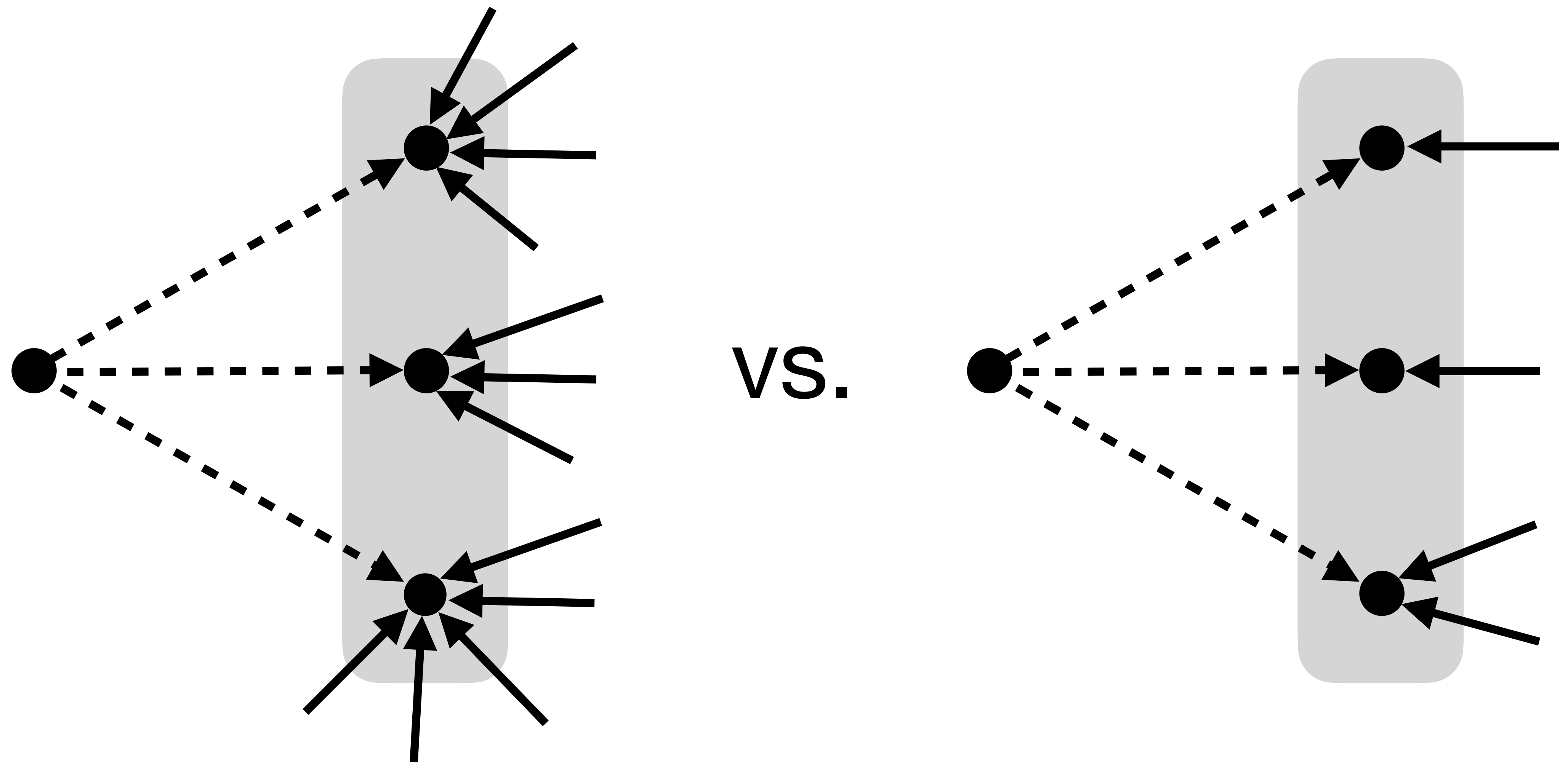
The choice set affects preferences

Context effects

(Huber et al., *Journal of Consumer Research* 1982)
(Simonson & Tversky, *Journal of Marketing Research* 1992)

e.g., *compromise effect*:
(Simonson, *Journal of Consumer Research* 1989)

In networks
e.g., how do preferences change when choosing from a popular group?



Our model:

Linear context logit (LCL)

$$\Pr(i, C) = \frac{\exp\left(\left[\theta + A x_C\right]^T x_i\right)}{\sum_{j \in C} \exp\left(\left[\theta + A x_C\right]^T x_j\right)}$$

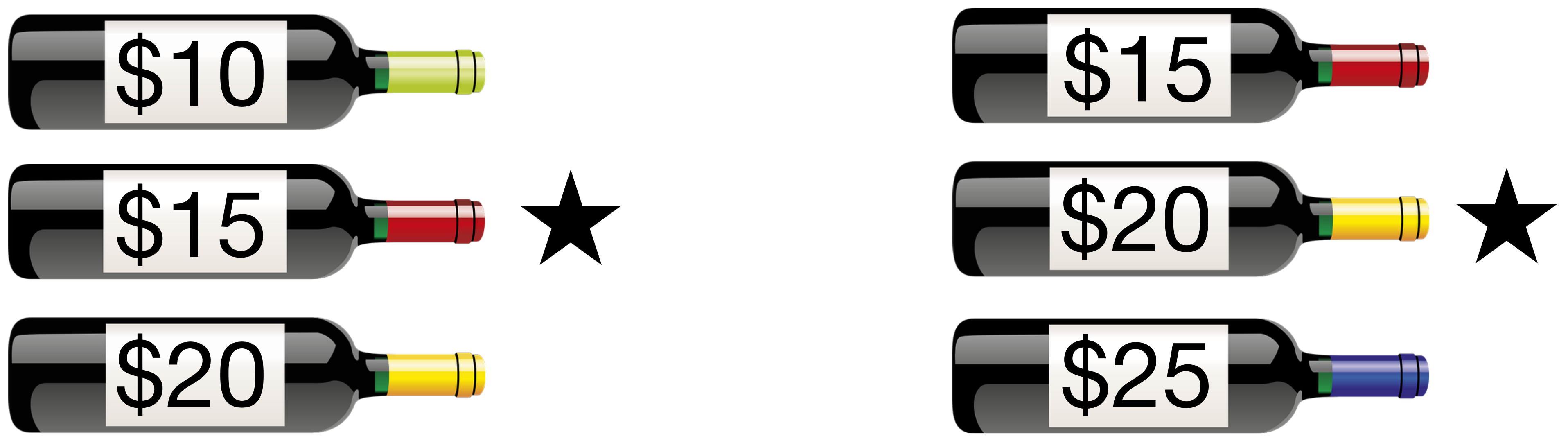
The choice set affects preferences

Context effects

(Huber et al., *Journal of Consumer Research* 1982)
(Simonson & Tversky, *Journal of Marketing Research* 1992)

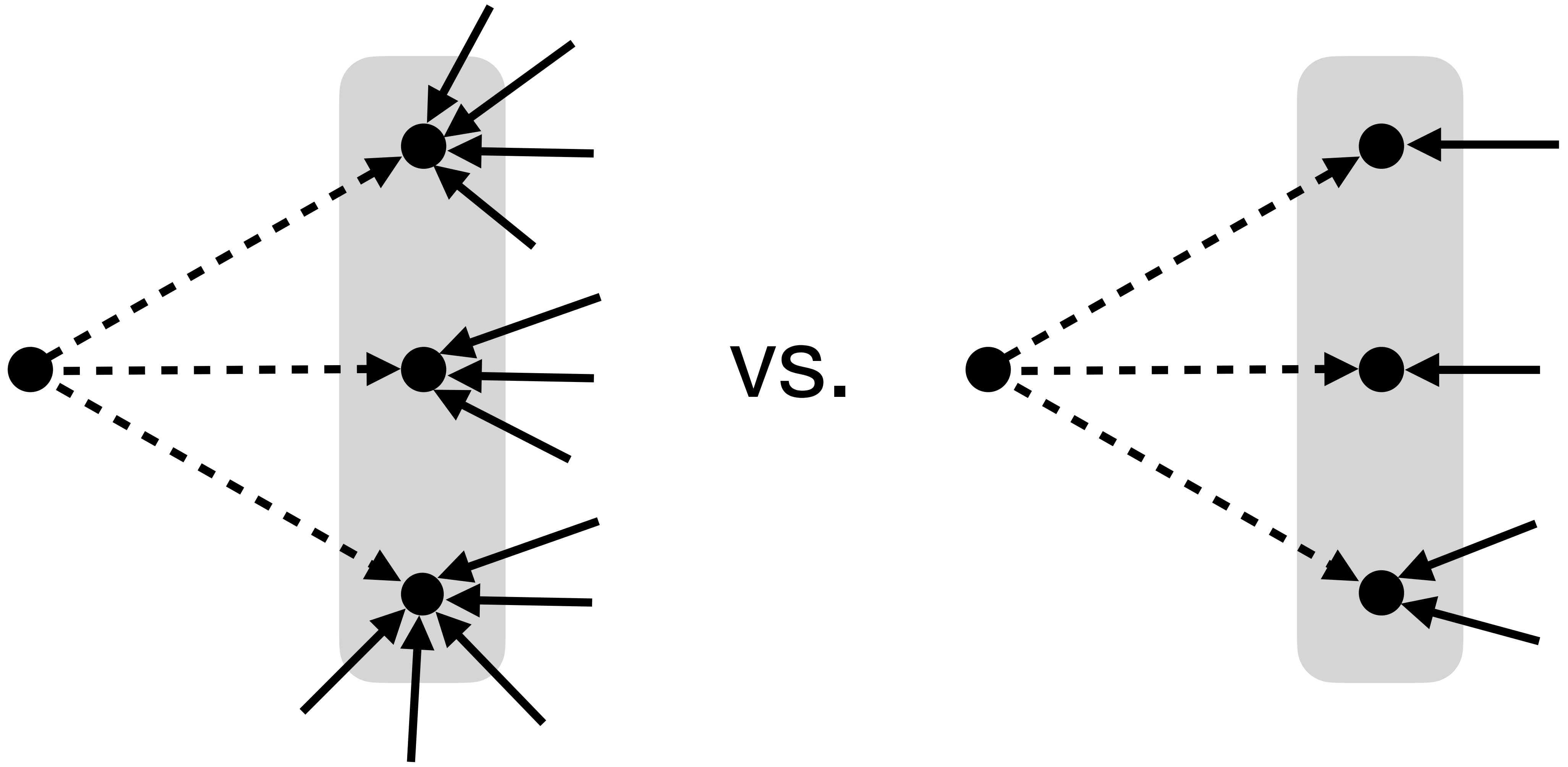
e.g., *compromise effect*:

(Simonson, *Journal of Consumer Research* 1989)



In networks

e.g., how do preferences change when choosing from a popular group?



Our model:

Linear context logit (LCL)

base preferences

$$\Pr(i, C) = \frac{\exp([\theta + A x_C]^T x_i)}{\sum_{j \in C} \exp([\theta + A x_C]^T x_j)}$$

The choice set affects preferences

Context effects

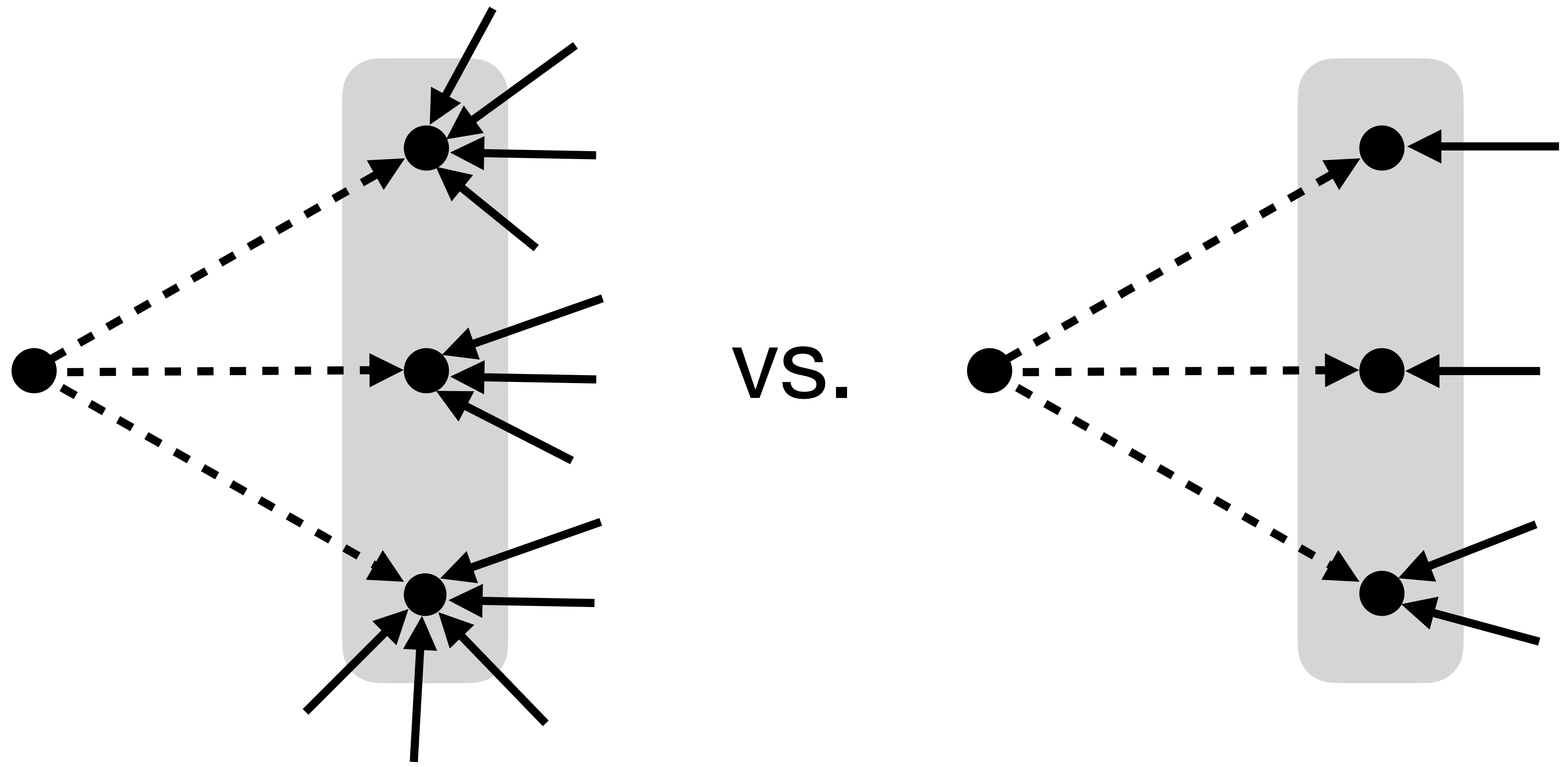
(Huber et al., *Journal of Consumer Research* 1982)
(Simonson & Tversky, *Journal of Marketing Research* 1992)

e.g., *compromise effect*:

(Simonson, *Journal of Consumer Research* 1989)

In networks

e.g., how do preferences change when choosing from a popular group?



Our model:

Linear context logit (LCL)

base preferences

context effect matrix

$$\Pr(i, C) = \frac{\exp([\theta + A x_C]^T x_i)}{\sum_{j \in C} \exp([\theta + A x_C]^T x_j)}$$

The choice set affects preferences

Context effects

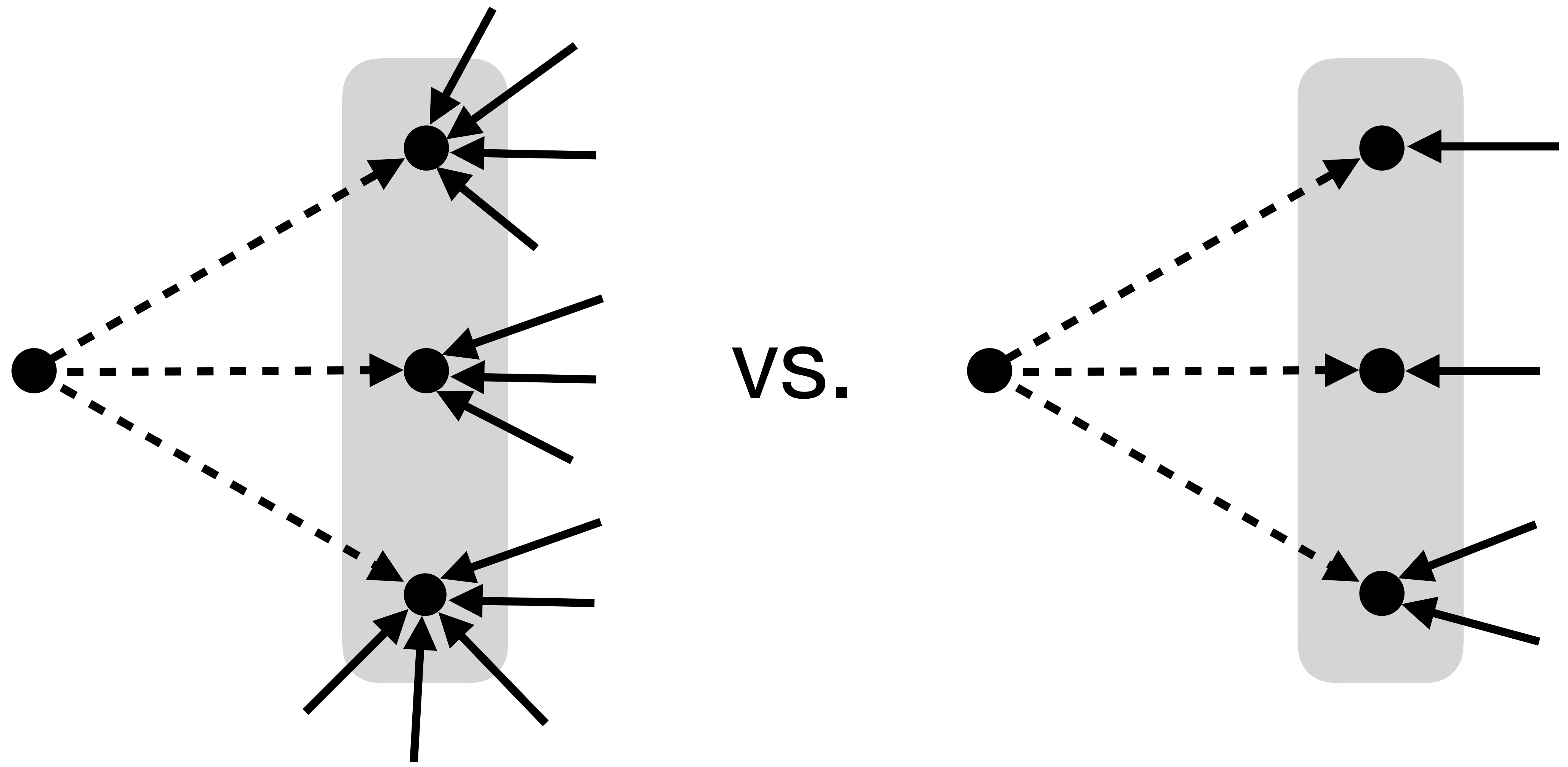
(Huber et al., *Journal of Consumer Research* 1982)
(Simonson & Tversky, *Journal of Marketing Research* 1992)

e.g., *compromise effect*:

(Simonson, *Journal of Consumer Research* 1989)

In networks

e.g., how do preferences change when choosing from a popular group?



