Simultaneously Approximating All ℓ_p-norms in Correlation Clustering

Heather Newman

Carnegie Mellon University (CMU)

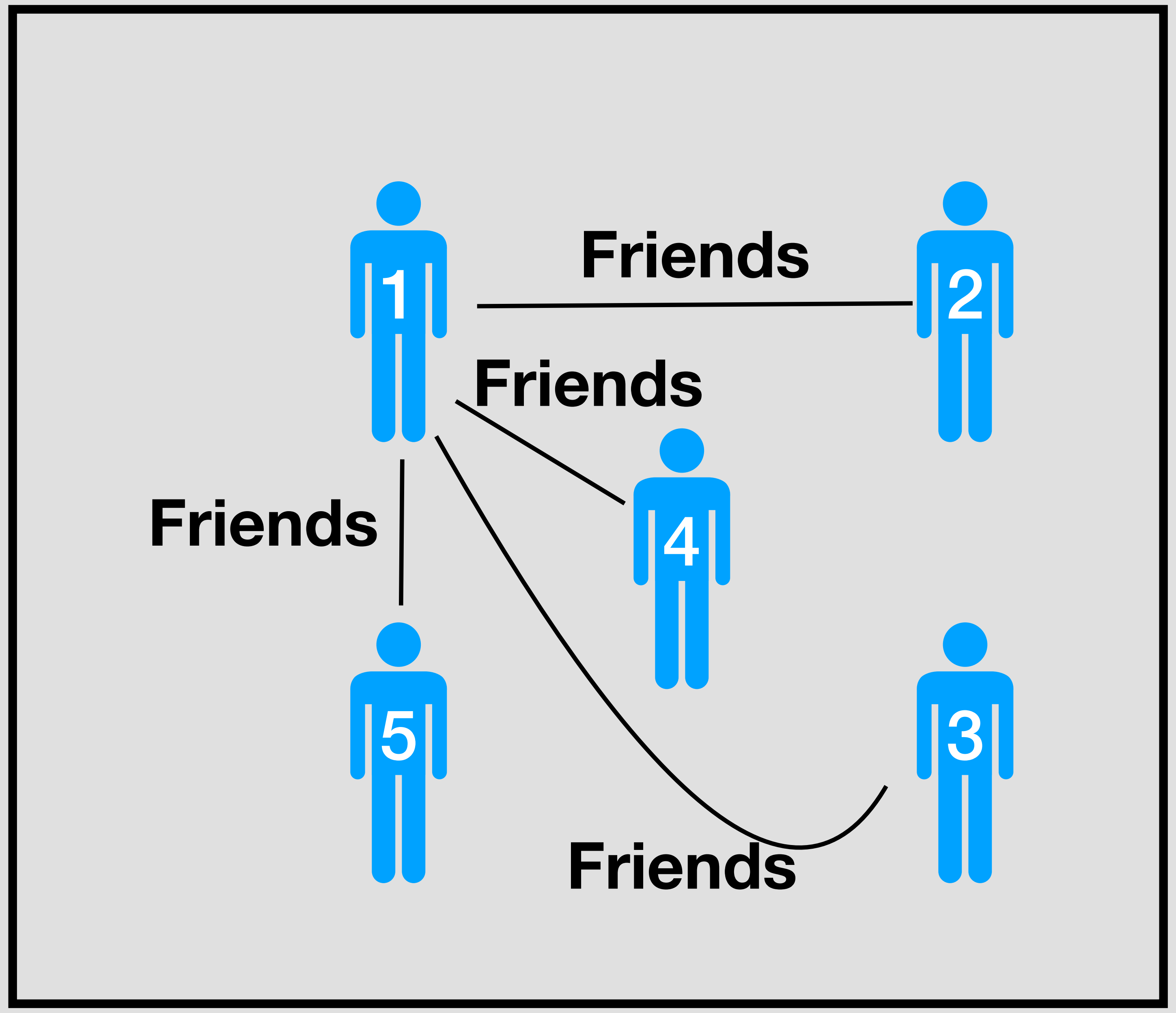
ICALP 2024

Joint work with: Sami Davies* (UC Berkeley/Simons) and Benjamin Moseley (CMU)

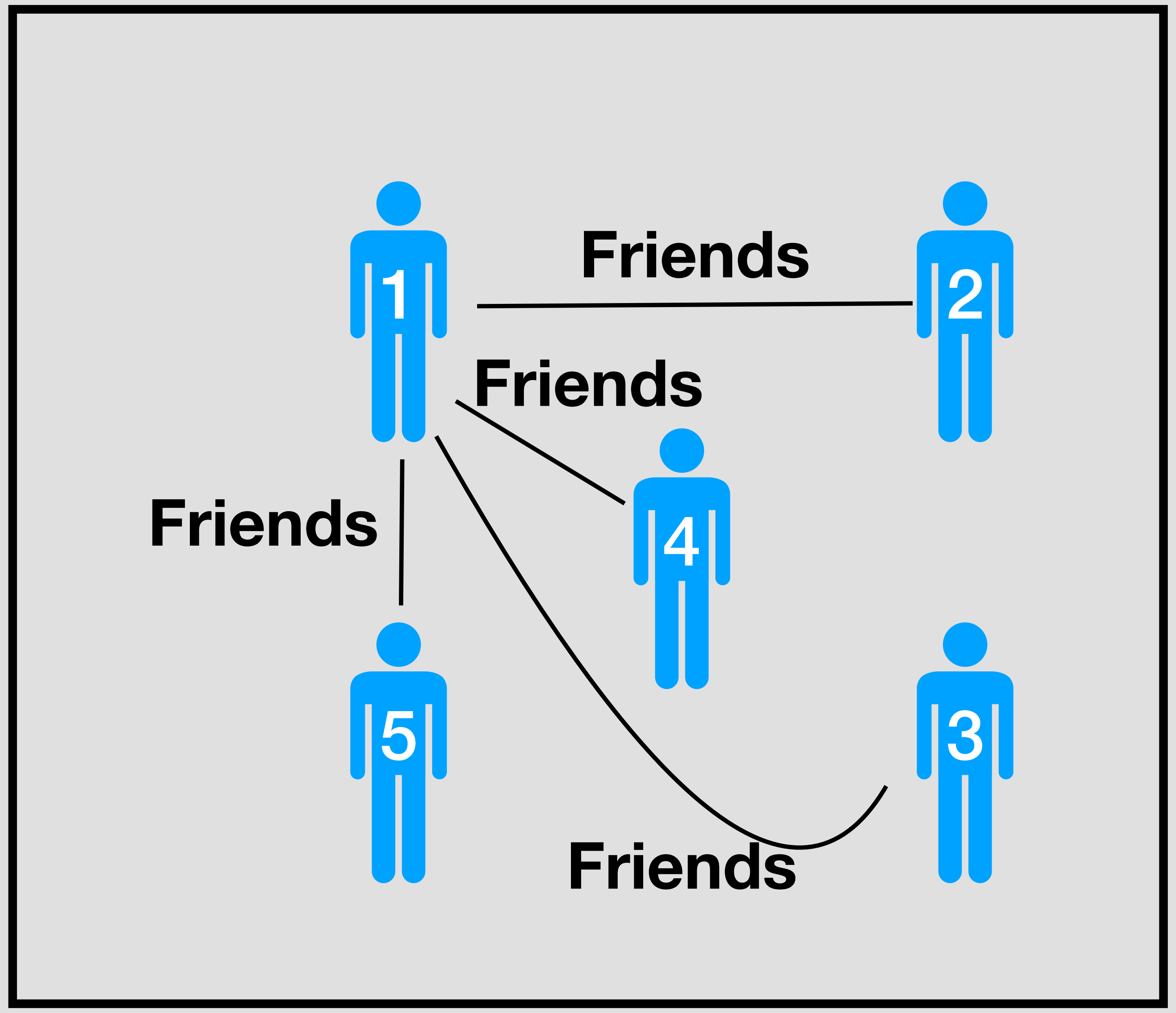
*Special thanks to Sami Davies for contributions to the slide deck.