Our model:

Linear context logit (LCL)

base preferences

context effect matrix

mean features over choice set

$$\Pr(i, C) = \frac{\exp([\theta + A x_C]^T x_i)}{\sum_{j \in C} \exp([\theta + A x_C]^T x_j)}$$

Choosing to close triangles

Triadic closure offers small choice sets
→ tractable inference
→ varied choice sets

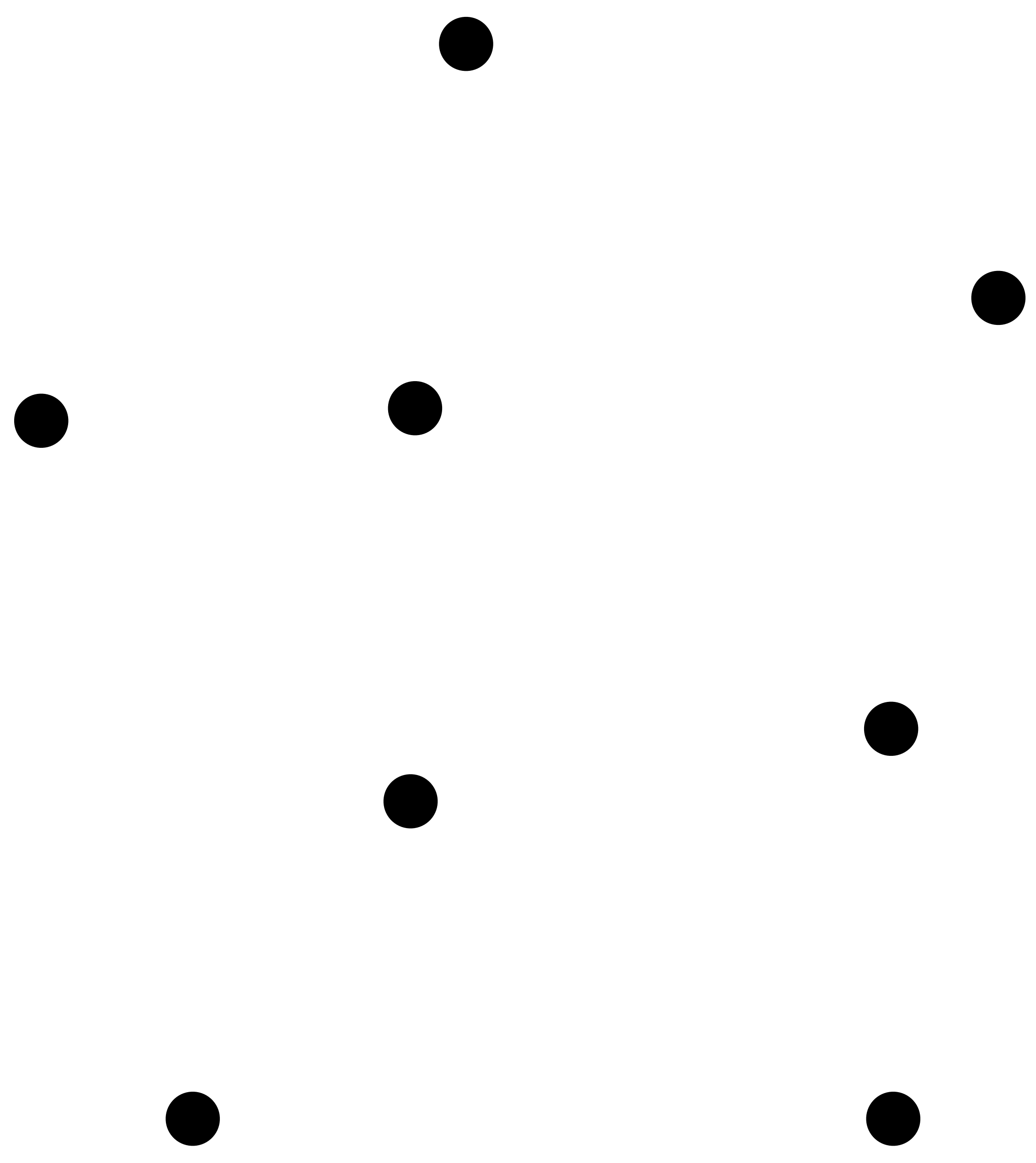
Choosing to close triangles

Triadic closure offers small choice sets
→ tractable inference
→ varied choice sets

Our data
Timestamped edges
(including repeats)

Choosing to close triangles

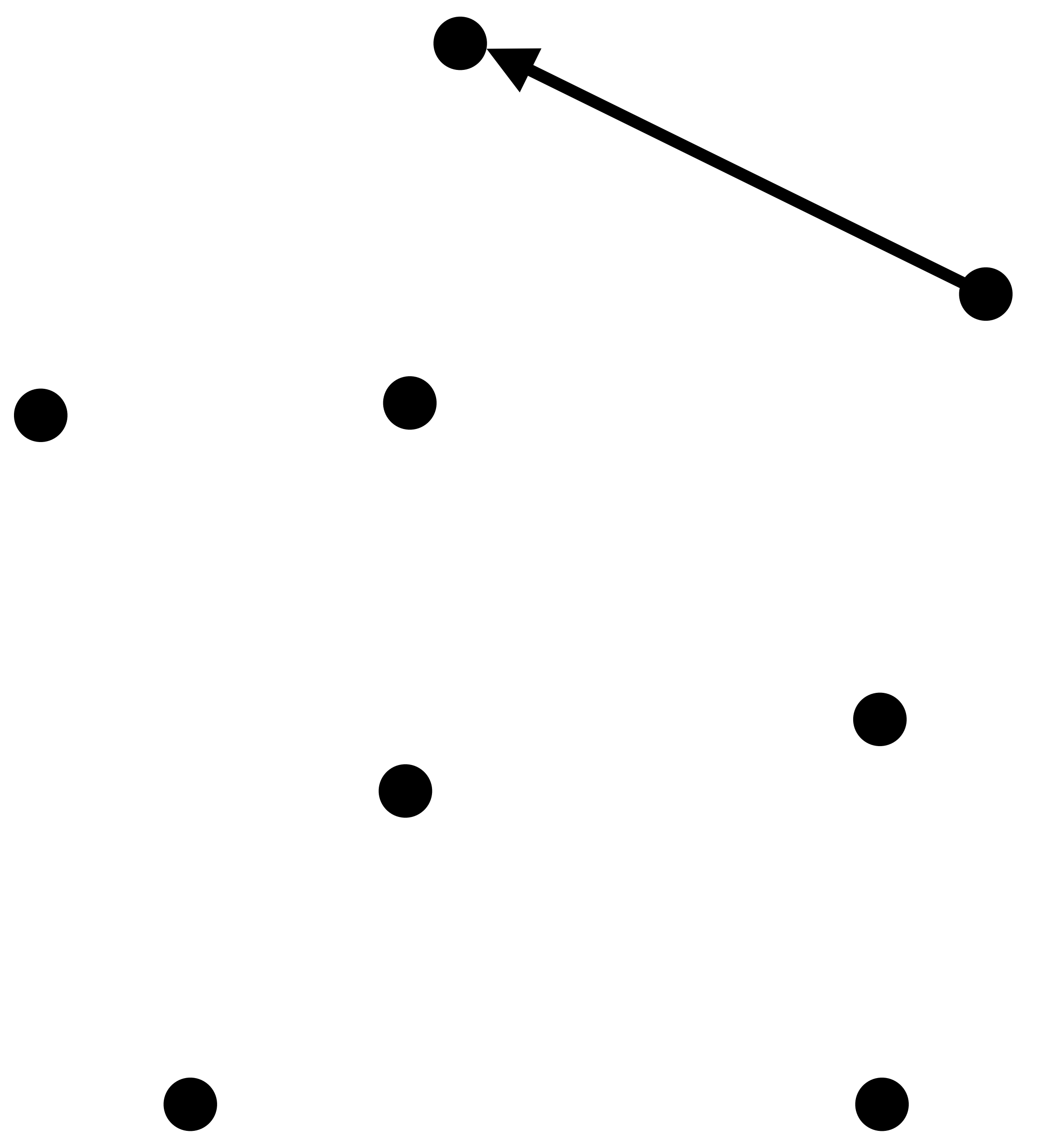
Triadic closure offers small choice sets
→ tractable inference
→ varied choice sets



Our data
Timestamped edges
(including repeats)

Choosing to close triangles

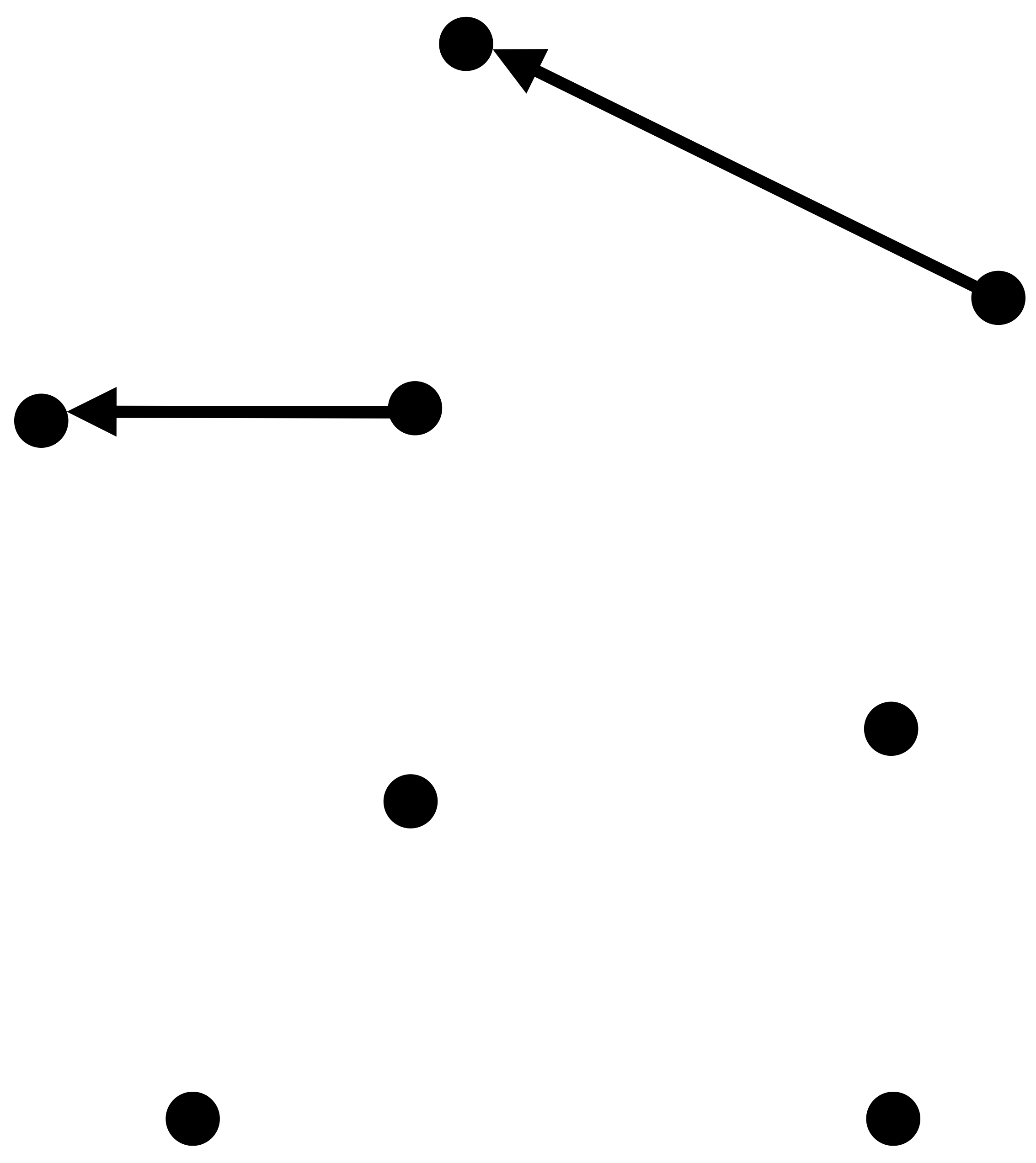
Triadic closure offers small choice sets
→ tractable inference
→ varied choice sets



Our data
Timestamped edges
(including repeats)

Choosing to close triangles

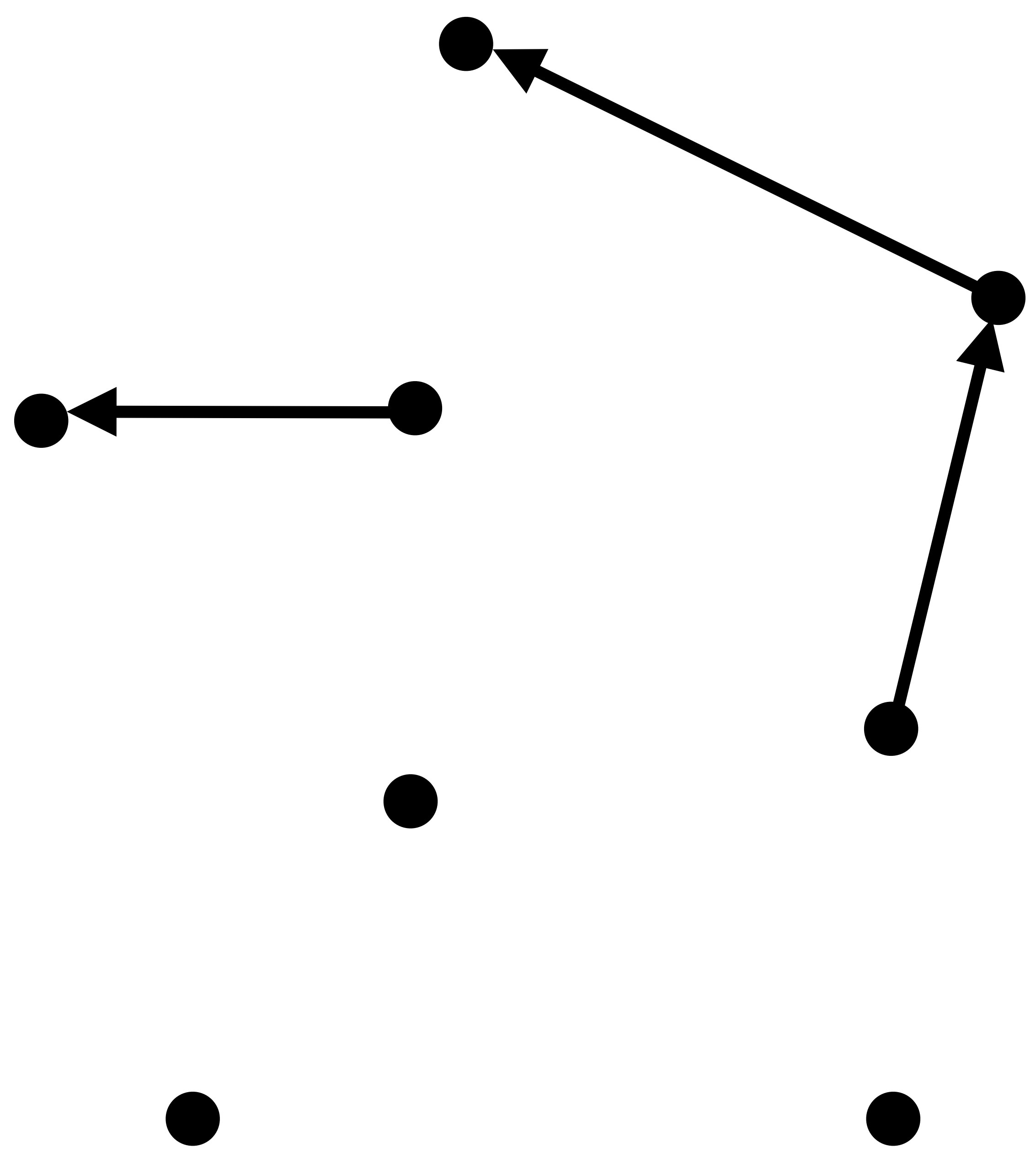
Triadic closure offers small choice sets
→ tractable inference
→ varied choice sets