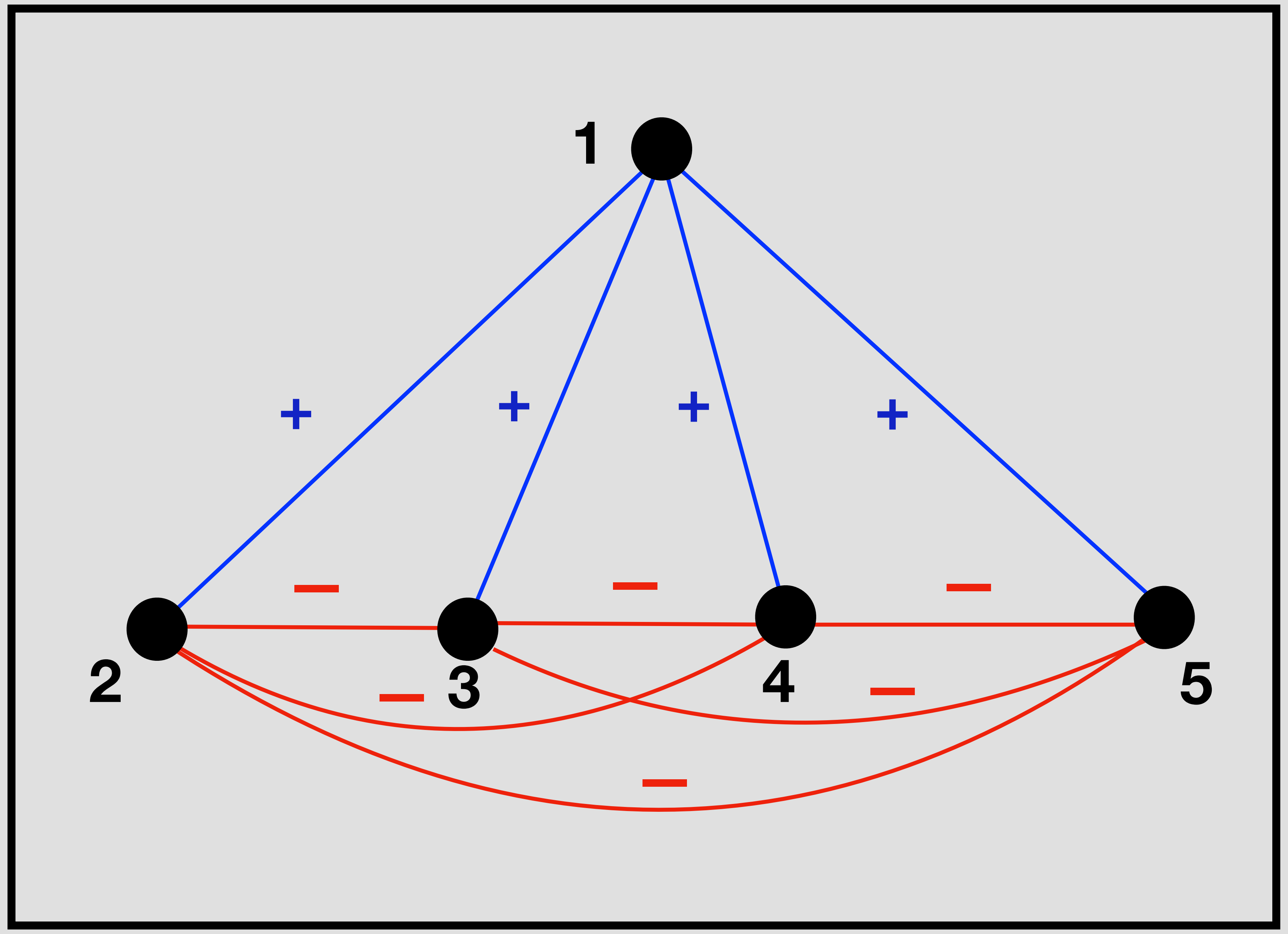
Correlation clustering

Correlation clustering



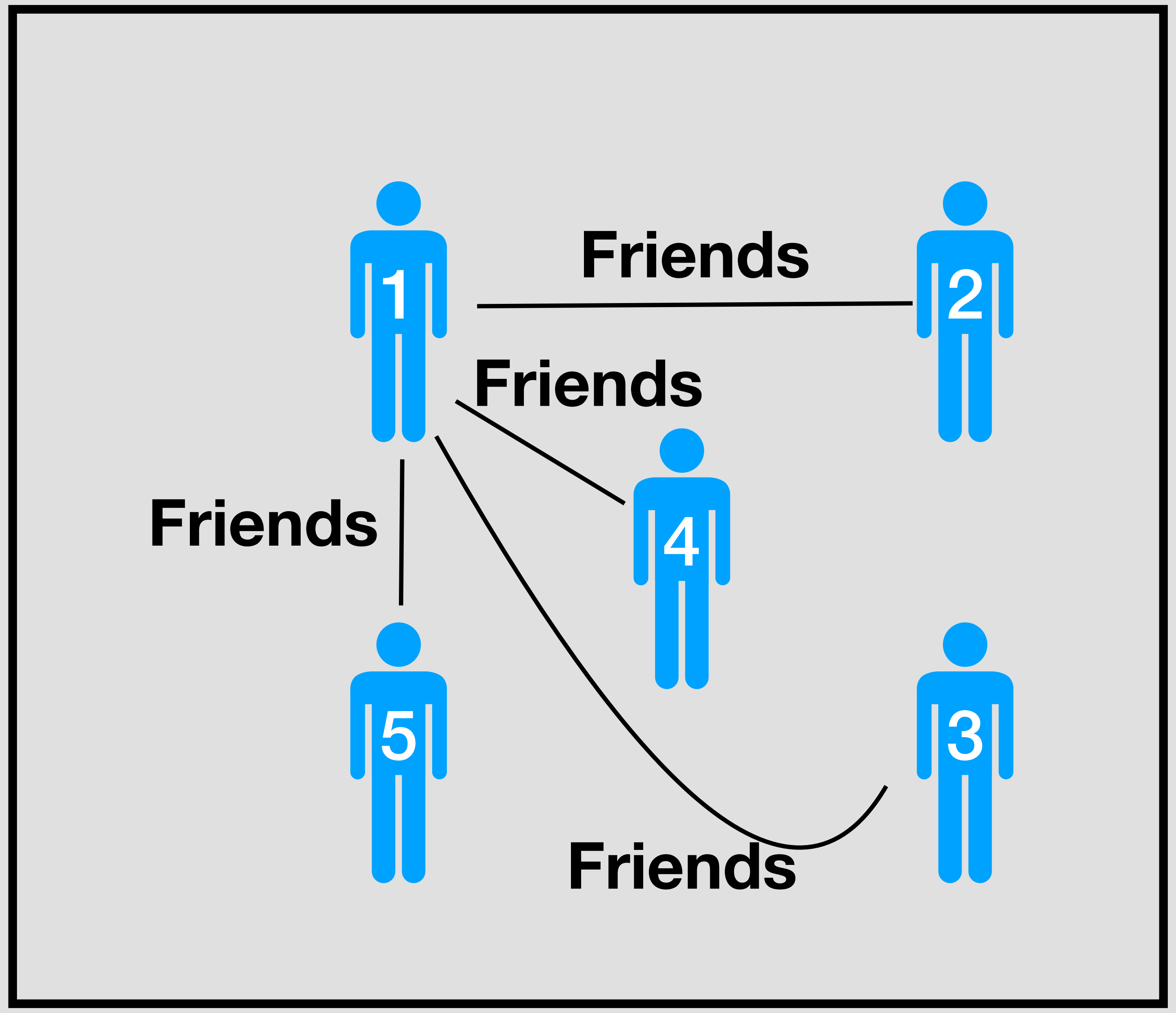
Correlation clustering

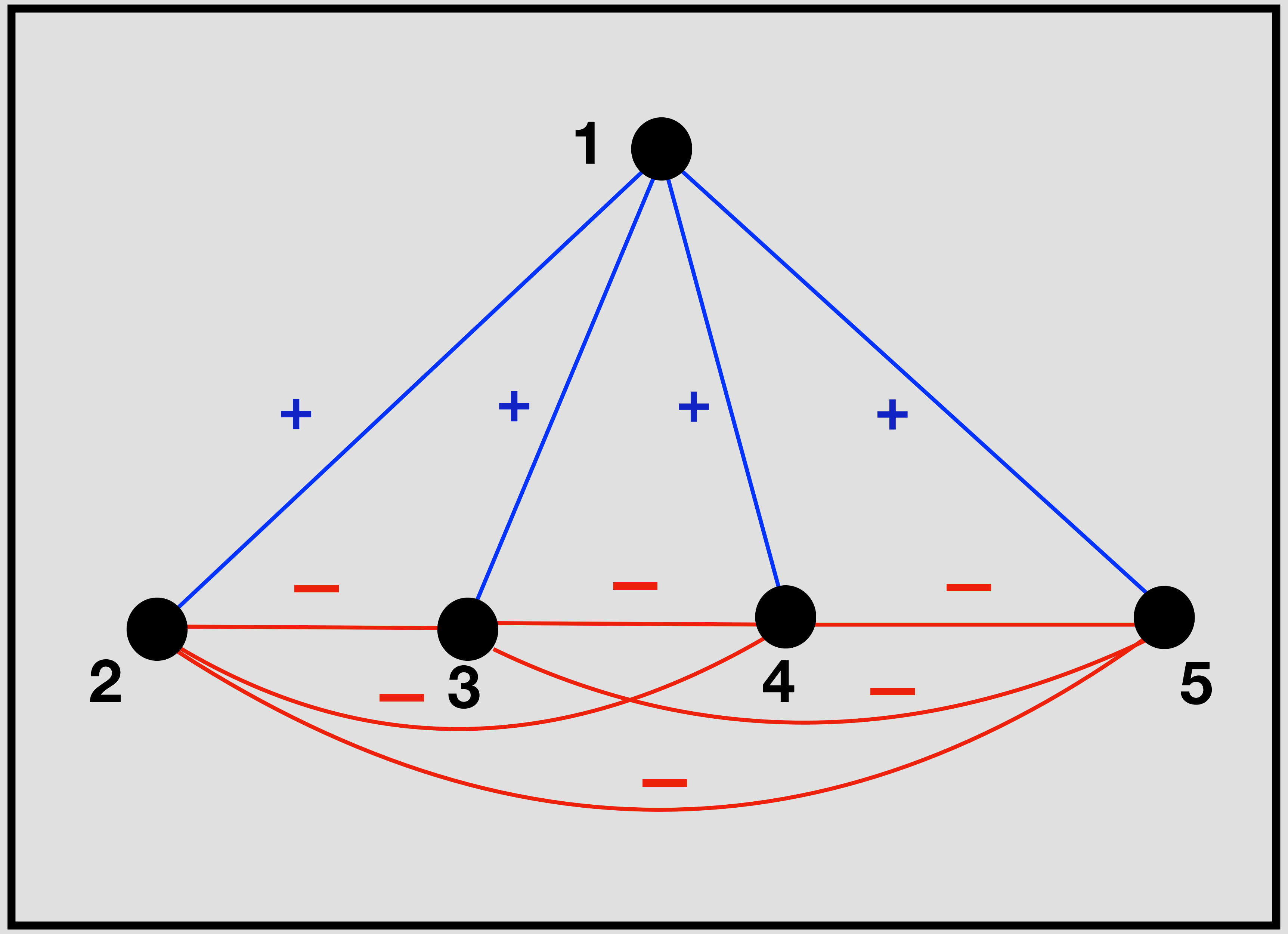




Correlation clustering

Model:

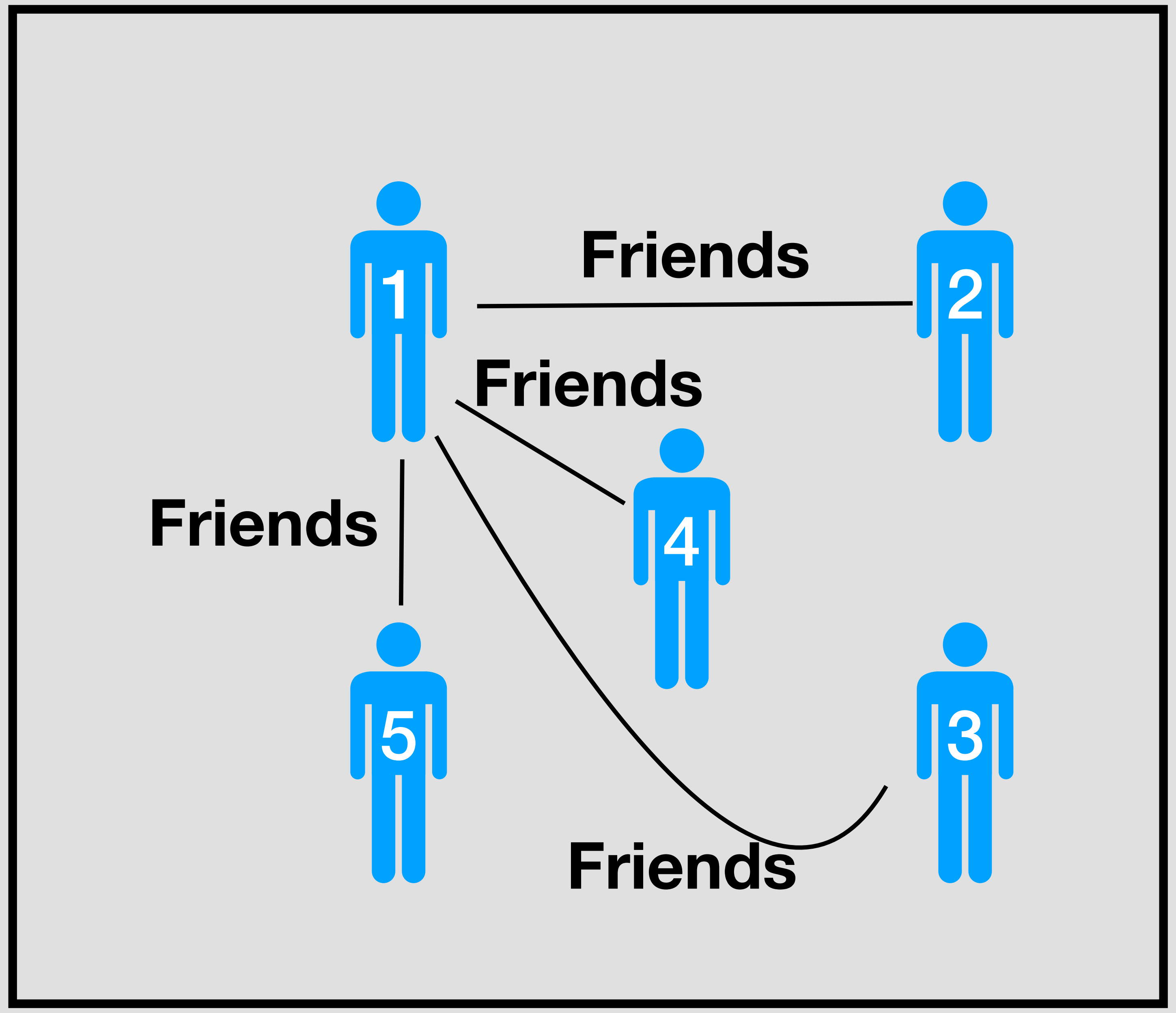


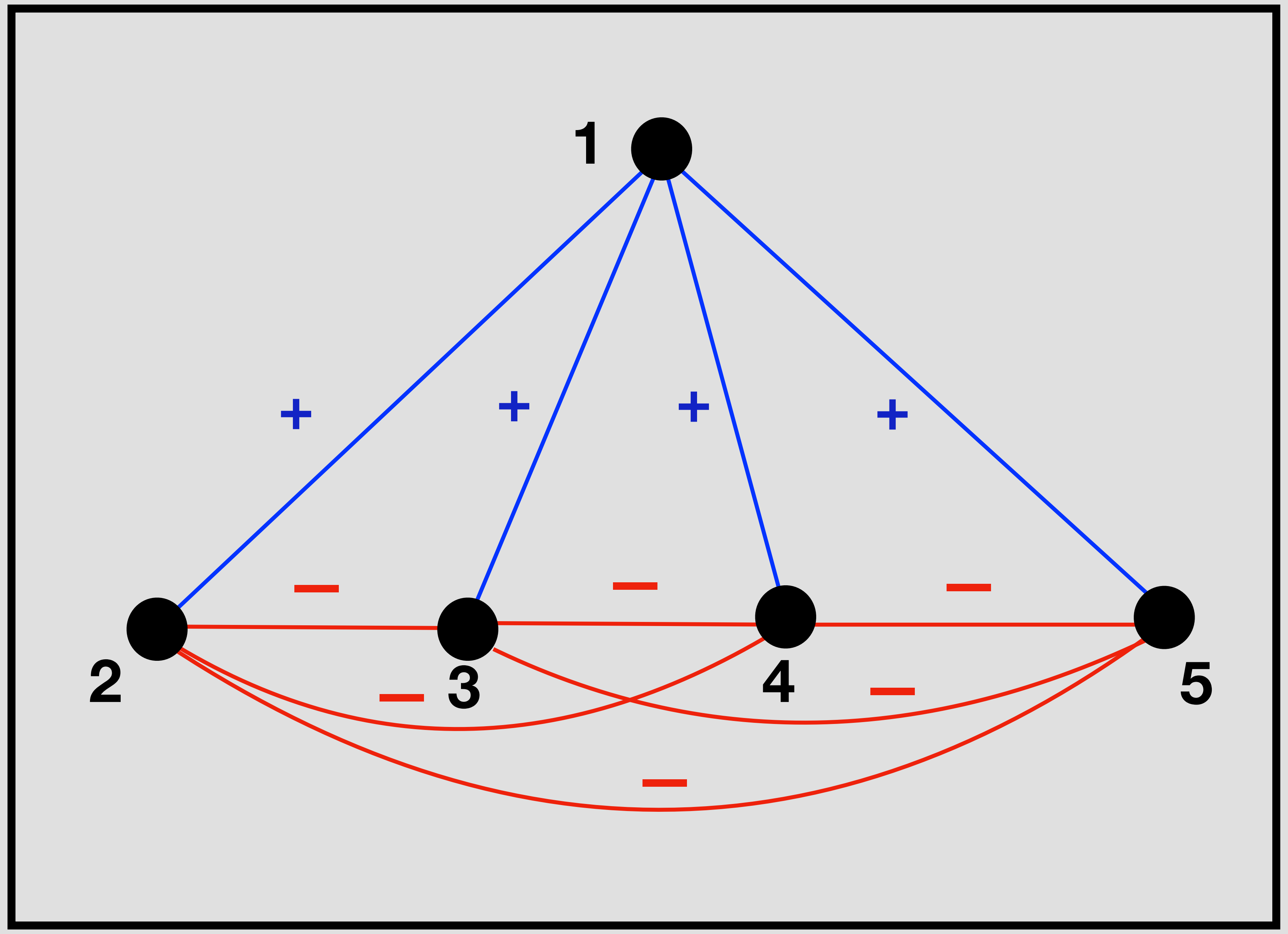


Correlation clustering

Model:

- Cluster ***similar*** nodes together, separate ***dissimilar*** nodes

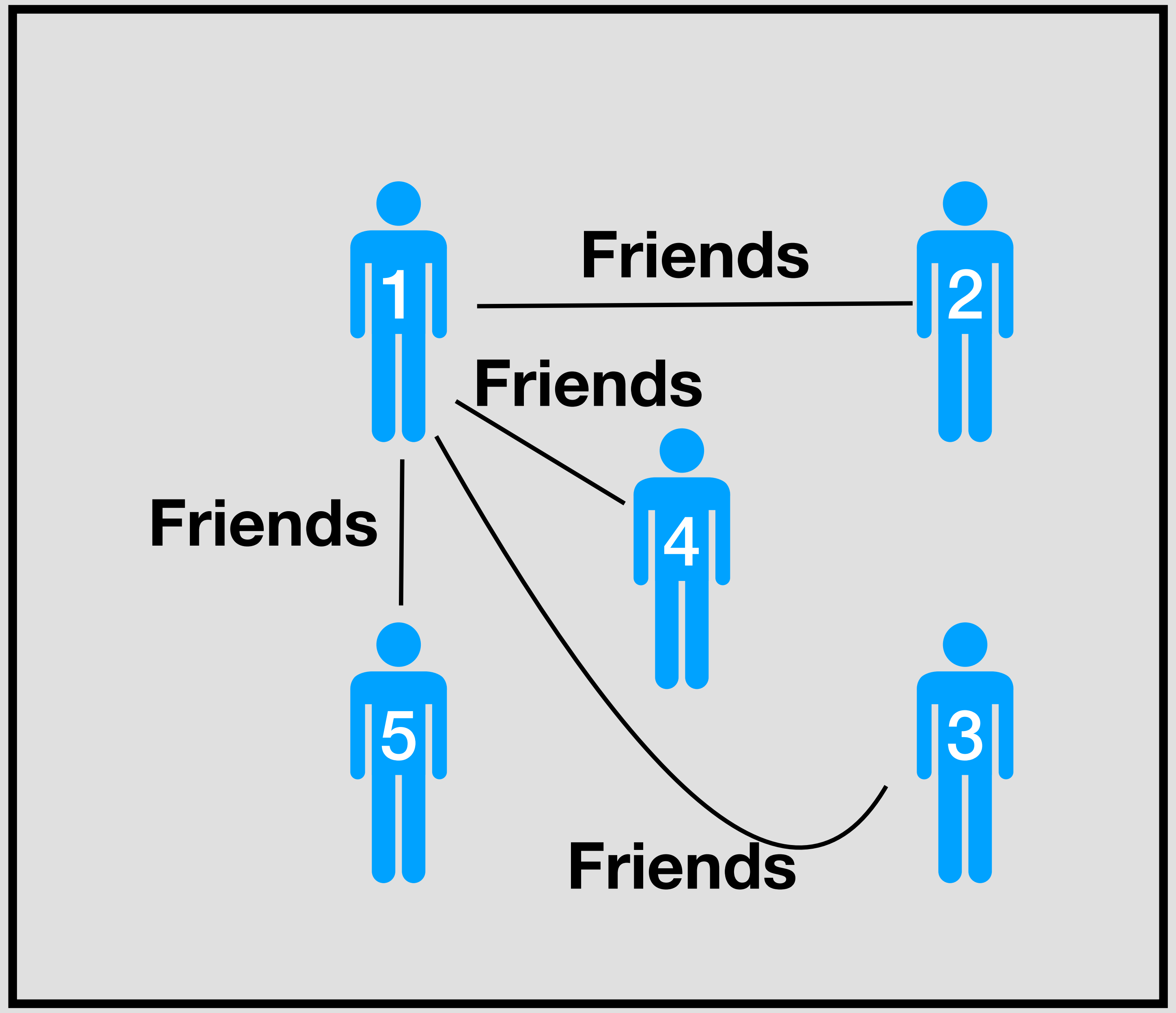


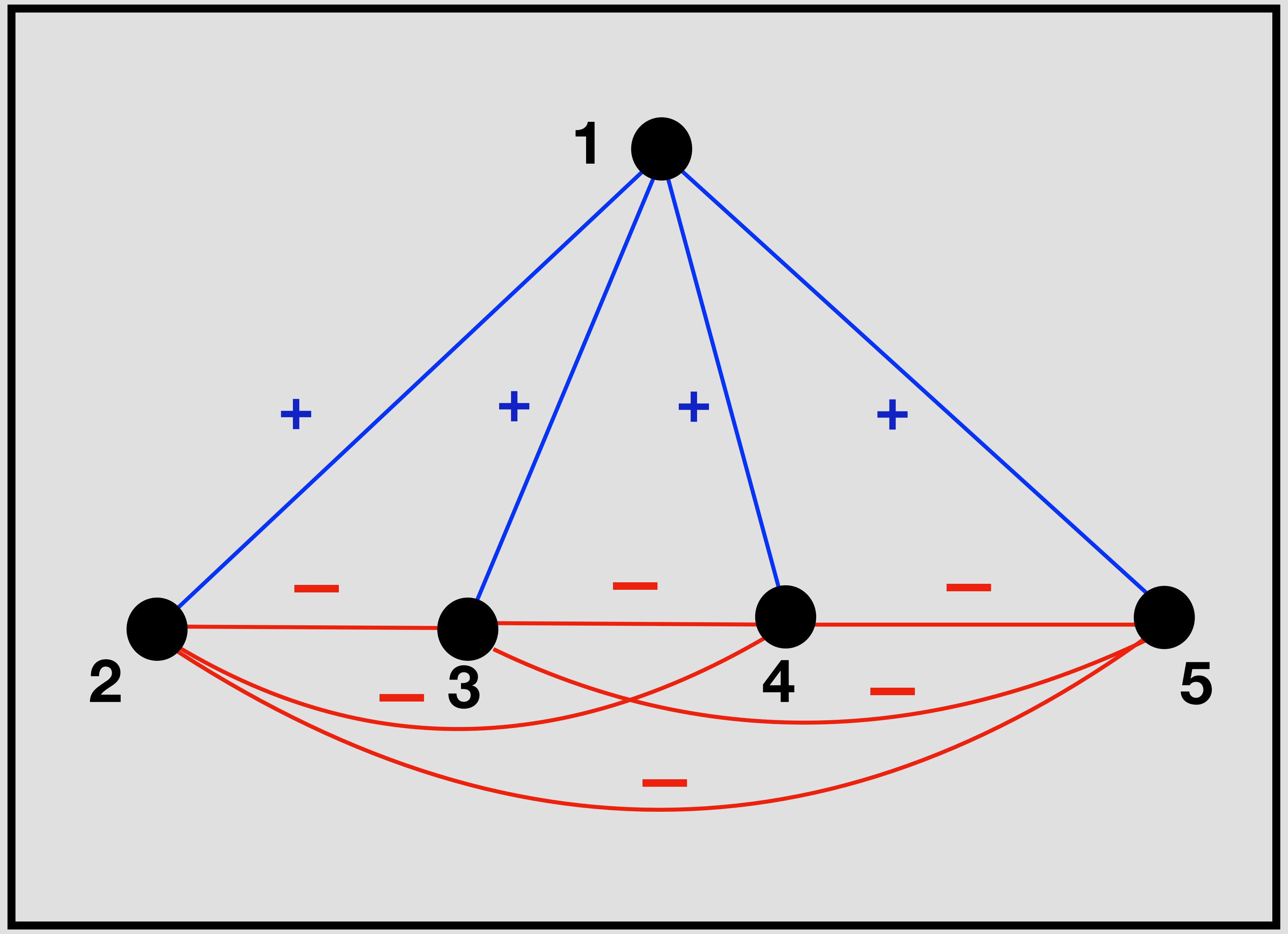


Correlation clustering

Model:

- Cluster ***similar*** nodes together, separate ***dissimilar*** nodes
- Number / sizes of clusters (of vertices) **not** pre-fixed / pre-specified

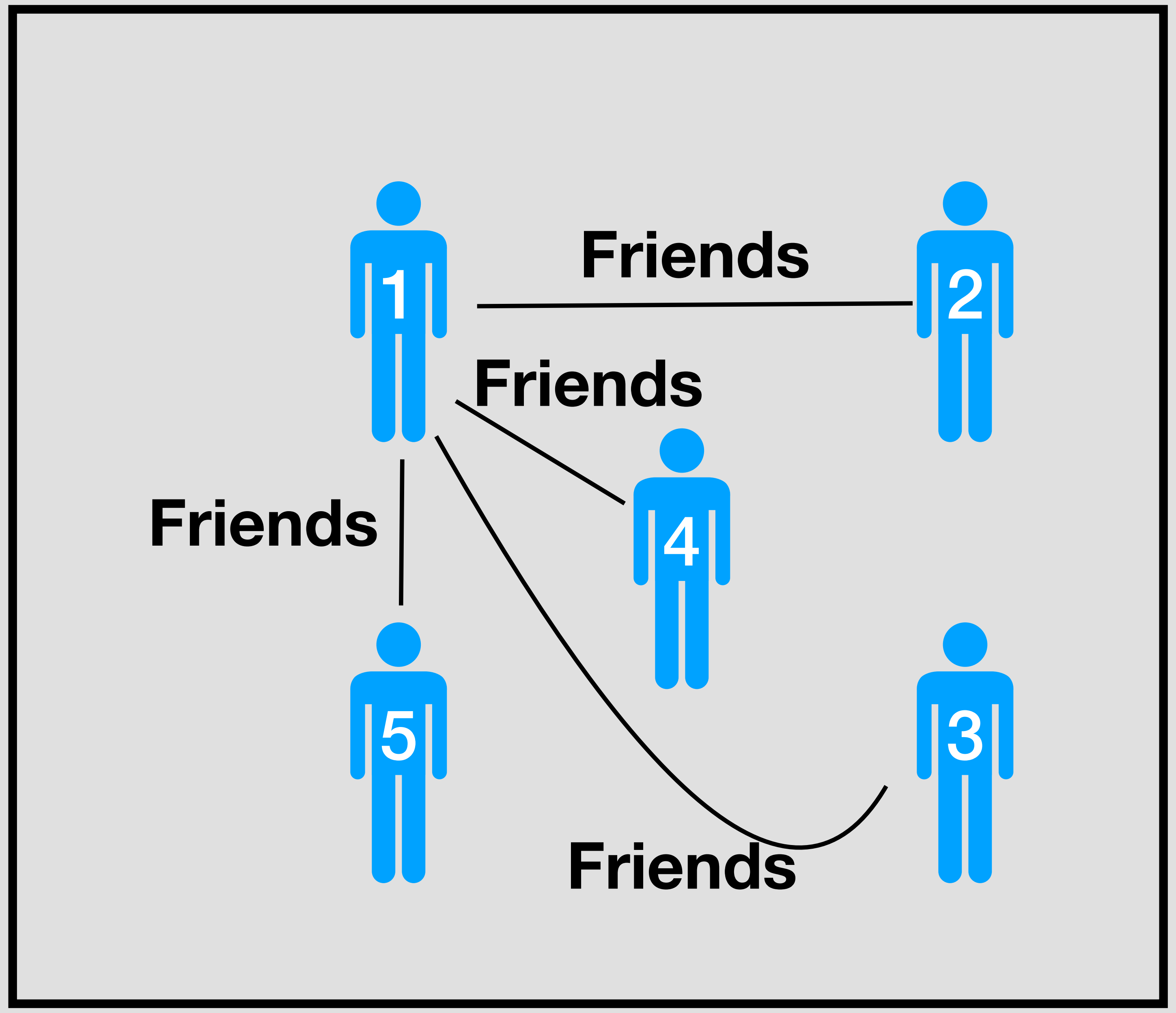


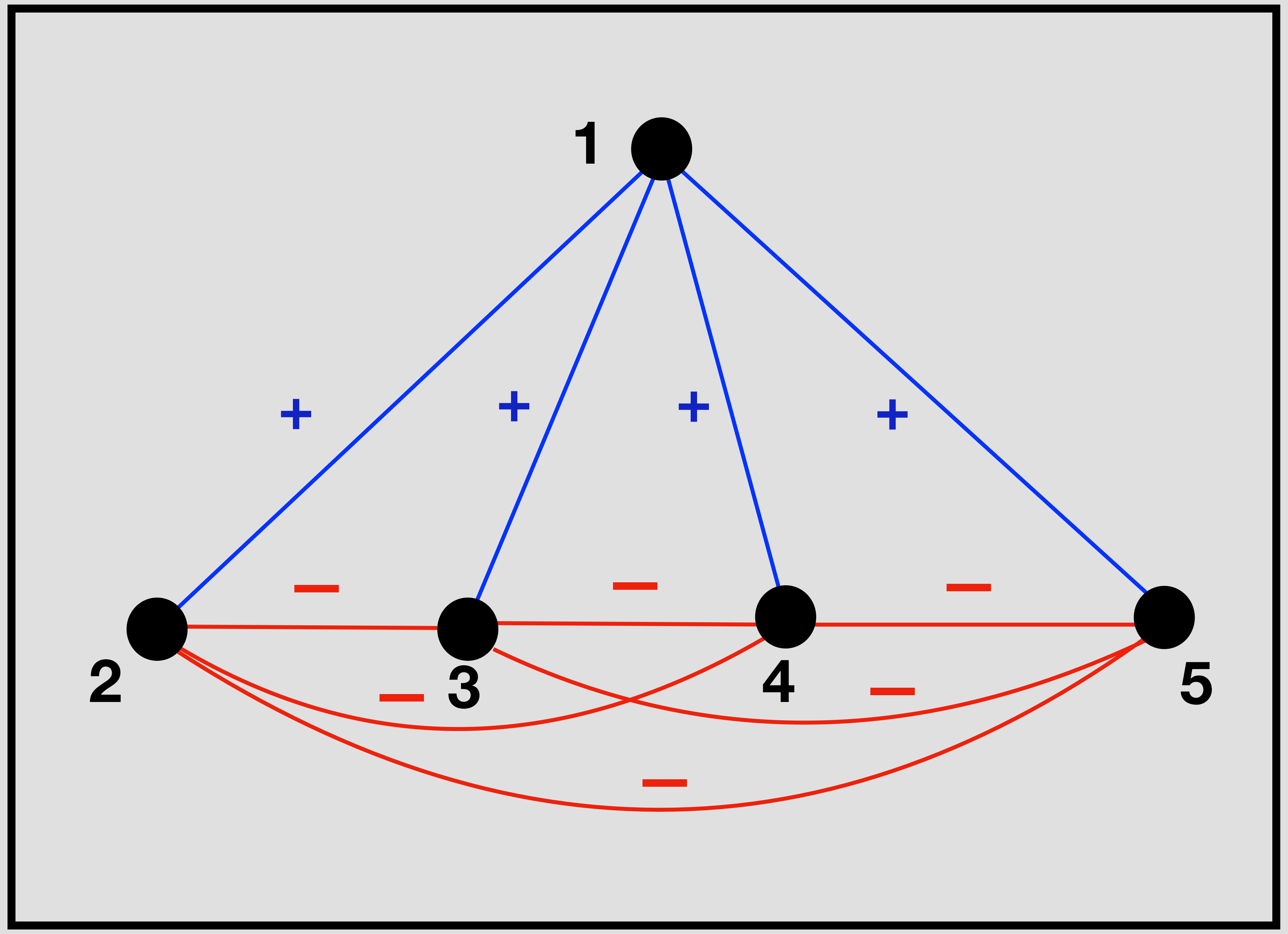


Correlation clustering

Model:

- Cluster ***similar*** nodes together, separate ***dissimilar*** nodes
- Number / sizes of clusters (of vertices) **not** pre-fixed / pre-specified
- **Complete, unweighted graph**: each edge (+) or (–)