Our data
Timestamped edges
(including repeats)

Choosing to close triangles

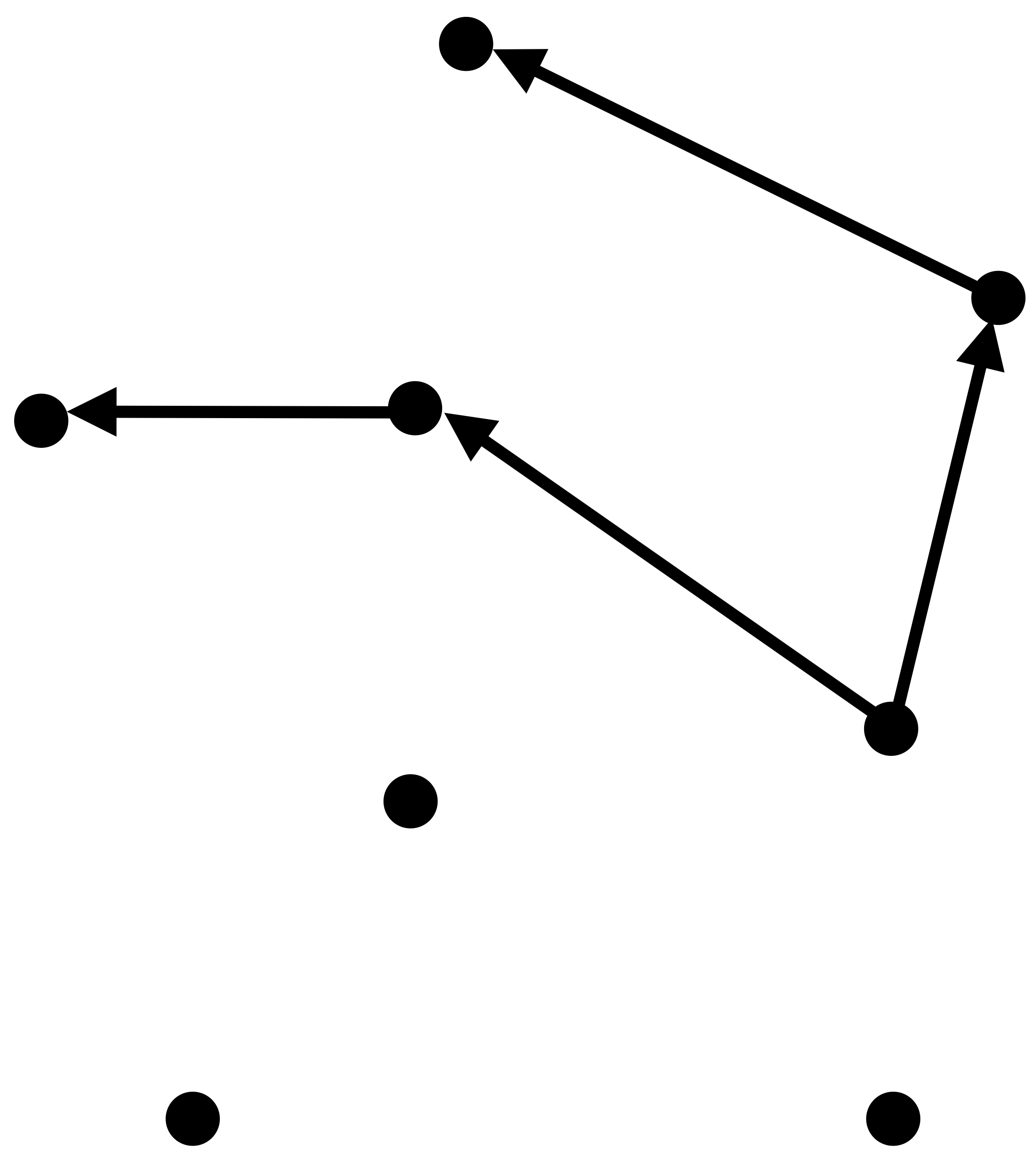
Triadic closure offers small choice sets
→ tractable inference
→ varied choice sets



Our data
Timestamped edges
(including repeats)

Choosing to close triangles

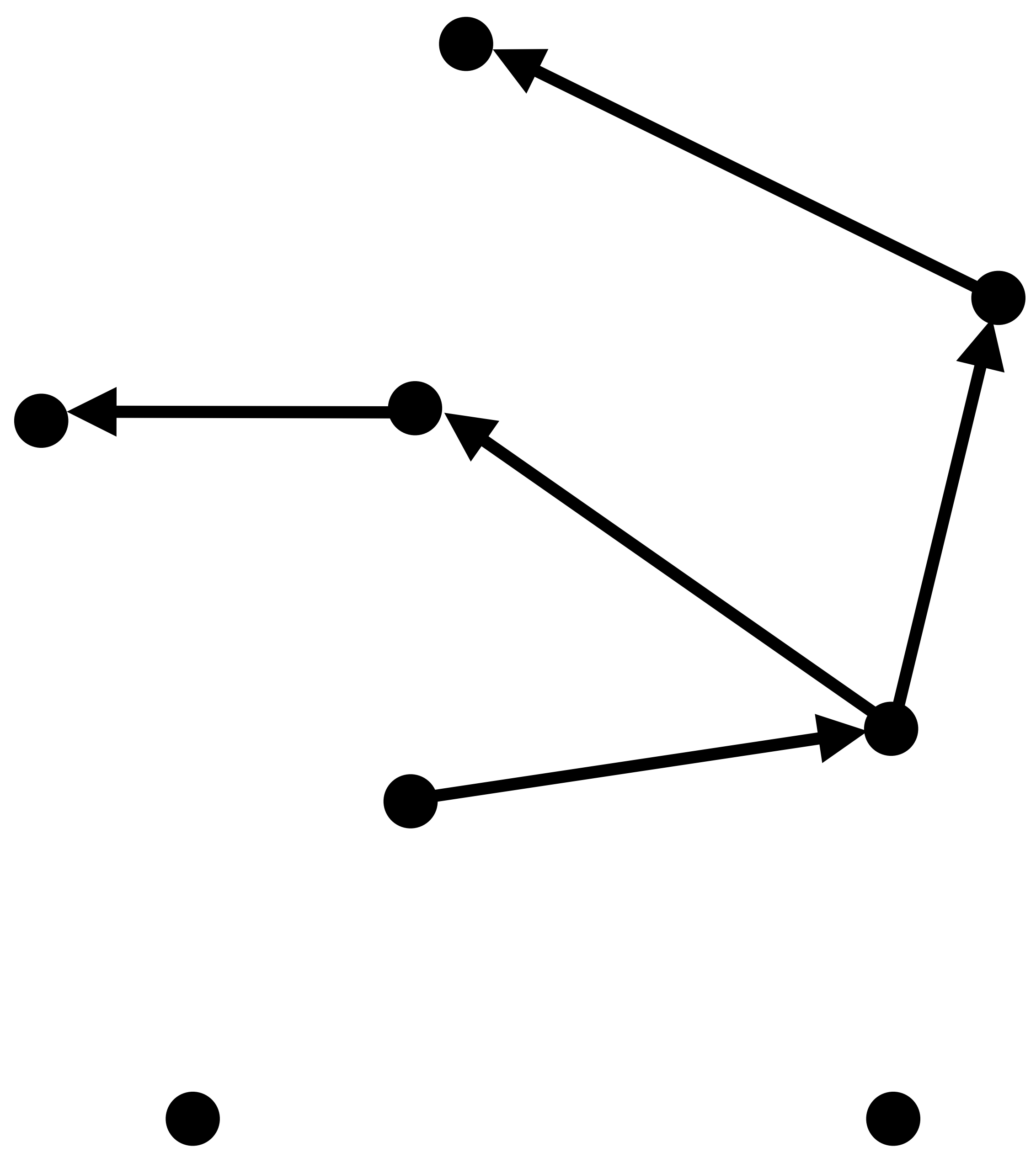
Triadic closure offers small choice sets
→ tractable inference
→ varied choice sets



Our data
Timestamped edges
(including repeats)

Choosing to close triangles

Triadic closure offers small choice sets
→ tractable inference
→ varied choice sets

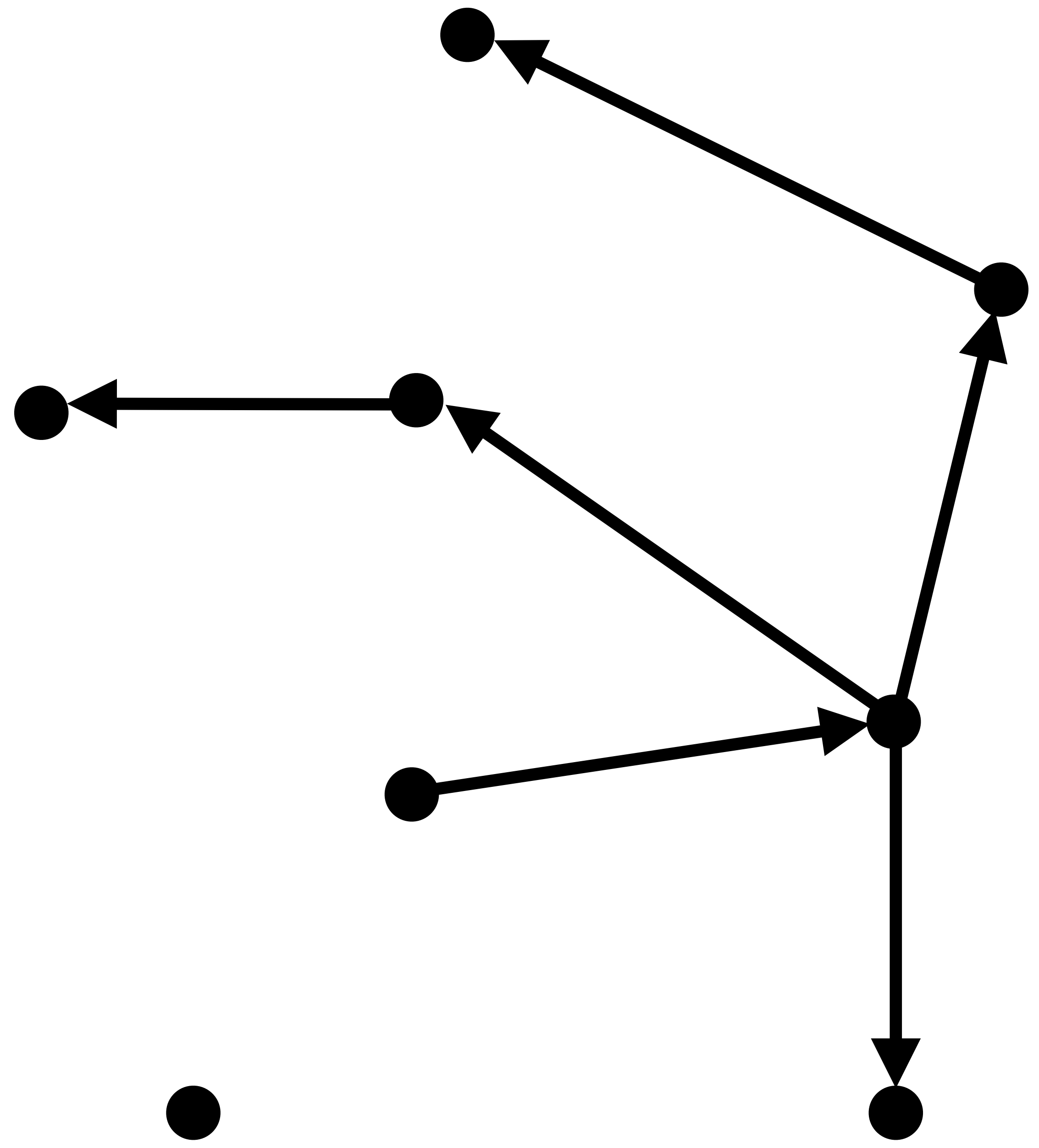


Our data
Timestamped edges
(including repeats)

Choosing to close triangles

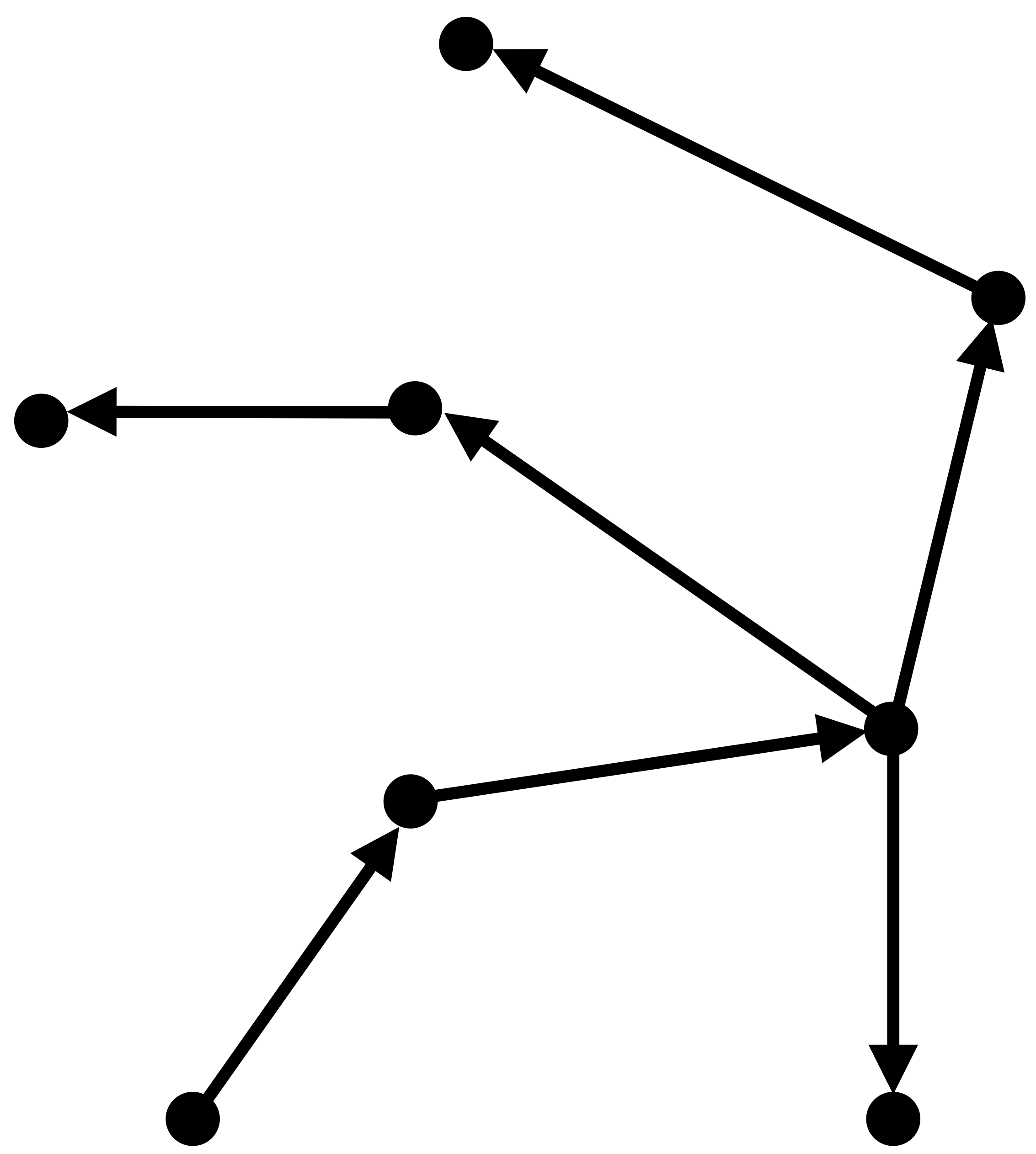
Triadic closure offers small choice sets
→ tractable inference
→ varied choice sets

Our data
Timestamped edges
(including repeats)



Choosing to close triangles

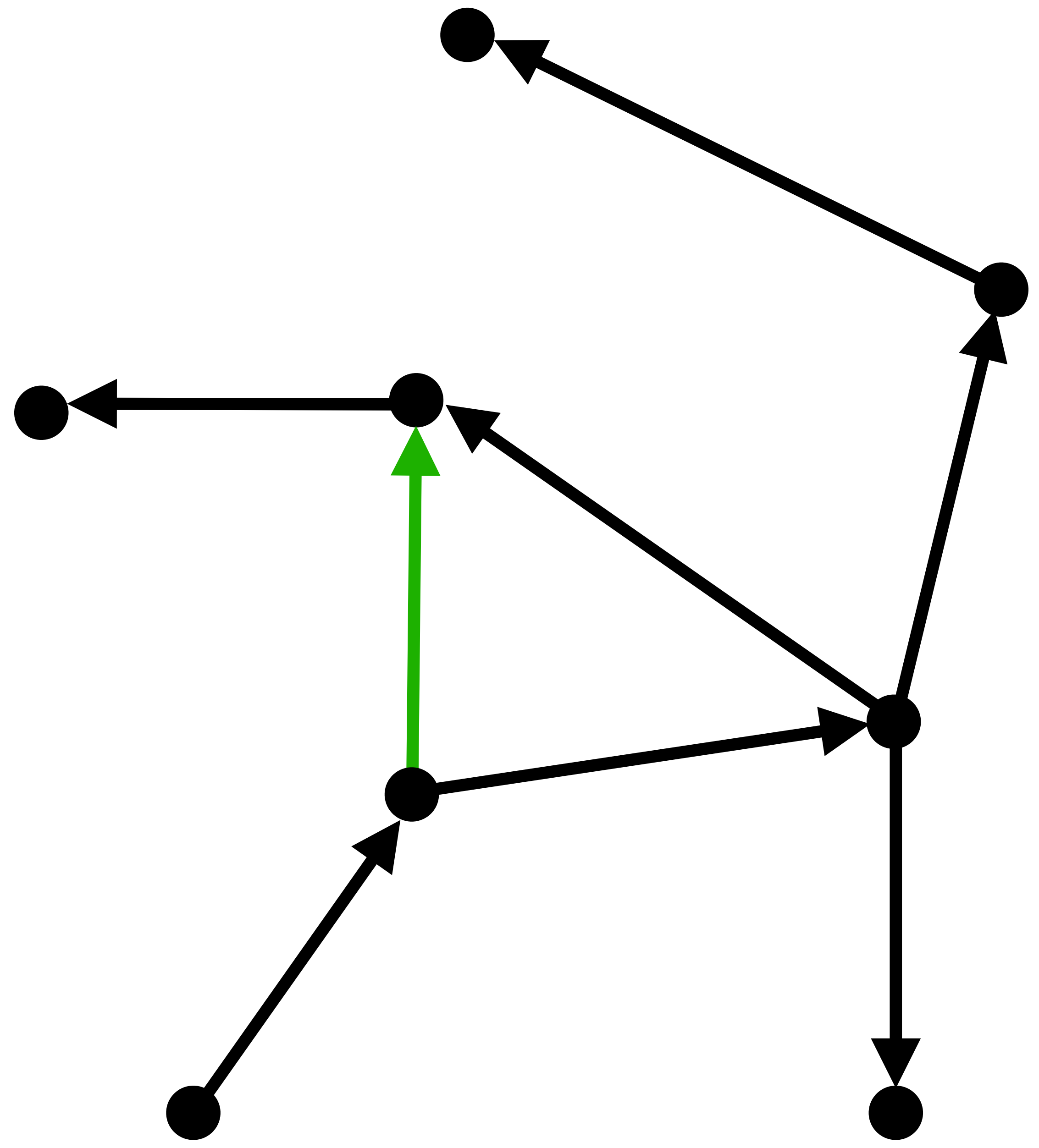
Triadic closure offers small choice sets
→ tractable inference
→ varied choice sets



Our data
Timestamped edges
(including repeats)

Choosing to close triangles

Triadic closure offers small choice sets
→ tractable inference
→ varied choice sets

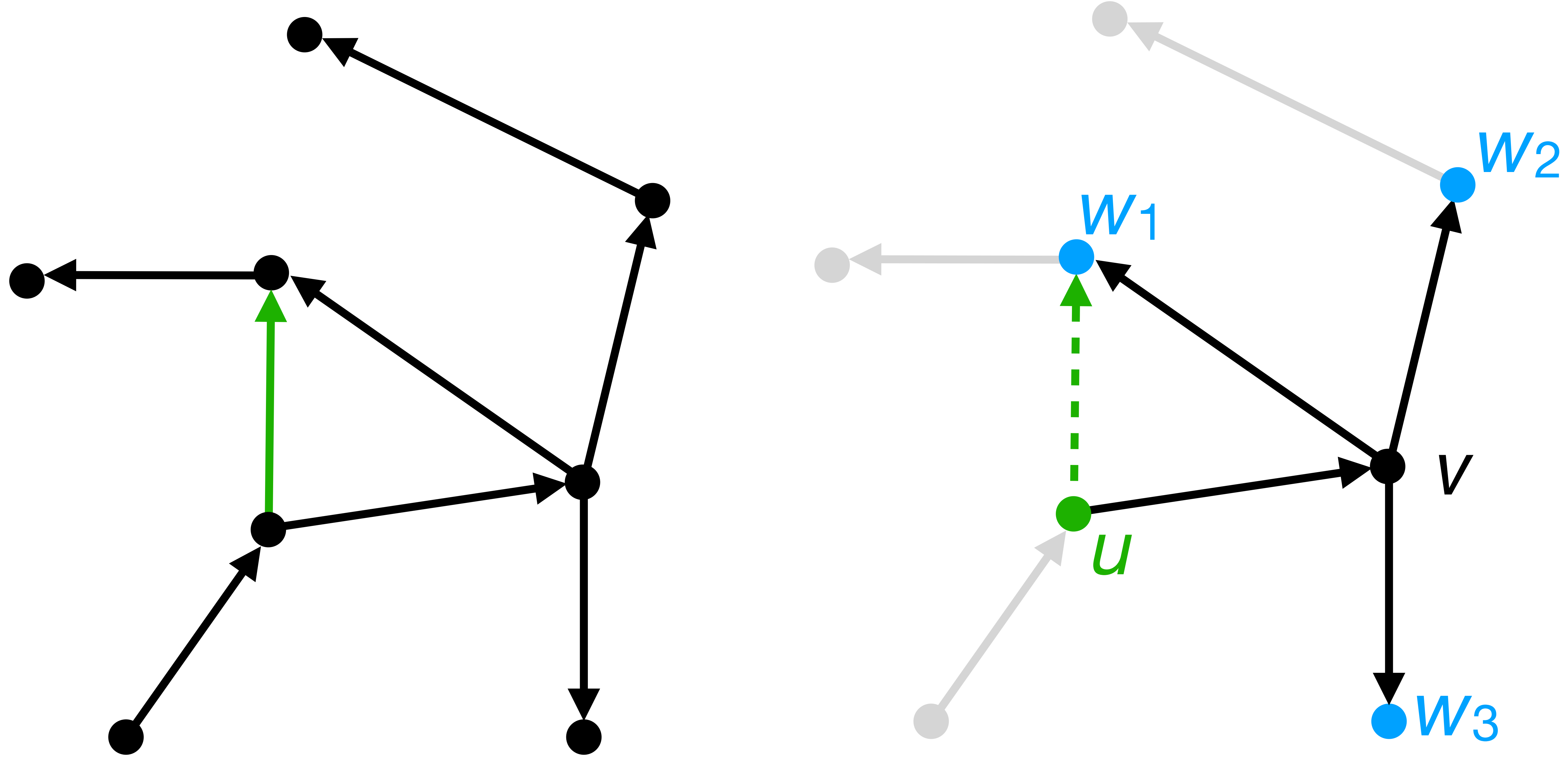


Our data
Timestamped edges
(including repeats)

Choosing to close triangles

Triadic closure offers small choice sets
→ tractable inference
→ varied choice sets

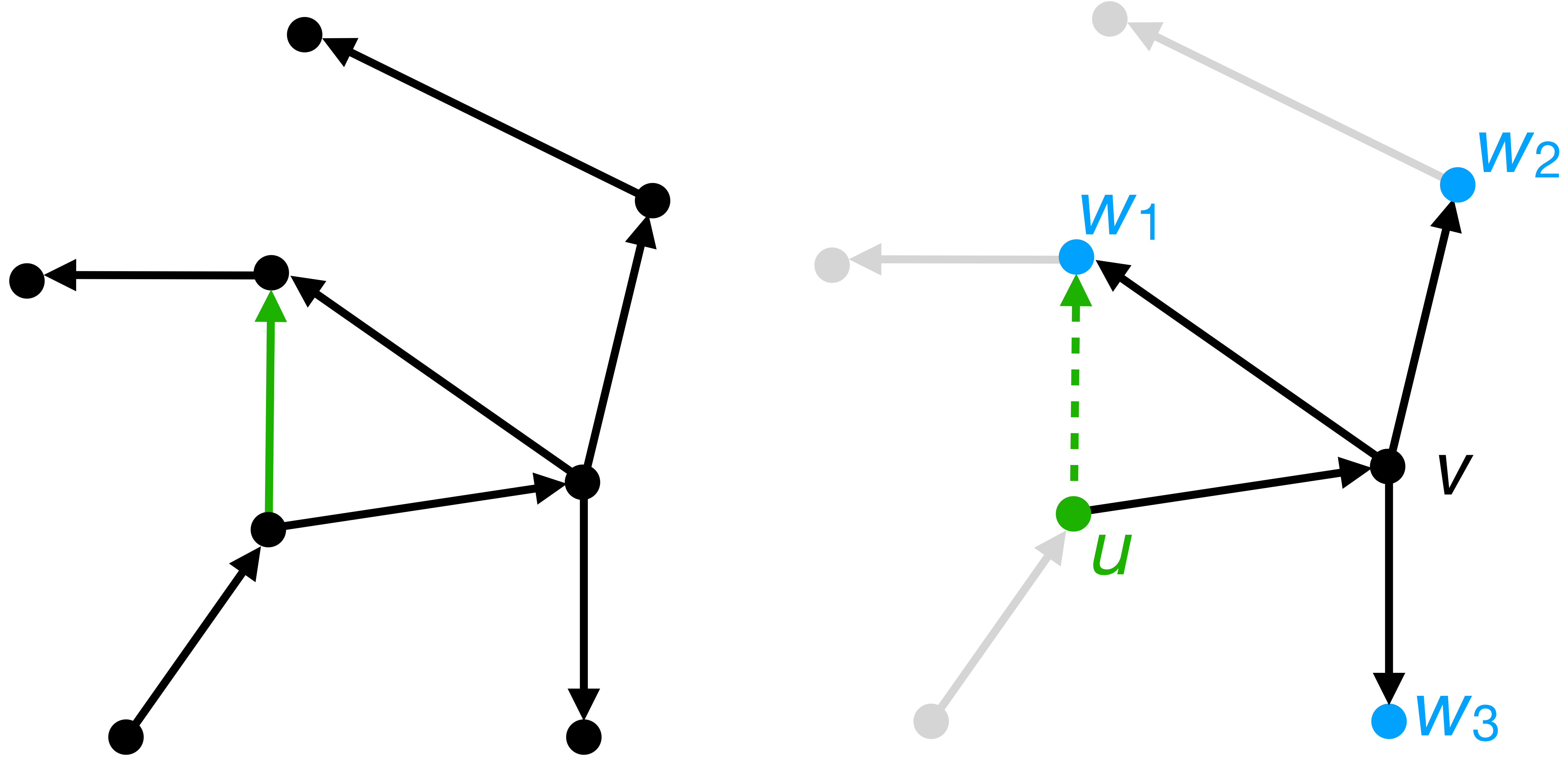
Our data
Timestamped edges
(including repeats)



Choosing to close triangles

Triadic closure offers small choice sets
→ tractable inference
→ varied choice sets

Our data
Timestamped edges
(including repeats)

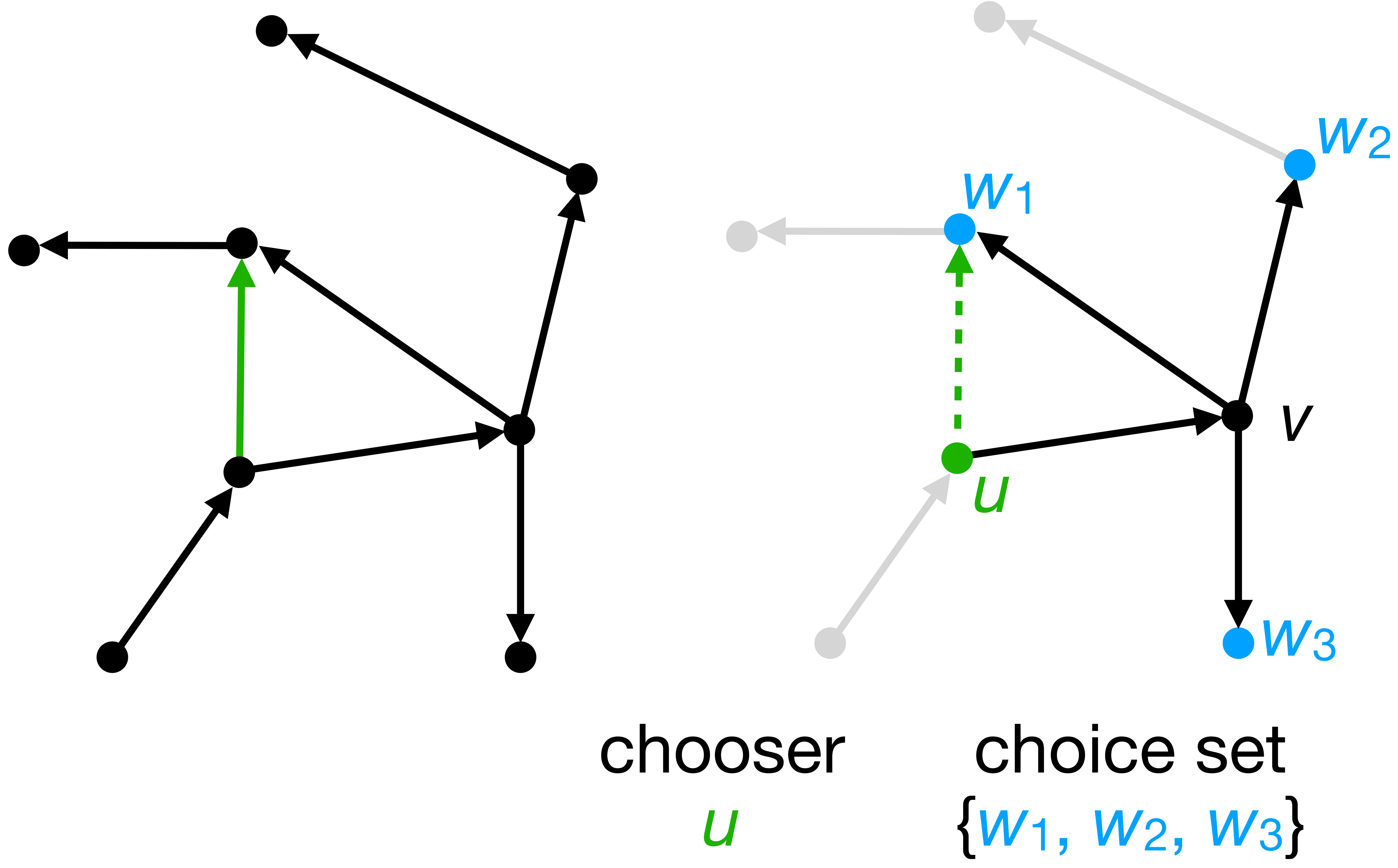


chooser
u

Choosing to close triangles

Triadic closure offers small choice sets
→ tractable inference
→ varied choice sets

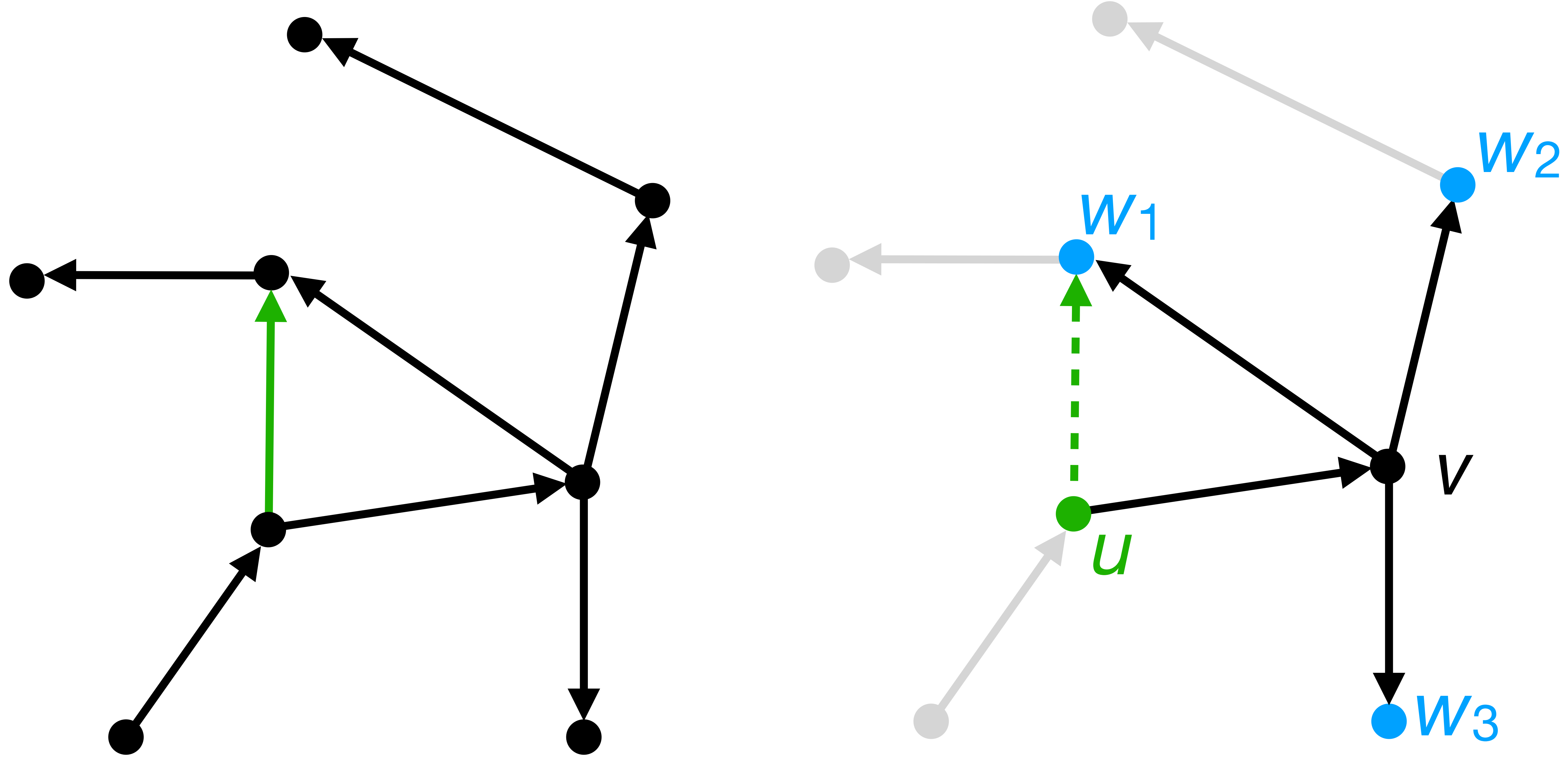
Our data
Timestamped edges
(including repeats)