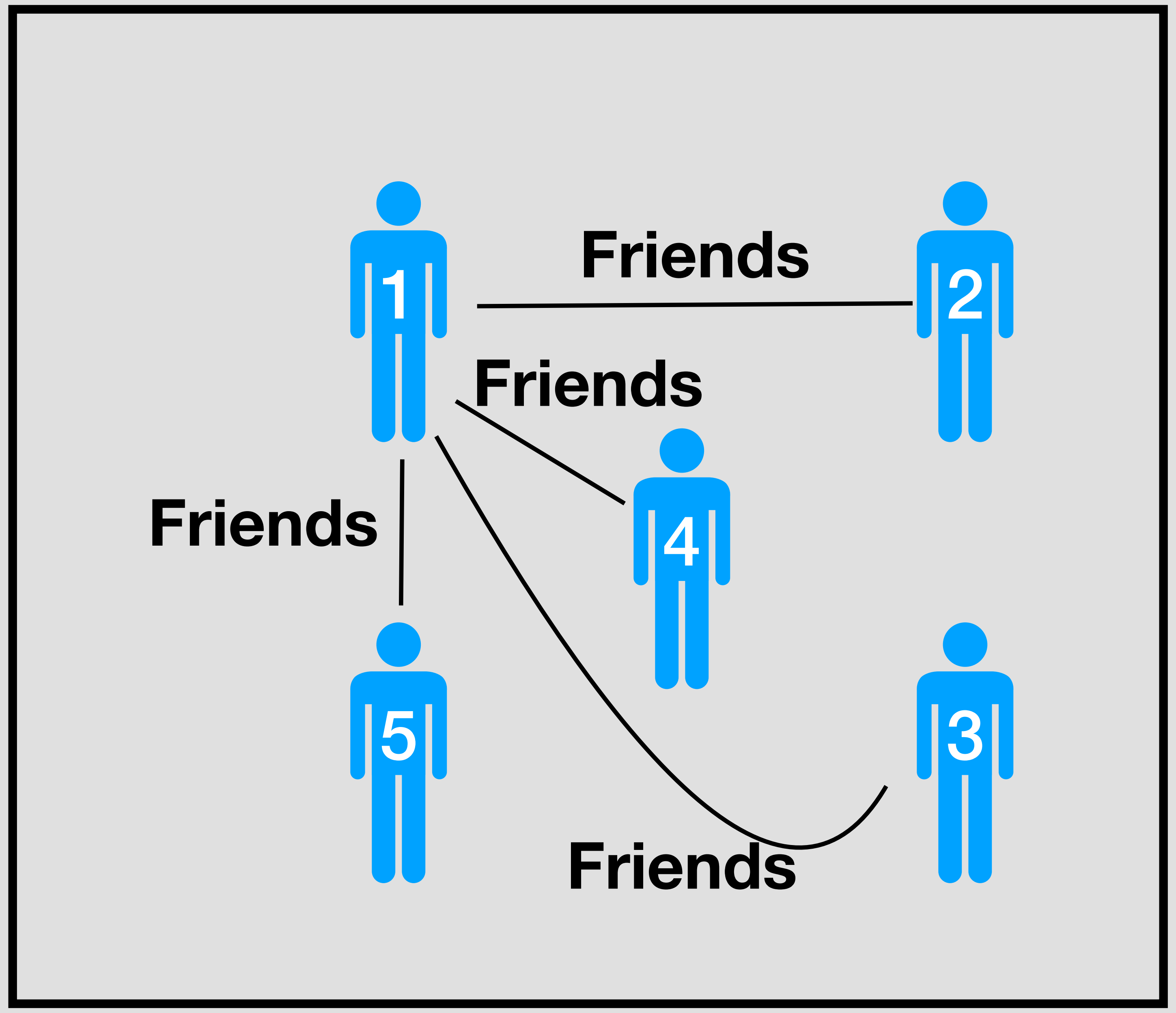


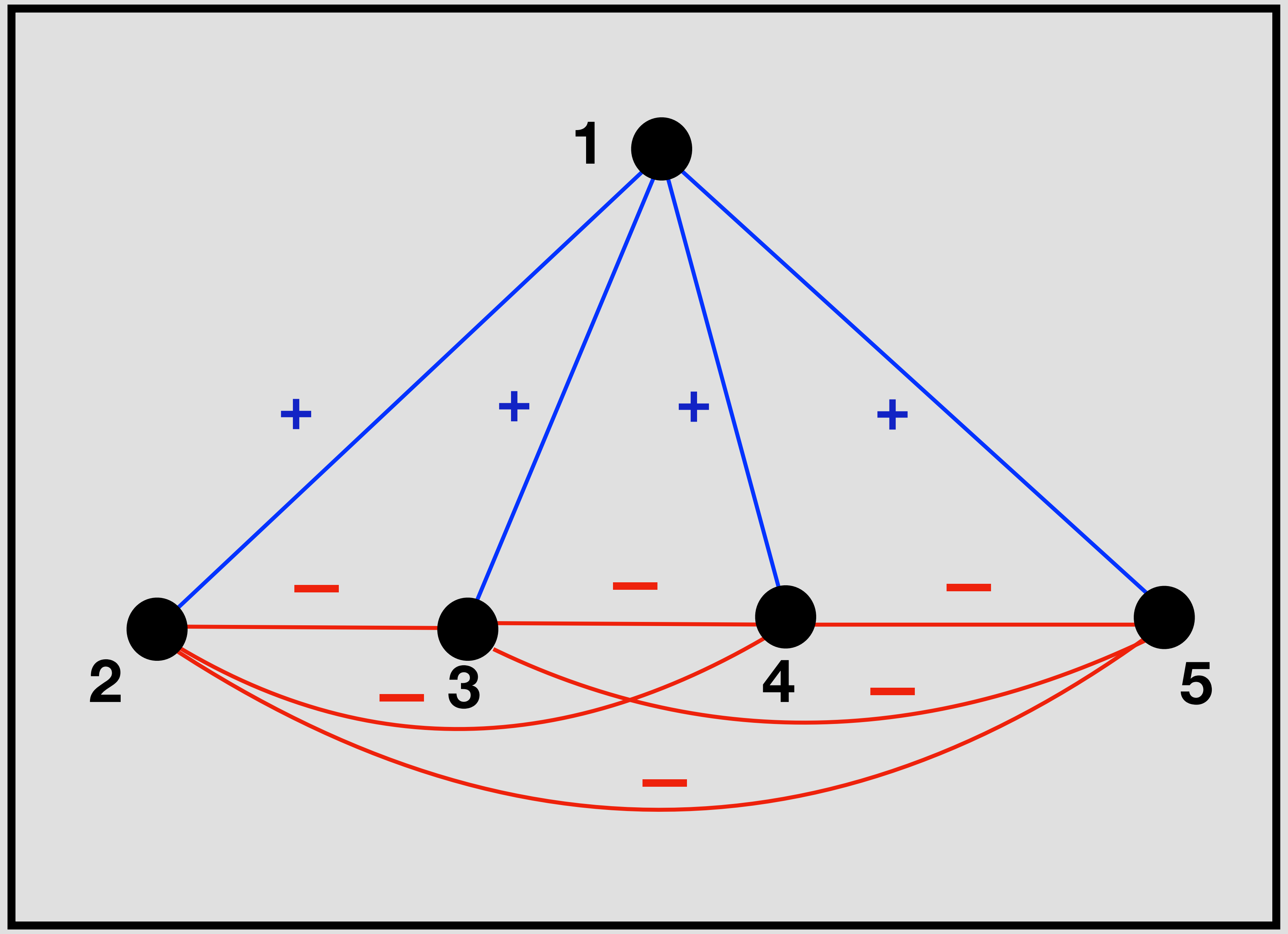


Correlation clustering

Model:

- Cluster ***similar*** nodes together, separate ***dissimilar*** nodes
- Number / sizes of clusters (of vertices) **not** pre-fixed / pre-specified
- **Complete, unweighted graph**: each edge (+) or (–)