Choosing to close triangles

Triadic closure offers small choice sets
→ tractable inference
→ varied choice sets

Our data
Timestamped edges
(including repeats)

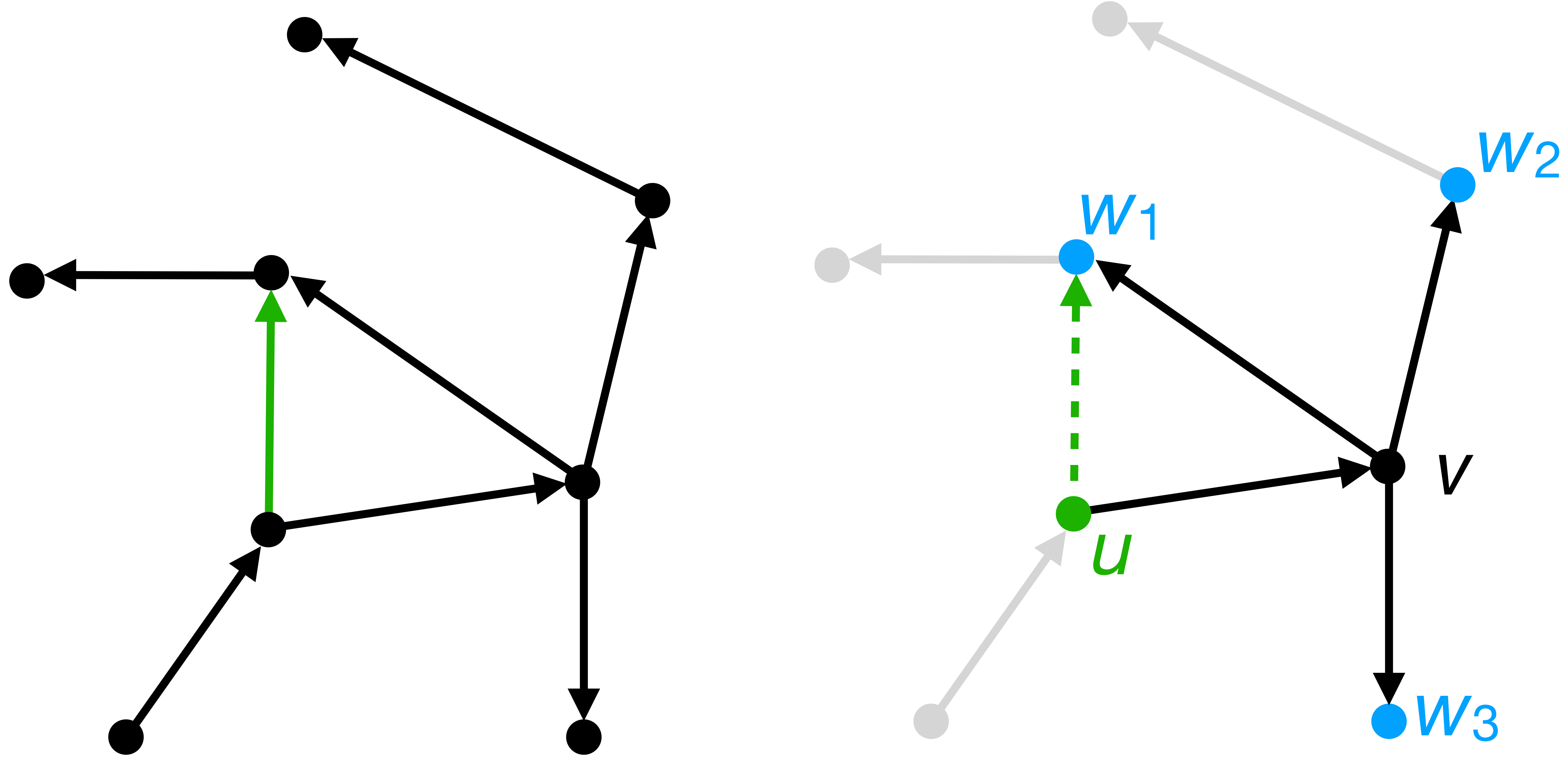


chooser	choice set	choice
u	$\{w_1, w_2, w_3\}$	w_1

Choosing to close triangles

Triadic closure offers small choice sets
→ tractable inference
→ varied choice sets

Our data
Timestamped edges
(including repeats)



Node features

1. in-degree of w
2. # shared neighbors of u, w
3. weight of edge $w \to u$
4. time since last edge into w
5. time since last edge out of w
6. time since last $w \to u$ edge

chooser	choice set	choice
u	$\{w_1, w_2, w_3\}$	w_1

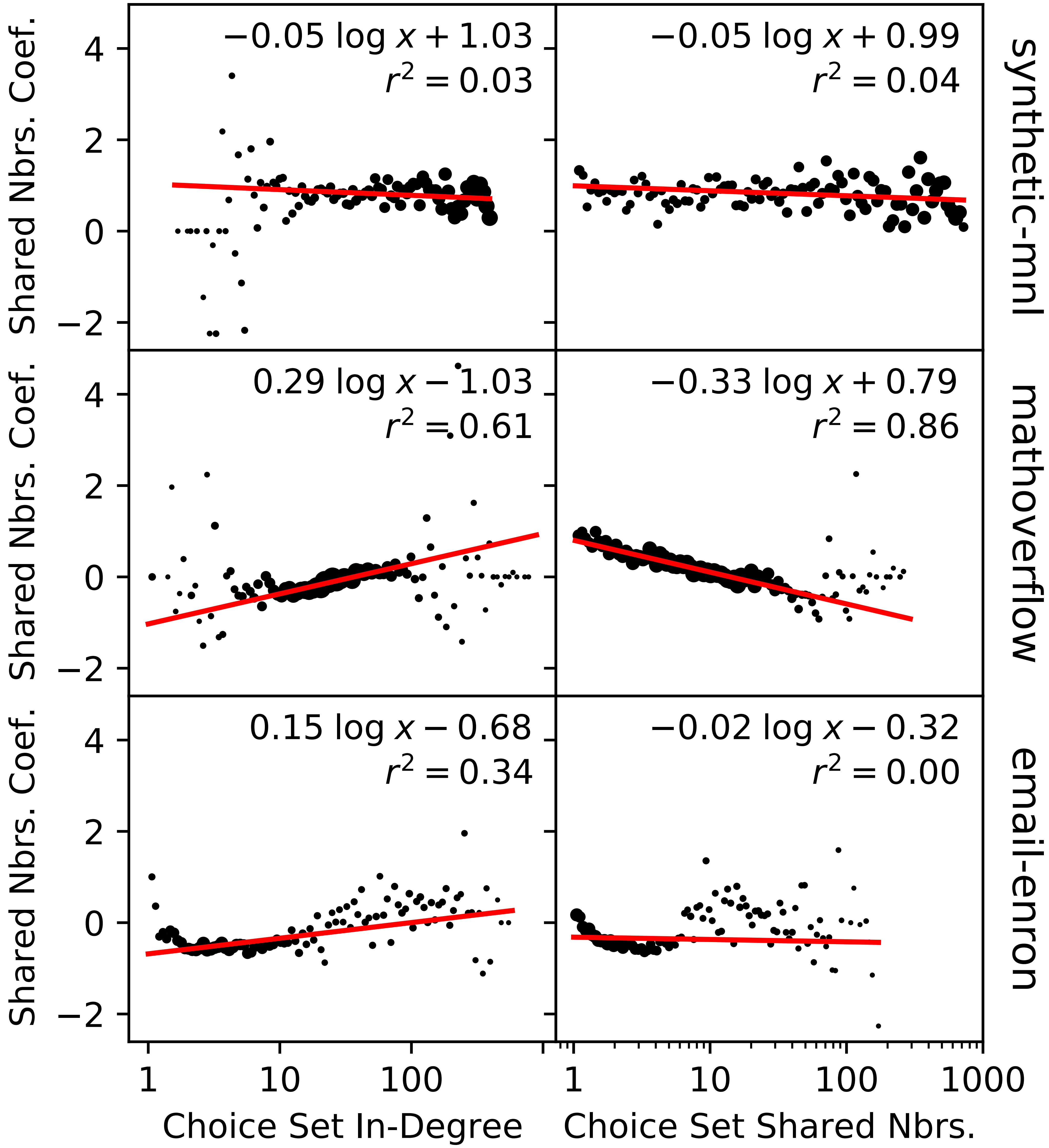
Context matters in triadic closure

Context matters in triadic closure

Datasets
email-enron
email-eu
email-w3c
wiki-talk
reddit-hyperlink
bitcoin-alpha
bitcoin-otc
mathoverflow
college-msg
facebook-wall
sms-a
sms-b
sms-c

bit.ly/lcl-data

Context matters in triadic closure

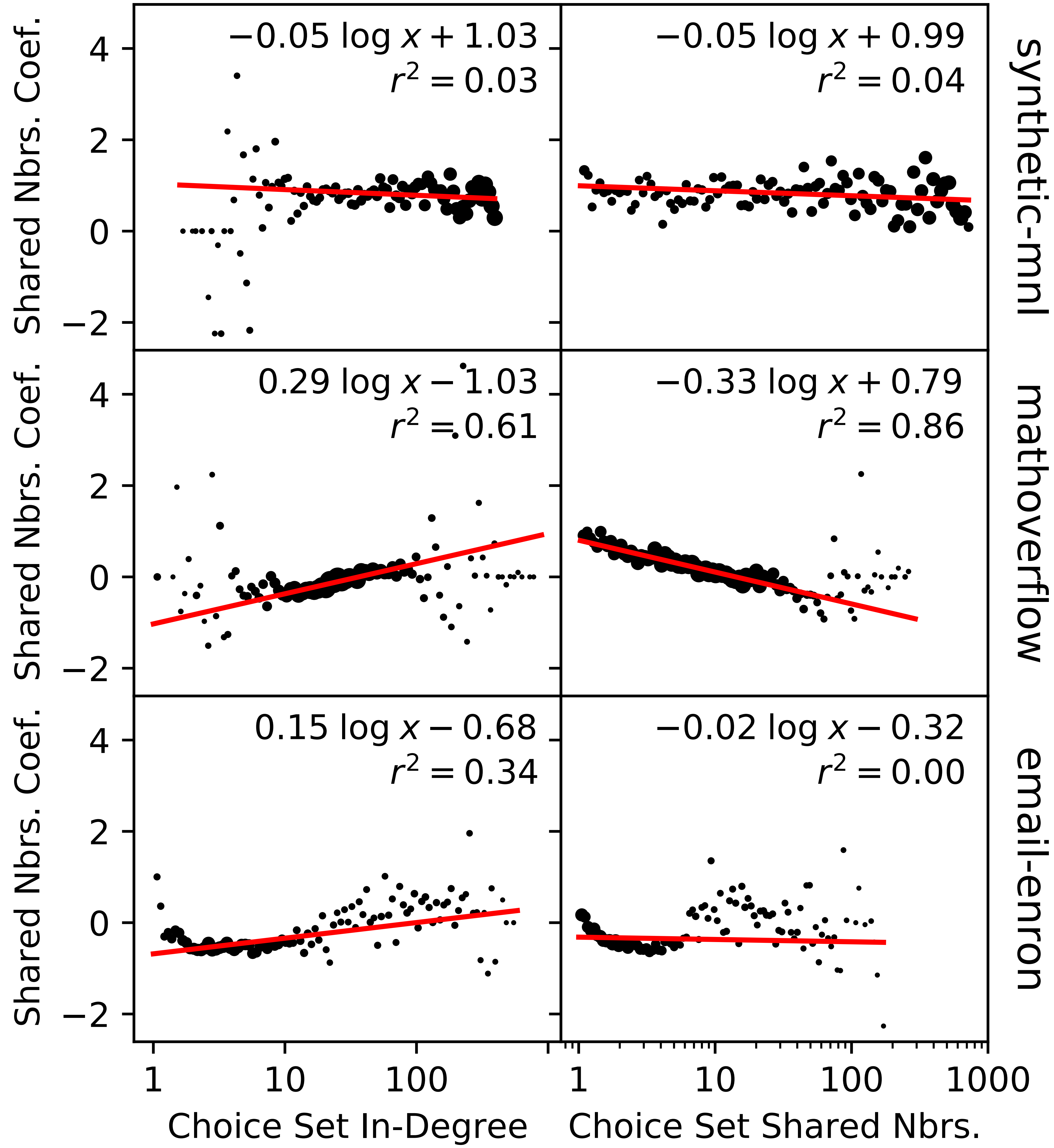


Datasets

email-enron
email-eu
email-w3c
wiki-talk
reddit-hyperlink
bitcoin-alpha
bitcoin-otc
mathoverflow
college-msg
facebook-wall
sms-a
sms-b
sms-c

bit.ly/lcl-data

Context matters in triadic closure

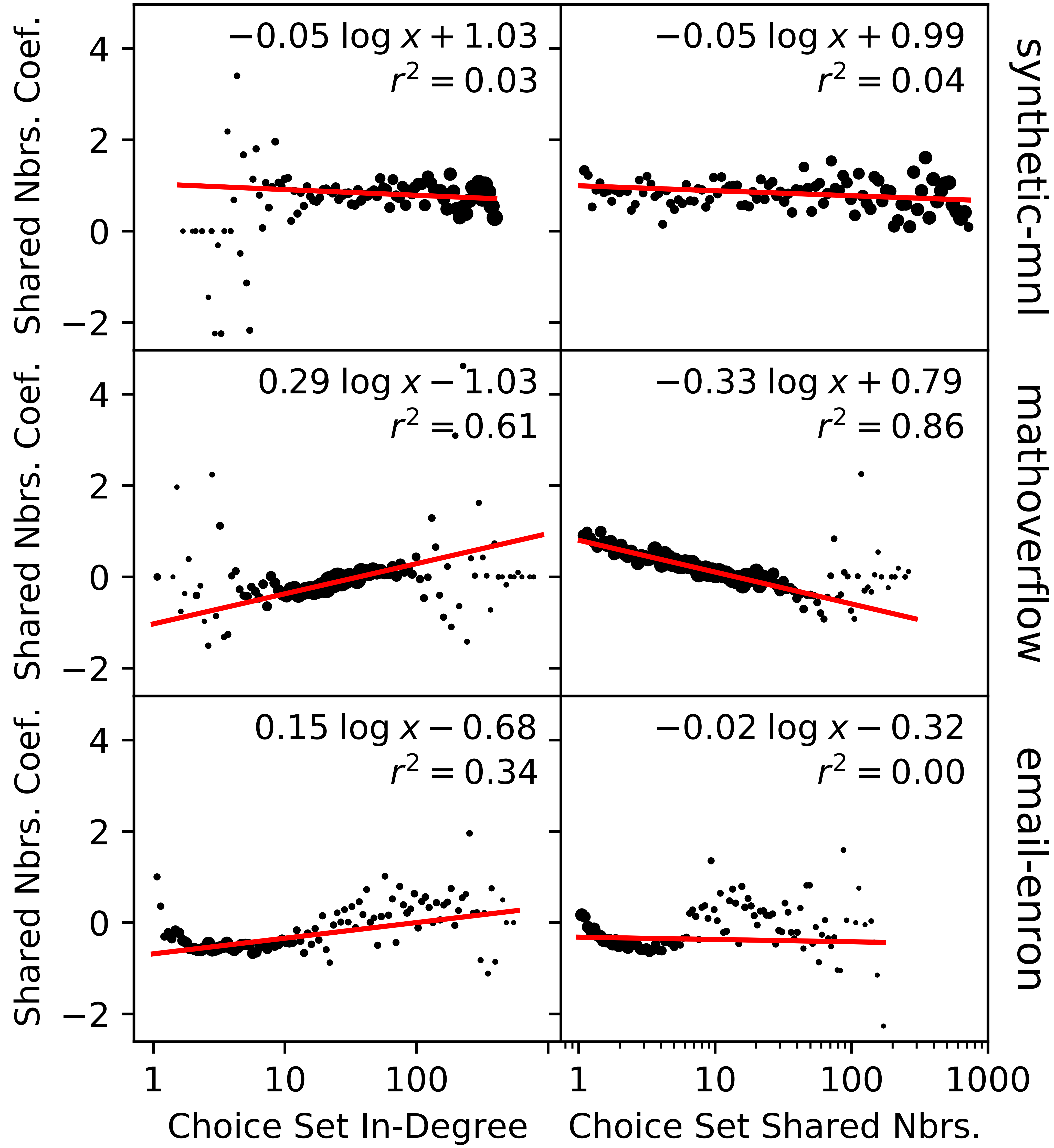


Synthetic data,
no context effects

Datasets
email-enron
email-eu
email-w3c
wiki-talk
reddit-hyperlink
bitcoin-alpha
bitcoin-otc
mathoverflow
college-msg
facebook-wall
sms-a
sms-b
sms-c

bit.ly/lcl-data

Context matters in triadic closure



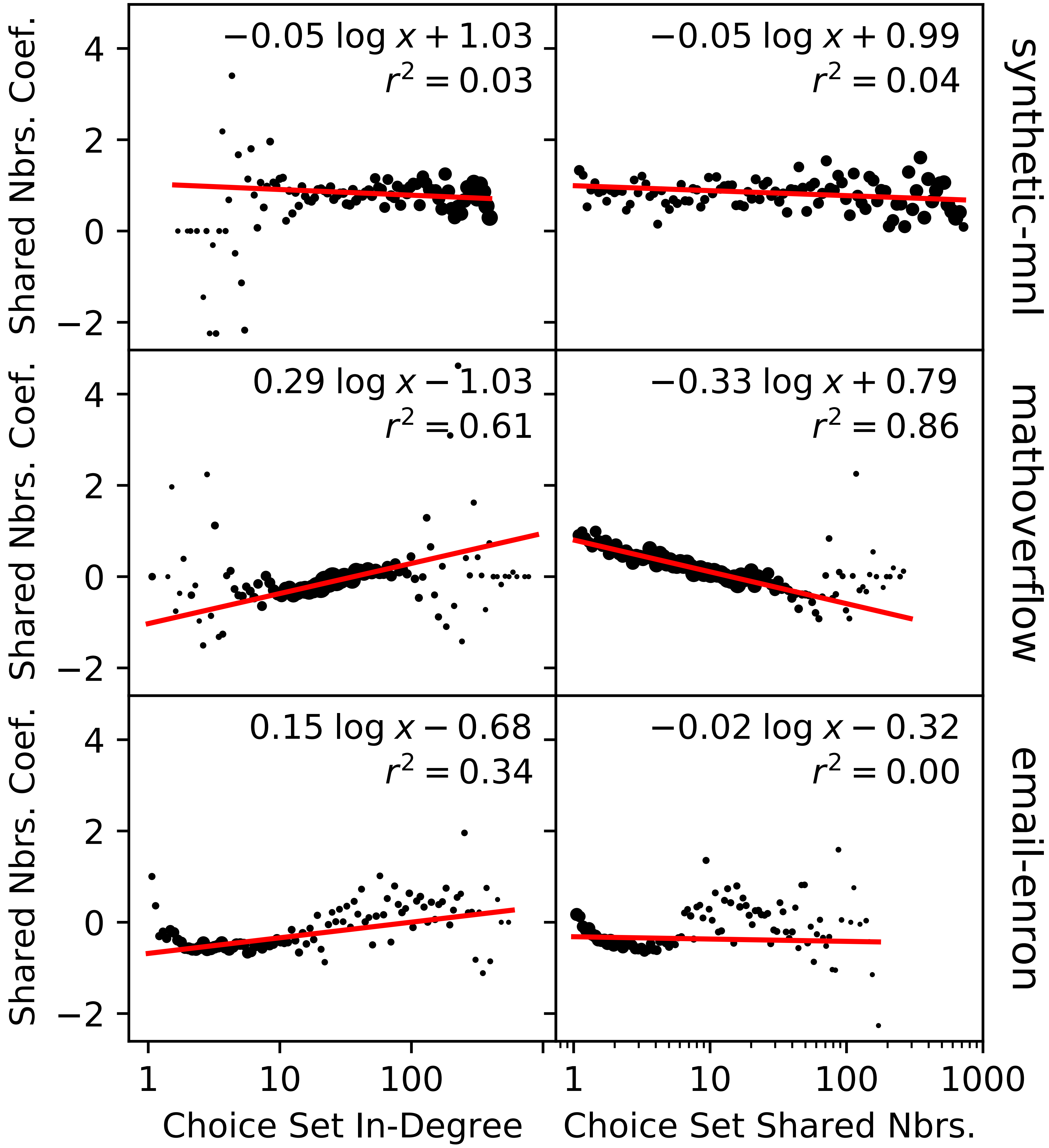
Synthetic data,
no context effects

Commenting network,
linear context effects

Datasets
email-enron
email-eu
email-w3c
wiki-talk
reddit-hyperlink
bitcoin-alpha
bitcoin-otc
mathoverflow
college-msg
facebook-wall
sms-a
sms-b
sms-c

bit.ly/lcl-data

Context matters in triadic closure



Synthetic data,
no context effects

Commenting network,
linear context effects

Email network,
nonlinear context effects?

Datasets
email-enron
email-eu
email-w3c
wiki-talk
reddit-hyperlink
bitcoin-alpha
bitcoin-otc
mathoverflow
college-msg
facebook-wall
sms-a
sms-b
sms-c

bit.ly/lcl-data

LCL reveals interpretable context effects

LCL reveals interpretable context effects

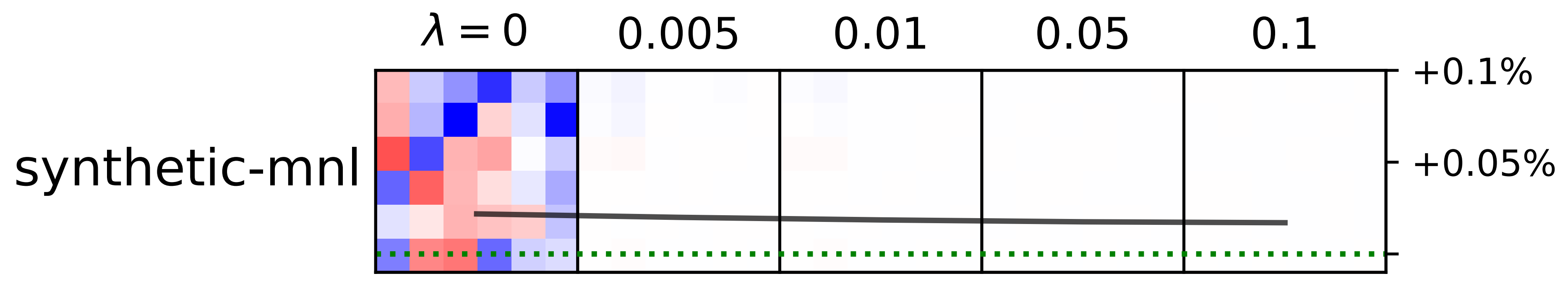
Estimation

MLE to infer LCL

$$\ell(\theta, A; \mathscr{D}) = \sum_{(i,C)\in\mathscr{D}} \left(\theta + Ax_C\right)^T x_i - \log \sum_{j\in C} \exp(\left[\theta + Ax_C\right]^T x_j)$$

(concave)

LCL reveals interpretable context effects



Estimation

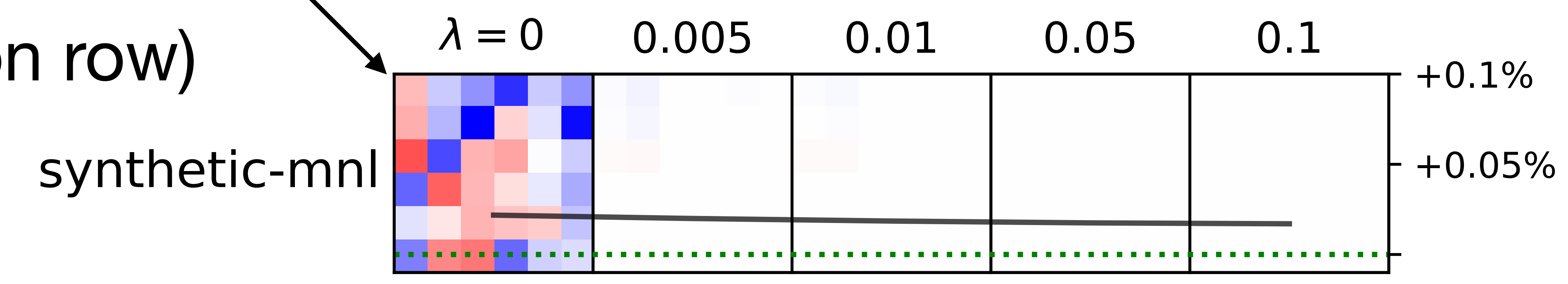
MLE to infer LCL

$$\ell(\theta, A; \mathcal{D}) = \sum_{(i,C)\in\mathcal{D}} \left(\theta + Ax_C\right)^T x_i - \log \sum_{j\in C} \exp(\left[\theta + Ax_C\right]^T x_j)$$

(concave)

LCL reveals interpretable context effects

context effect matrix A
red: +, blue: -, white: 0
(column acts on row)



Node features
(left-right, top-bottom)

1. in-degree
2. shared neighbors
3. reciprocal weight
4. send recency
5. receive recency
6. reciprocal recency

Estimation

MLE to infer LCL

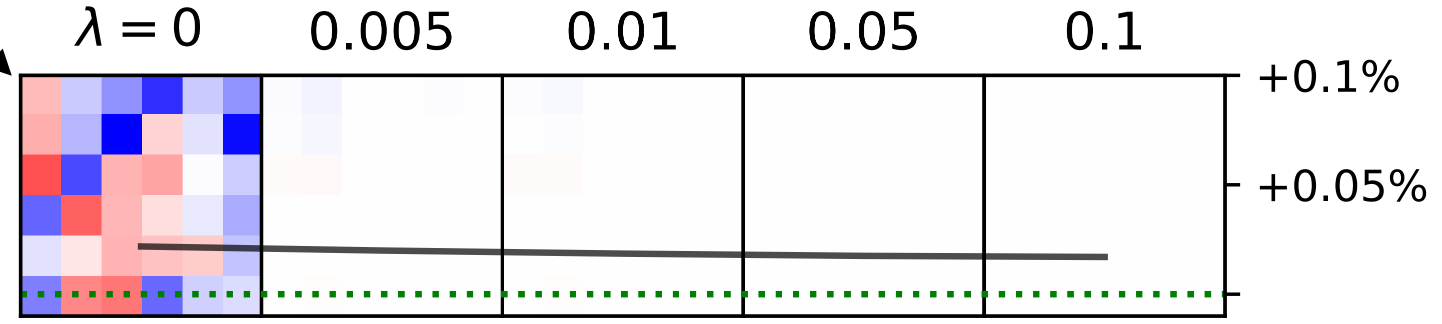
$$\ell(\theta, A; \mathscr{D}) = \sum_{(i,C)\in\mathscr{D}} (\theta + A x_C)^T x_i - \log \sum_{j\in C} \exp([\theta + A x_C]^T x_j)$$

(concave)

LCL reveals interpretable context effects

context effect matrix A
red: +, blue: -, white: 0
(column acts on row)

L_1 regularization level



Node features
(left-right, top-bottom)

1. in-degree
2. shared neighbors
3. reciprocal weight
4. send recency
5. receive recency
6. reciprocal recency

Estimation

MLE to infer LCL

$$\ell(\theta, A; \mathscr{D}) = \sum_{(i,C)\in\mathscr{D}} \left(\theta + Ax_C\right)^T x_i - \log \sum_{j\in C} \exp(\left[\theta + Ax_C\right]^T x_j)$$

(concave)

LCL reveals interpretable context effects

Node features
(left-right, top-bottom)

1. in-degree
2. shared neighbors
3. reciprocal weight
4. send recency
5. receive recency
6. reciprocal recency

Estimation

MLE to infer LCL

$$\ell(\theta, A; \mathcal{D}) = \sum_{(i,C)\in\mathcal{D}} (\theta + Ax_C)^T x_i - \log \sum_{j\in C} \exp([\theta + Ax_C]^T x_j)$$

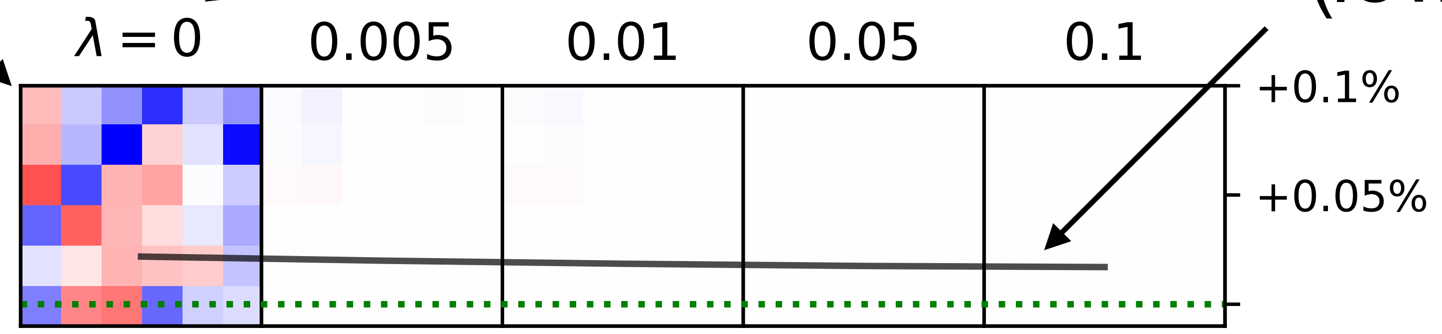
(concave)

context effect matrix A
red: +, blue: -, white: 0
(column acts on row)

L_1 regularization level

LCL negative log-likelihood
(lower = better)

synthetic-mnl



LCL reveals interpretable context effects

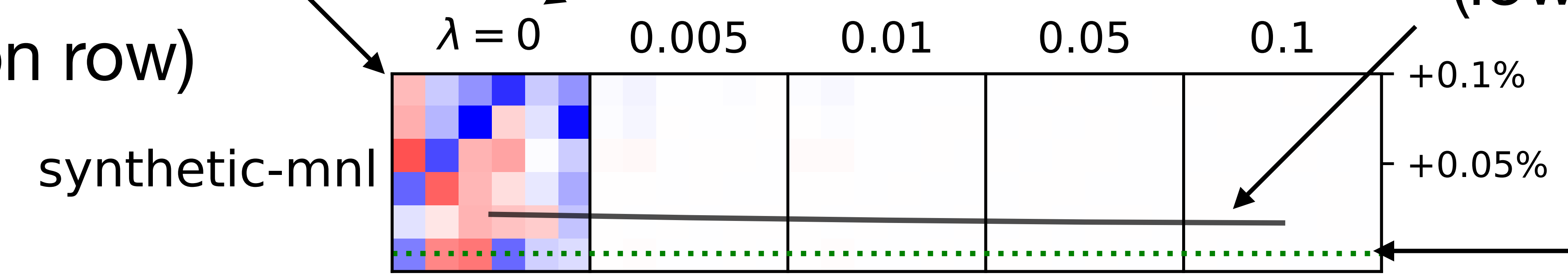
Node features
(left-right, top-bottom)

1. in-degree
2. shared neighbors
3. reciprocal weight
4. send recency
5. receive recency
6. reciprocal recency

context effect matrix A
red: +, blue: -, white: 0
(column acts on row)

L_1 regularization level

LCL negative log-likelihood
(lower = better)



likelihood-ratio test vs MNL
significance threshold
($p < 0.001$)