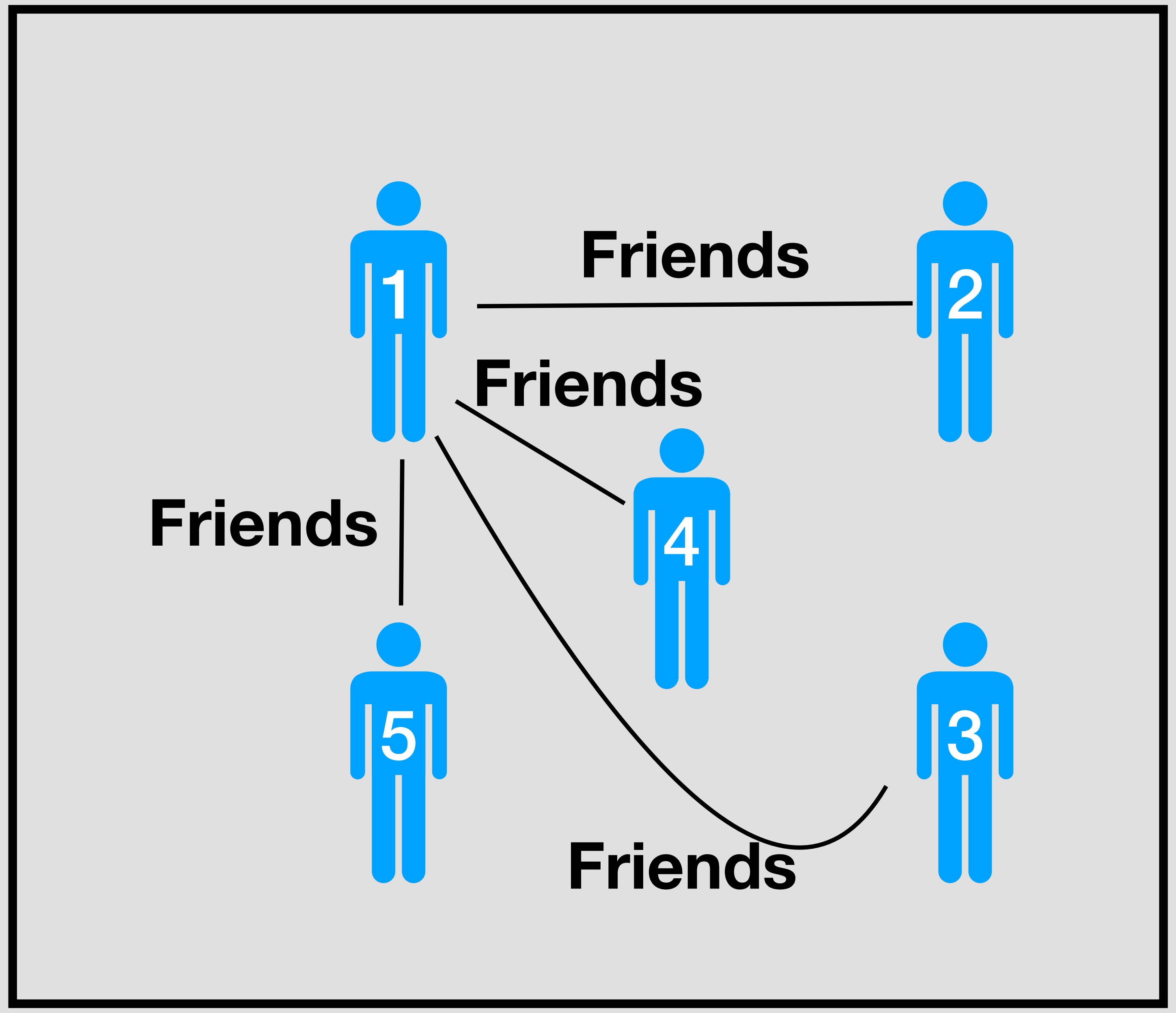


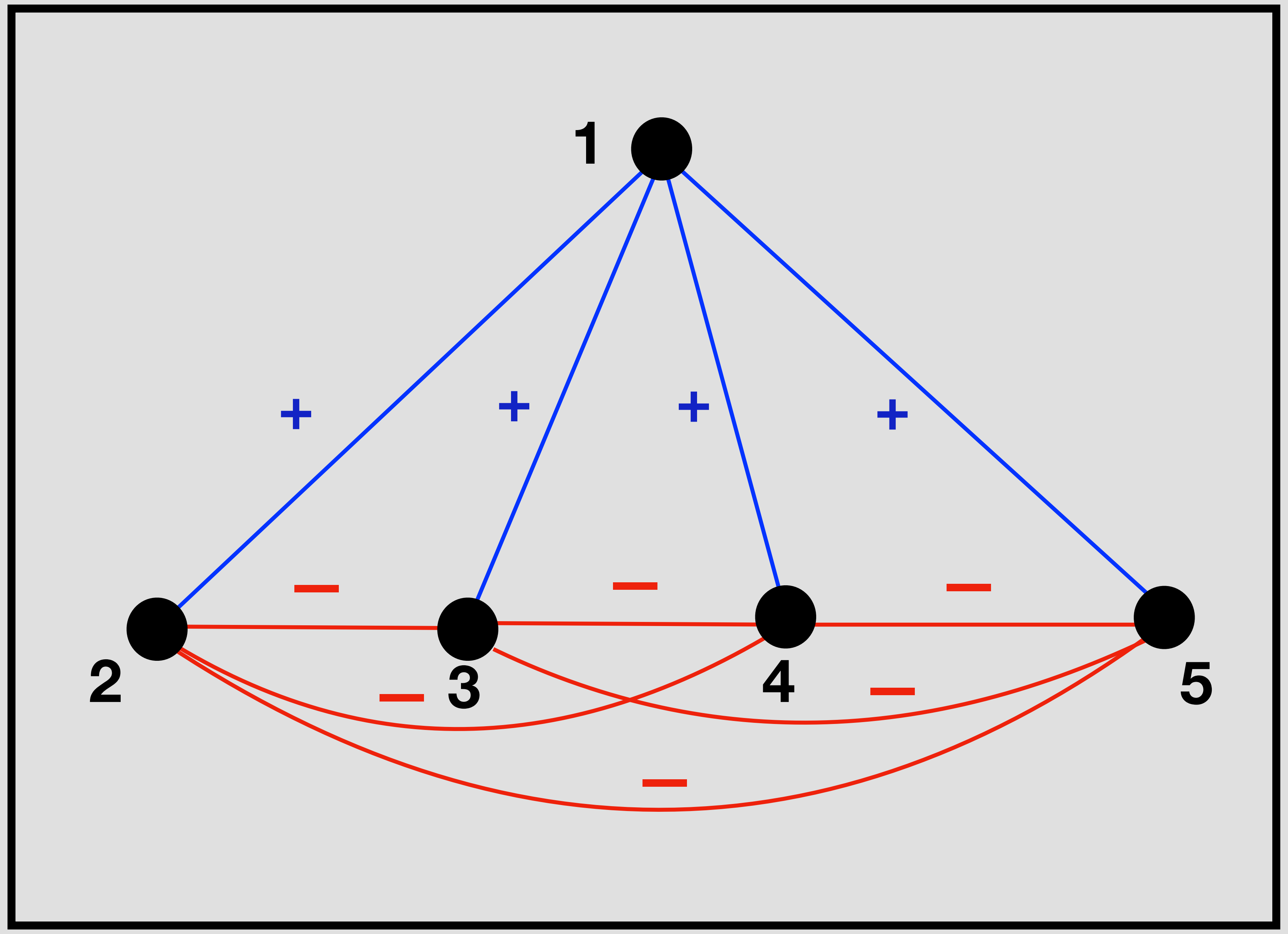


Correlation clustering

Model:

- Cluster ***similar*** nodes together, separate ***dissimilar*** nodes