Estimation

MLE to infer LCL

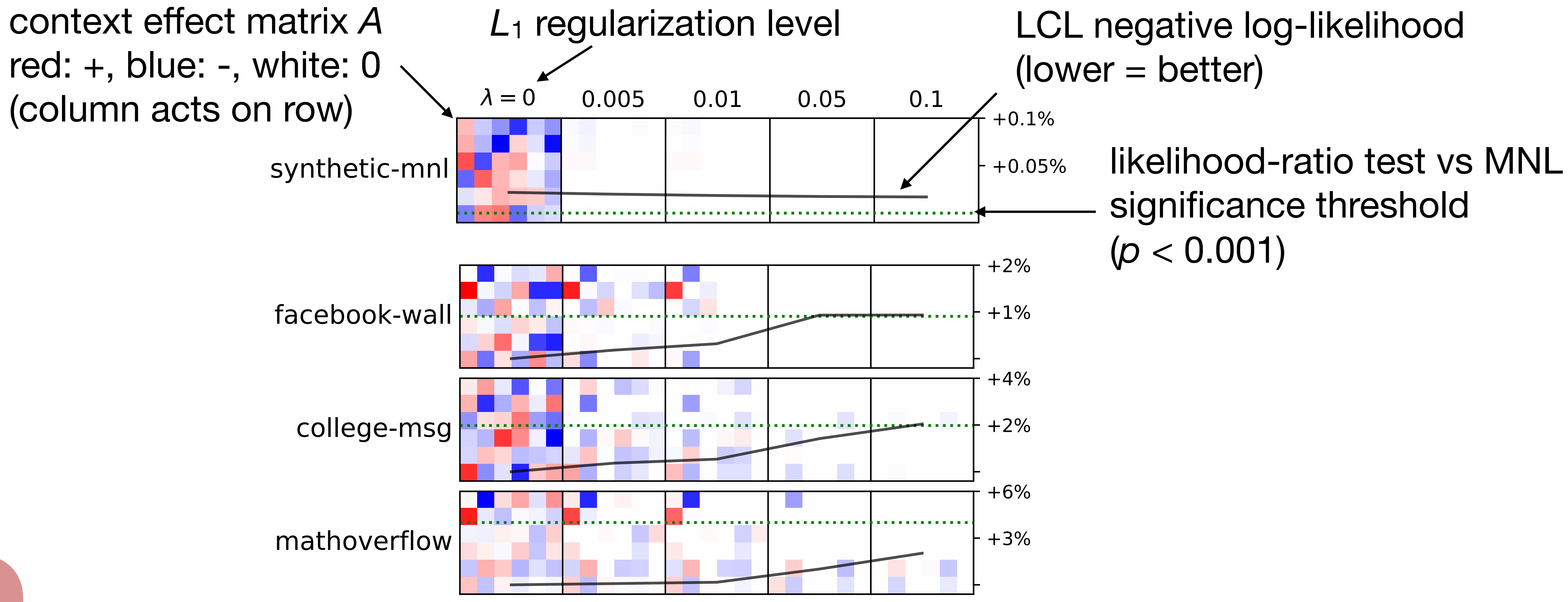
$$\ell(\theta, A; \mathscr{D}) = \sum_{(i,C)\in\mathscr{D}} \left(\theta + A x_C\right)^T x_i - \log \sum_{j\in C} \exp(\left[\theta + A x_C\right]^T x_j)$$

(concave)

LCL reveals interpretable context effects

Node features
(left-right, top-bottom)

1. in-degree
2. shared neighbors
3. reciprocal weight
4. send recency
5. receive recency
6. reciprocal recency



Estimation

MLE to infer LCL

$$\ell(\theta, A; \mathscr{D}) = \sum_{(i,C)\in\mathscr{D}} (\theta + A x_C)^T x_i - \log \sum_{j\in C} \exp([\theta + A x_C]^T x_j)$$

(concave)

LCL reveals interpretable context effects

Node features
(left-right, top-bottom)

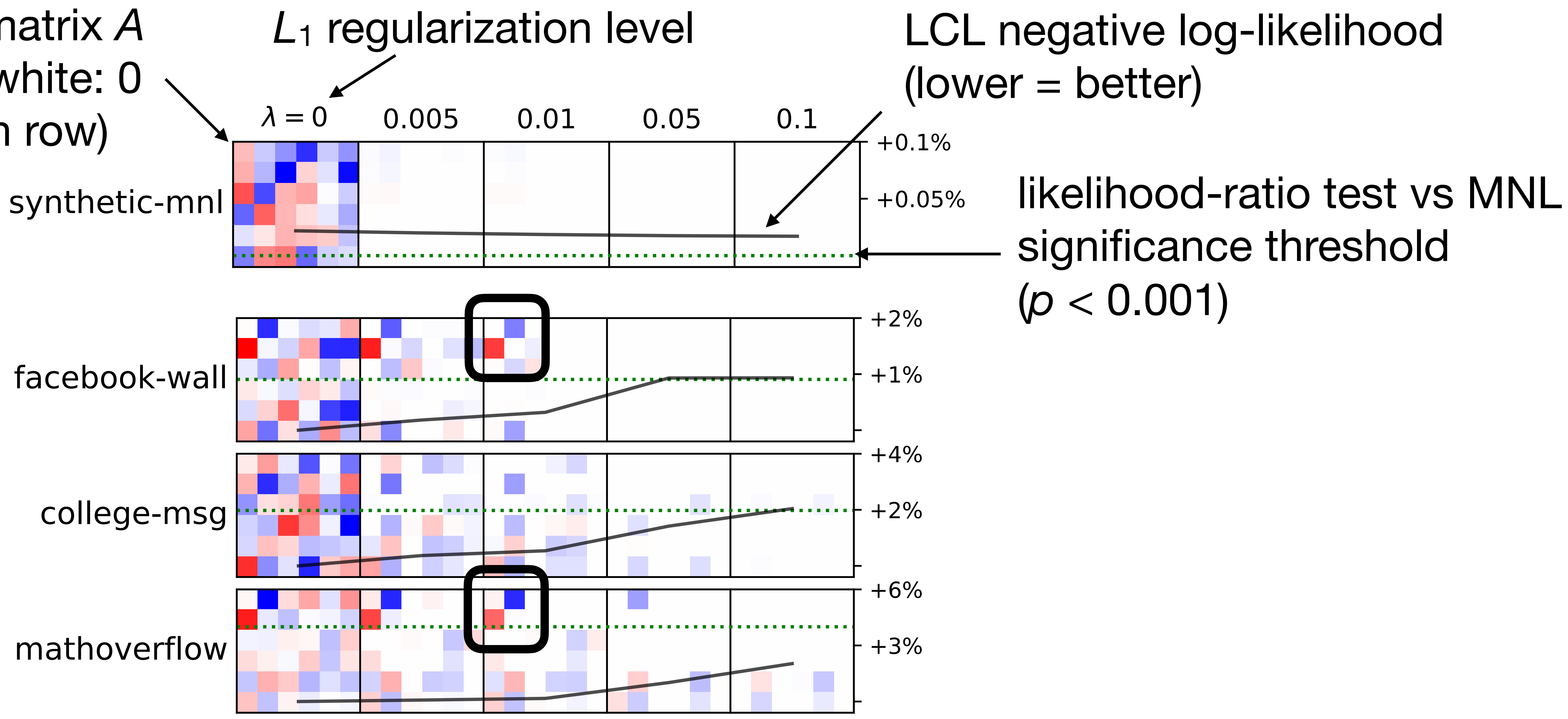
1. in-degree
2. shared neighbors
3. reciprocal weight
4. send recency
5. receive recency
6. reciprocal recency

Estimation

MLE to infer LCL

$$\ell(\theta, A; \mathscr{D}) = \sum_{(i,C)\in\mathscr{D}} (\theta + Ax_C)^T x_i - \log \sum_{j\in C} \exp([\theta + Ax_C]^T x_j)$$

(concave)



LCL reveals interpretable context effects

Node features
(left-right, top-bottom)

1. in-degree
2. shared neighbors
3. reciprocal weight
4. send recency
5. receive recency
6. reciprocal recency

Estimation

MLE to infer LCL

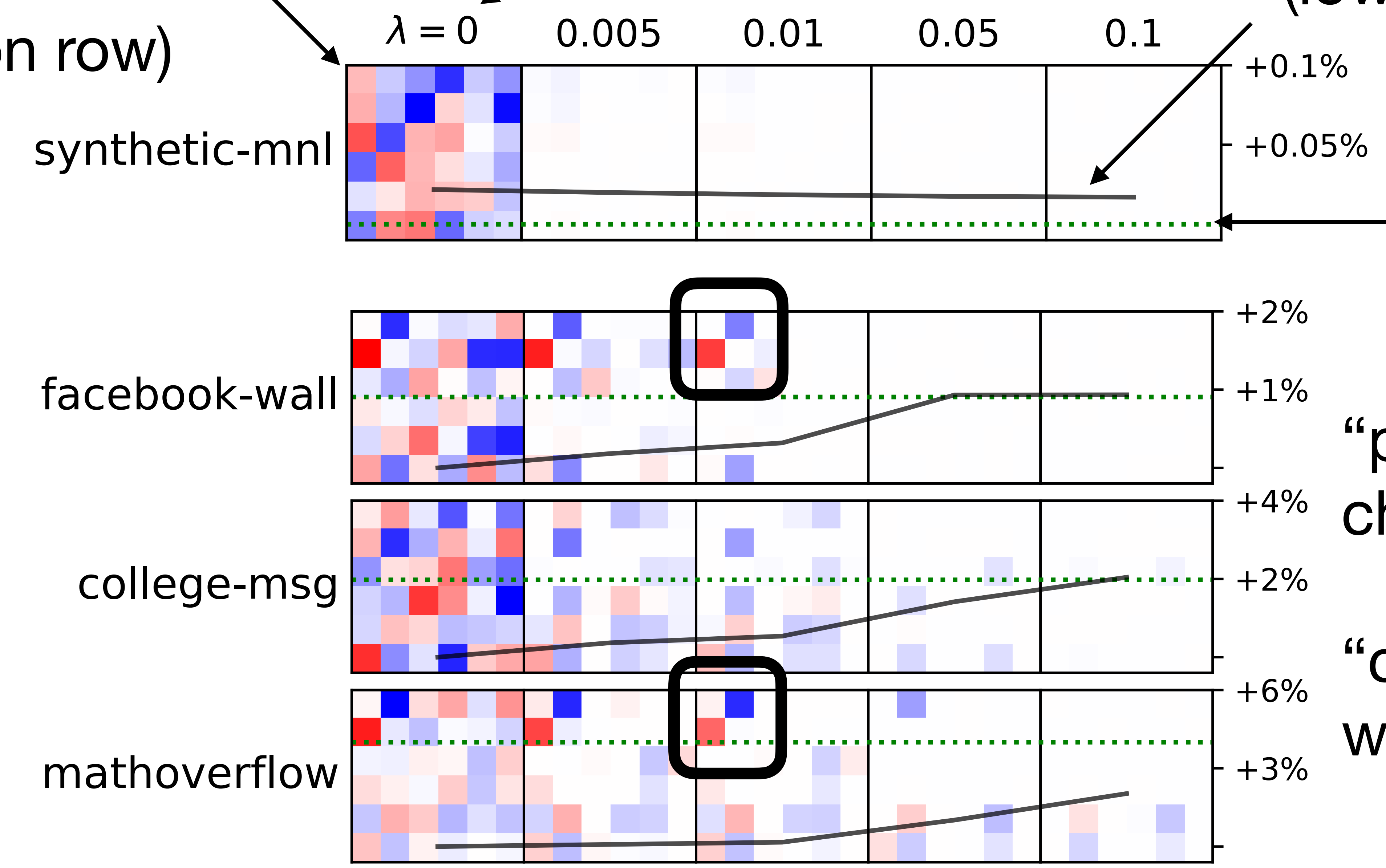
$$\ell(\theta, A; \mathscr{D}) = \sum_{(i,C)\in\mathscr{D}} (\theta + A x_C)^T x_i - \log \sum_{j \in C} \exp([\theta + A x_C]^T x_j)$$

(concave)

context effect matrix A
red: +, blue: -, white: 0
(column acts on row)

L_1 regularization level

LCL negative log-likelihood
(lower = better)



likelihood-ratio test vs MNL
significance threshold
($p < 0.001$)

“popularity matters less when choosing from close connections”

“close connections matter more when choosing from the popular”

LCL reveals interpretable context effects

Node features
(left-right, top-bottom)

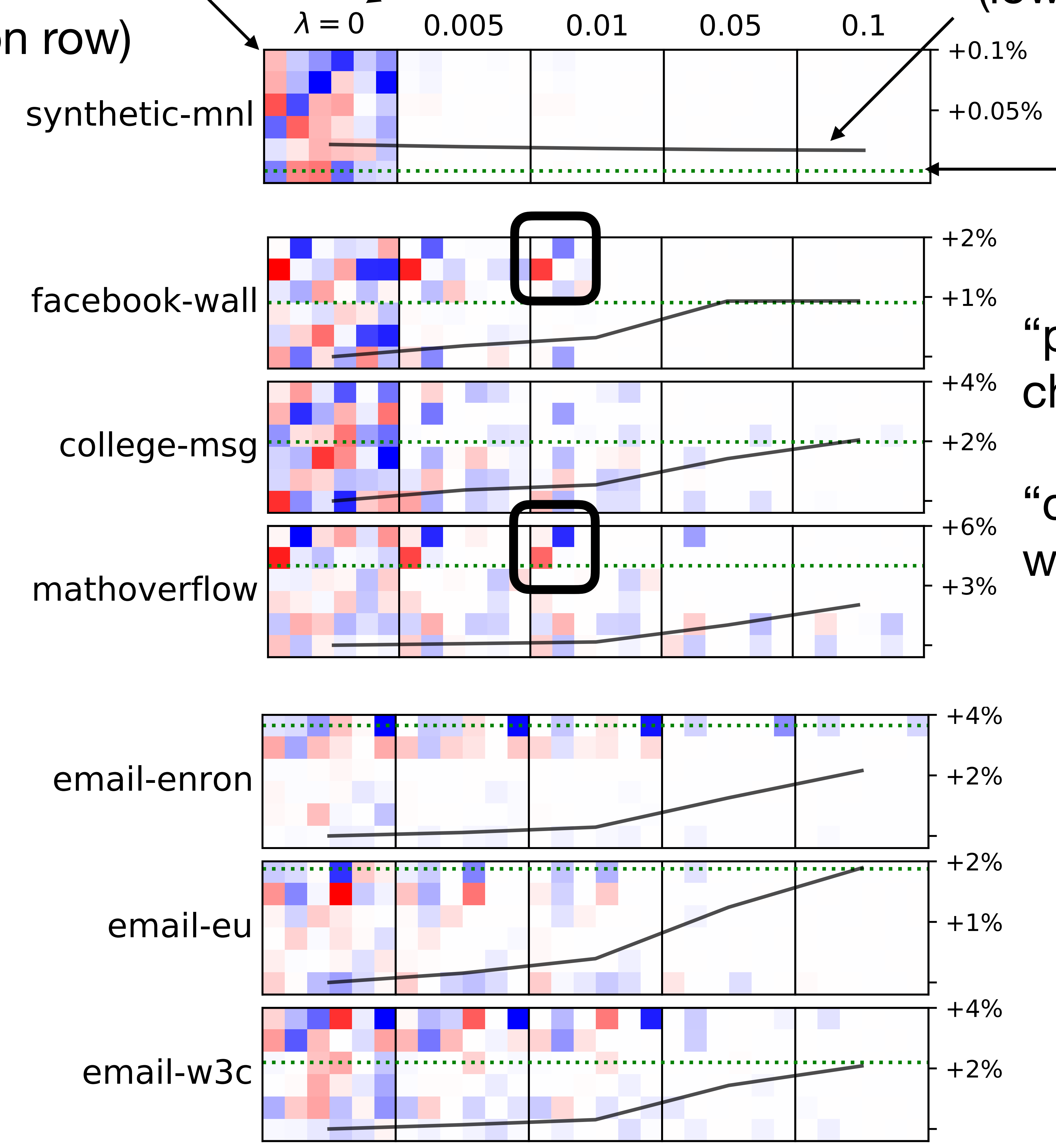
1. in-degree
2. shared neighbors
3. reciprocal weight
4. send recency
5. receive recency
6. reciprocal recency

Estimation

MLE to infer LCL

$$\ell(\theta, A; \mathscr{D}) = \sum_{(i,C)\in\mathscr{D}} (\theta + A x_C)^T x_i - \log \sum_{j\in C} \exp([\theta + A x_C]^T x_j)$$

(concave)



LCL reveals interpretable context effects

Node features
(left-right, top-bottom)

1. in-degree
2. shared neighbors
3. reciprocal weight
4. send recency
5. receive recency
6. reciprocal recency

Estimation

MLE to infer LCL

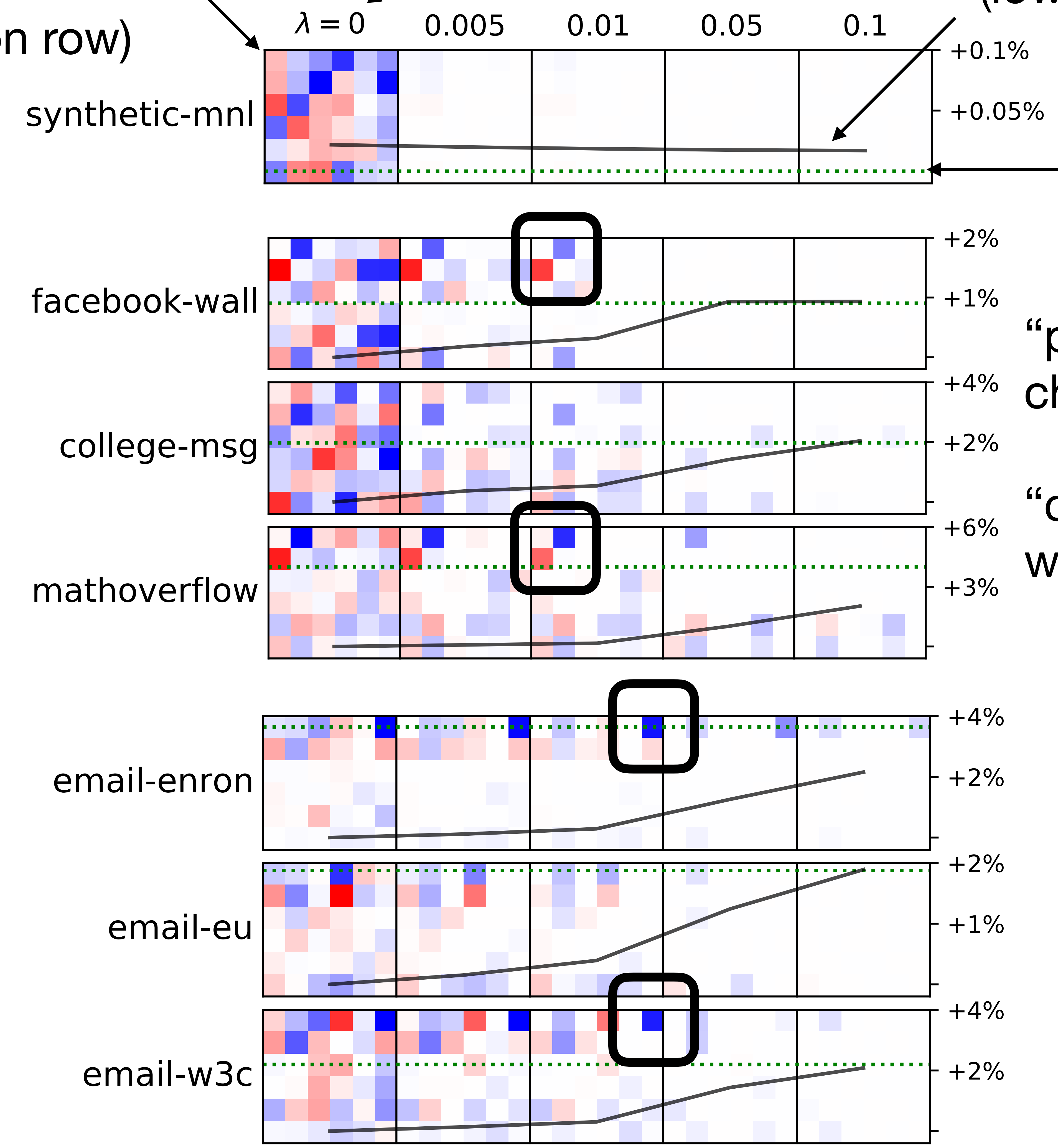
$$\ell(\theta, A; \mathcal{D}) = \sum_{(i,C)\in\mathcal{D}} (\theta + A x_C)^T x_i - \log \sum_{j \in C} \exp([\theta + A x_C]^T x_j)$$

(concave)

context effect matrix A
red: +, blue: -, white: 0
(column acts on row)

L_1 regularization level

LCL negative log-likelihood
(lower = better)



likelihood-ratio test vs MNL
significance threshold
($p < 0.001$)

“popularity matters less when choosing from close connections”

“close connections matter more when choosing from the popular”

LCL reveals interpretable context effects

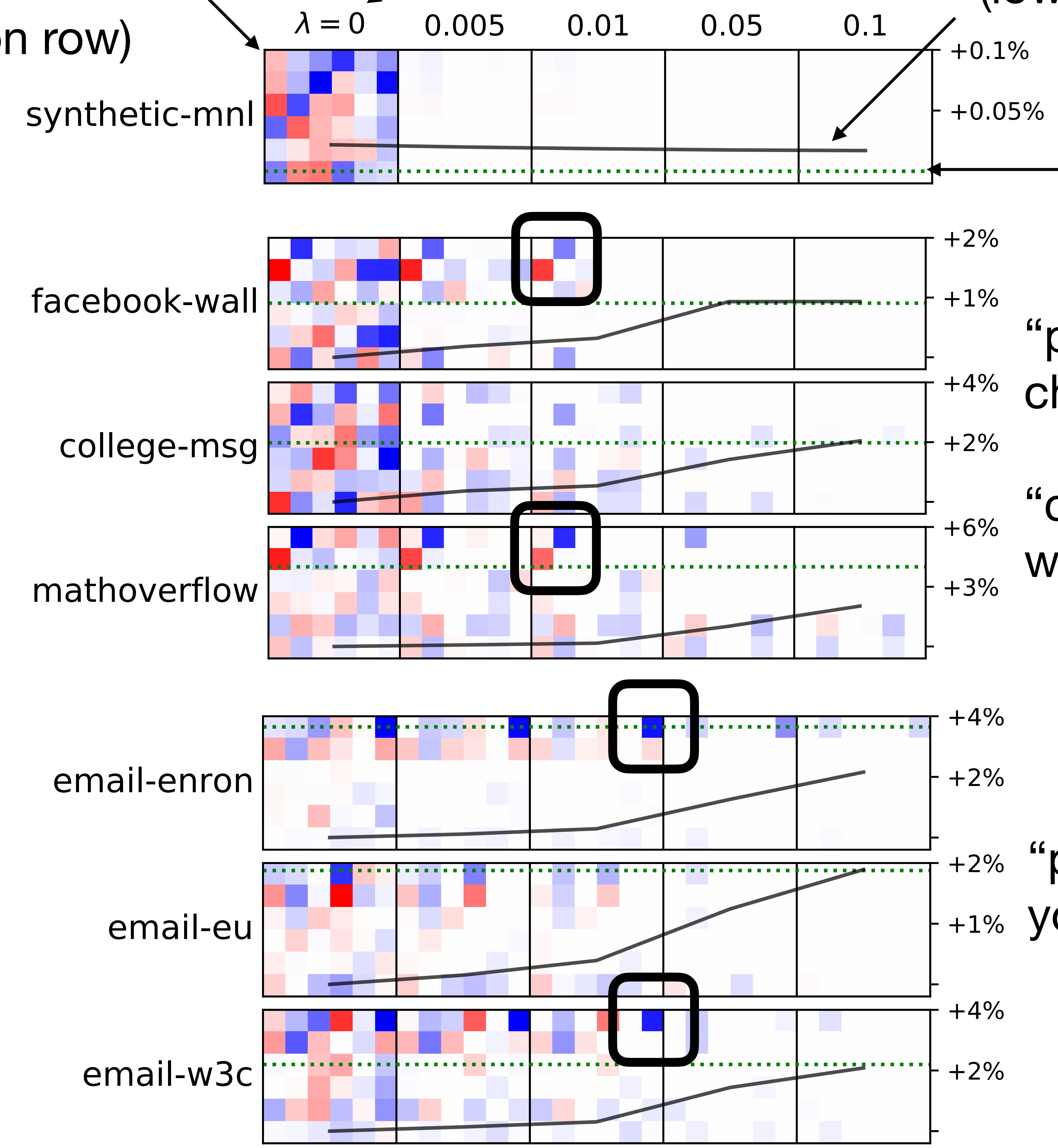
Node features
(left-right, top-bottom)

1. in-degree
2. shared neighbors
3. reciprocal weight
4. send recency
5. receive recency
6. reciprocal recency

context effect matrix A
red: +, blue: -, white: 0
(column acts on row)

L_1 regularization level

LCL negative log-likelihood
(lower = better)



likelihood-ratio test vs MNL significance threshold ($p < 0.001$)

“popularity matters less when choosing from close connections”

“close connections matter more when choosing from the popular”

“popularity matters less when your inbox is full of recent emails”

Estimation

MLE to infer LCL

$$\ell(\theta, A; \mathscr{D}) = \sum_{(i,C)\in\mathscr{D}} (\theta + A x_C)^T x_i - \log \sum_{j\in C} \exp([\theta + A x_C]^T x_j)$$

(concave)

Other things in our paper

Kiran Tomlinson and Austin R. Benson
Learning Interpretable Feature Context Effects in Discrete Choice
arXiv: 2009.03417, September 2020
bit.ly/lcl-paper

Other things in our paper

Kiran Tomlinson and Austin R. Benson
Learning Interpretable Feature Context Effects in Discrete Choice
***arXiv: 2009.03417*, September 2020**
bit.ly/lcl-paper

- LCL derivation from simple assumptions

Other things in our paper

$$\Pr(i, C) = \sum_{k=1}^{d} \pi_k \frac{\exp\left([B_k + A_k (x_C)_k]^T x_i\right)}{\sum_{j \in C} \exp\left([B_k + A_k (x_C)_k]^T x_j\right)}$$

Kiran Tomlinson and Austin R. Benson
Learning Interpretable Feature Context Effects in Discrete Choice
***arXiv: 2009.03417*, September 2020**
bit.ly/lcl-paper

- LCL derivation from simple assumptions
- More flexible model: decomposed LCL

Other things in our paper

$$\Pr(i, C) = \sum_{k=1}^{d} \pi_k \frac{\exp\left([B_k + A_k(x_C)_k]^T x_i\right)}{\sum_{j \in C} \exp\left([B_k + A_k(x_C)_k]^T x_j\right)}$$

Kiran Tomlinson and Austin R. Benson
Learning Interpretable Feature Context Effects in Discrete Choice
***arXiv: 2009.03417*, September 2020**
bit.ly/lcl-paper

Theorem 1. A d-feature linear context logit is identifiable from a dataset $\mathcal{D}$ if and only if

$$\operatorname{span}\left\{\begin{bmatrix} x_C \\ 1 \end{bmatrix} \otimes (x_i - x_C) \mid C \in \mathcal{C}_\mathcal{D}, i \in C\right\} = \mathbb{R}^{d^2+d}. \qquad (6)$$

- LCL derivation from simple assumptions
- More flexible model: decomposed LCL
- LCL identifiability condition

Other things in our paper

Kiran Tomlinson and Austin R. Benson
Learning Interpretable Feature Context Effects in Discrete Choice
***arXiv: 2009.03417*, September 2020**
bit.ly/lcl-paper

- LCL derivation from simple assumptions
- More flexible model: decomposed LCL
- LCL identifiability condition
- Application to general choice data

$$\Pr(i, C) = \sum_{k=1}^{d} \pi_k \frac{\exp\left([B_k + A_k(x_C)_k]^T x_i\right)}{\sum_{j \in C} \exp\left([B_k + A_k(x_C)_k]^T x_j\right)}$$

Theorem 1. A d-feature linear context logit is identifiable from a dataset $\mathcal{D}$ if and only if

$$\operatorname{span}\left\{\begin{bmatrix} x_C \\ 1 \end{bmatrix} \otimes (x_i - x_C) \mid C \in \mathcal{C}_{\mathcal{D}}, i \in C\right\} = \mathbb{R}^{d^2+d}. \tag{6}$$

Dataset
DISTRICT
DISTRICT-SMART
SUSHI
EXPEDIA
CAR-A
CAR-B
CAR-ALT

Other things in our paper

$$\Pr(i, C) = \sum_{k=1}^{d} \pi_k \frac{\exp\left([B_k + A_k(x_C)_k]^T x_i\right)}{\sum_{j \in C} \exp\left([B_k + A_k(x_C)_k]^T x_j\right)}$$

Kiran Tomlinson and Austin R. Benson
Learning Interpretable Feature Context Effects in Discrete Choice
***arXiv: 2009.03417*, September 2020**
bit.ly/lcl-paper

Theorem 1. A d-feature linear context logit is identifiable from a dataset $\mathcal{D}$ if and only if

$$\text{span}\left\{\begin{bmatrix} x_C \\ 1 \end{bmatrix} \otimes (x_i - x_C) \mid C \in \mathcal{C}_{\mathcal{D}}, i \in C\right\} = \mathbb{R}^{d^2+d}. \qquad (6)$$

- LCL derivation from simple assumptions
- More flexible model: decomposed LCL
- LCL identifiability condition
- Application to general choice data
- Accounting for context improves prediction

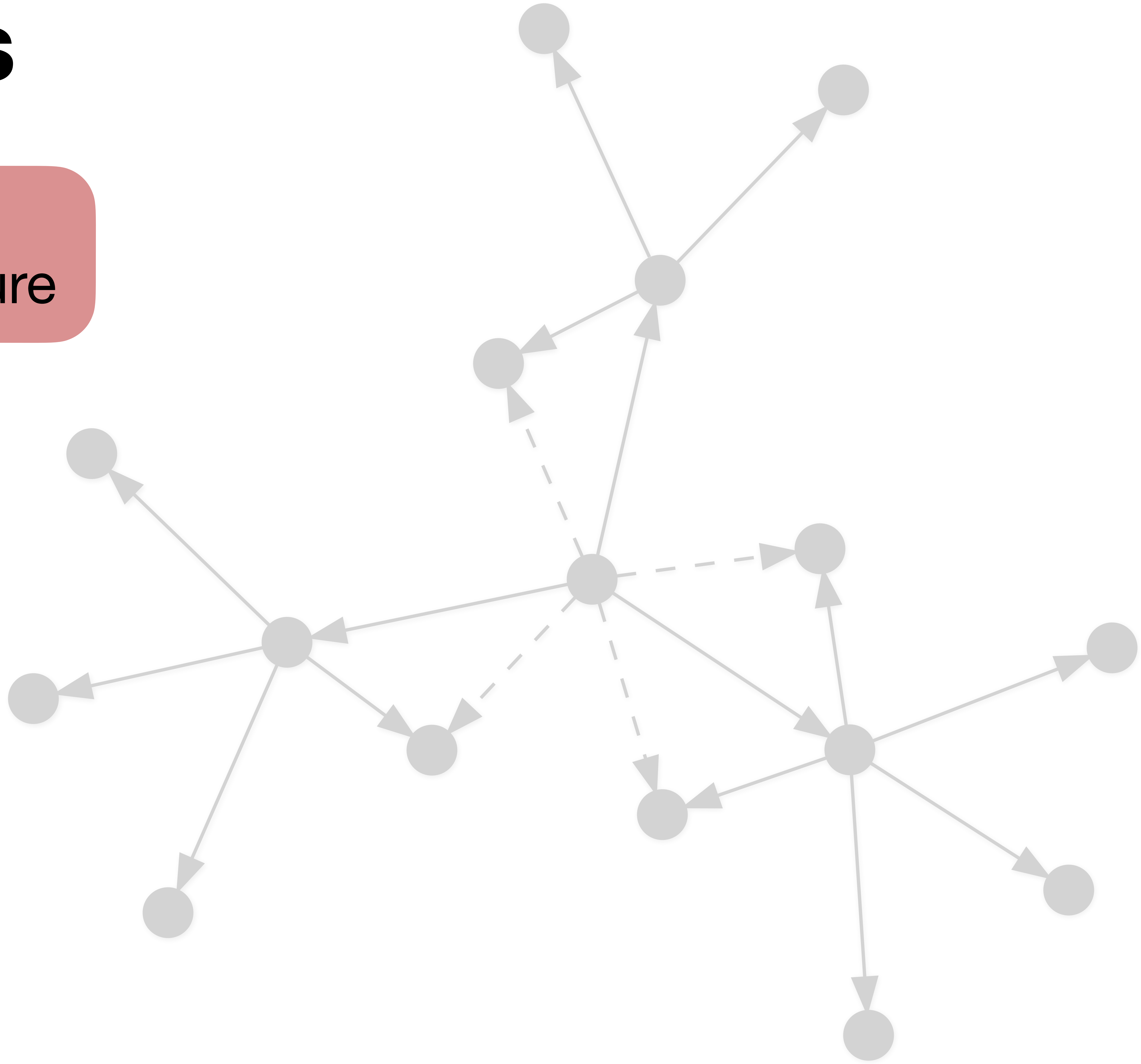
Dataset
DISTRICT
DISTRICT-SMART
SUSHI
EXPEDIA
CAR-A
CAR-B
CAR-ALT

	MNL	**LCL**
DISTRICT	.3680 (.4823)	.3327 (.4712)
DISTRICT-SMART	.4006 (.4900)	.3894 (.4876)
EXPEDIA	.3859 (.2954)	.3696* (.2926)
SUSHI	.2727 (.2751)	.2741 (.2771)
CAR-A	.3570 (.4791)	**.3514** (.4774)
CAR-B	.3326 (.4711)	.3326 (.4711)
CAR-ALT	.2944 (.2875)	**.2650*** (.2804)
SYNTHETIC-MNL	.1513 (.1865)	**.1512** (.1864)
SYNTHETIC-LCL	.1360 (.1684)	**.1357*** (.1683)
WIKI-TALK	.2946 (.2916)	**.2666*** (.2773)
REDDIT-HYPERLINK	.2859 (.2611)	.2761* (.2606)
BITCOIN-ALPHA	.2724 (.3246)	.2591* (.3178)
BITCOIN-OTC	.1891 (.2756)	.1529* (.2468)
SMS-A	.2825 (.3250)	**.2661*** (.3193)
SMS-B	.3045 (.3419)	**.2848*** (.3273)
SMS-C	.3115 (.3455)	**.3070** (.3477)
EMAIL-ENRON	.1265 (.2068)	.1244* (.2115)
EMAIL-EU	.2683 (.3021)	.2665 (.3037)
EMAIL-W3C	.1332 (.2070)	.1210* (.1845)
FACEBOOK-WALL	.2176 (.2895)	.2109* (.2871)
COLLEGE-MSG	.1850 (.2726)	**.1723*** (.2655)
MATHOVERFLOW	.1385 (.2503)	.1153* (.2200)

Concluding thoughts

Key takeaway
Context effects matter in triadic closure

Challenges
Features correlate
Causal context effects?
Handling nonlinearity?
Global edge formation modes?
Missing timestamps?



Concluding thoughts

Slides: bit.ly/lcl-slides
Preprint: bit.ly/lcl-paper
Code: bit.ly/lcl-code
Data: bit.ly/lcl-data

Key takeaway
Context effects matter in triadic closure

Challenges
Features correlate
Causal context effects?
Handling nonlinearity?
Global edge formation modes?
Missing timestamps?

Thank you!

More questions or ideas?
Email me: kt@cs.cornell.edu

Acknowledgments
Funding from NSF, ARO

Thanks to Johan Ugander and Jan Overgoor