- Number / sizes of clusters (of vertices) **not** pre-fixed / pre-specified
- **Complete, unweighted graph**: each edge (+) or (–)

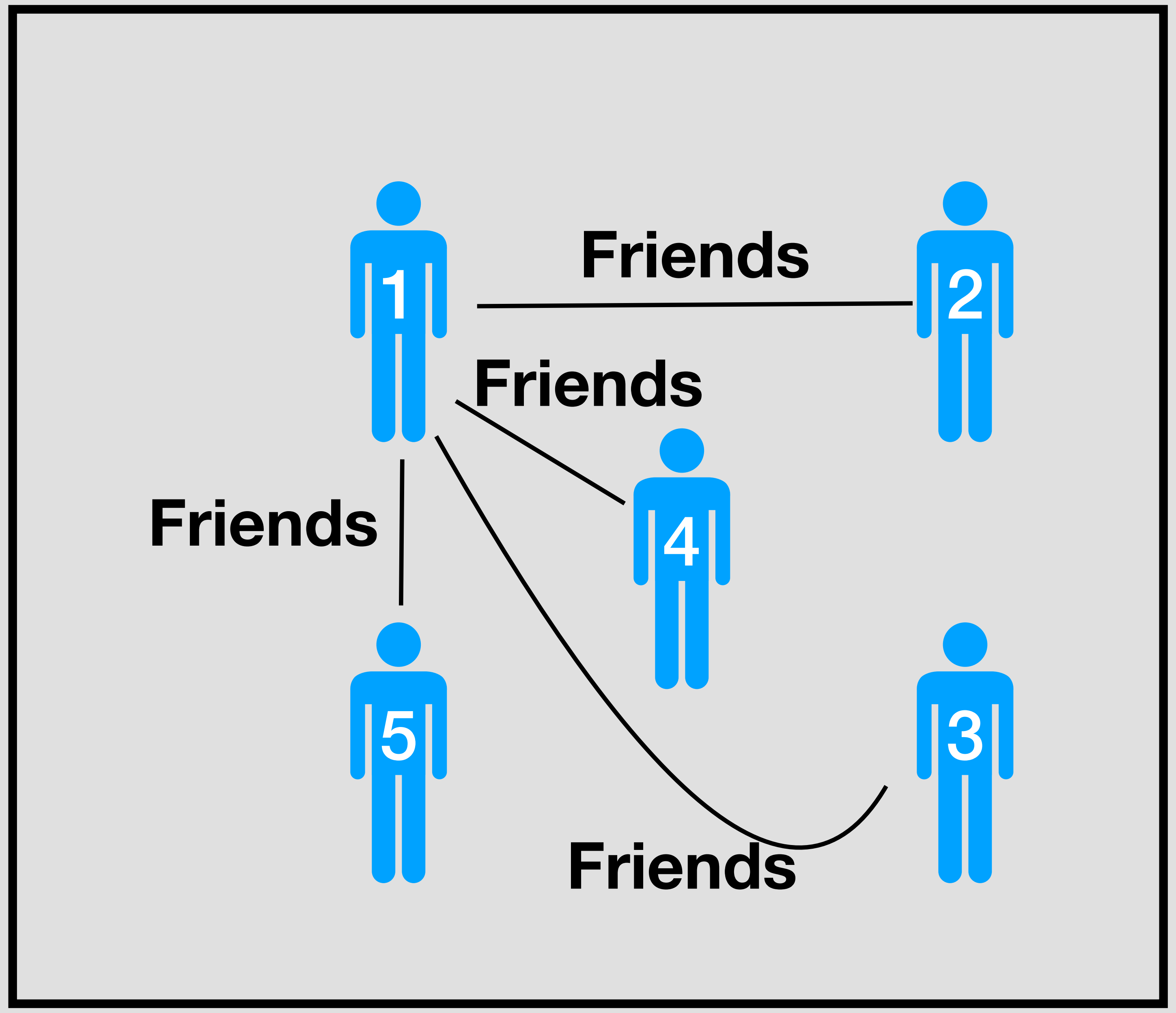


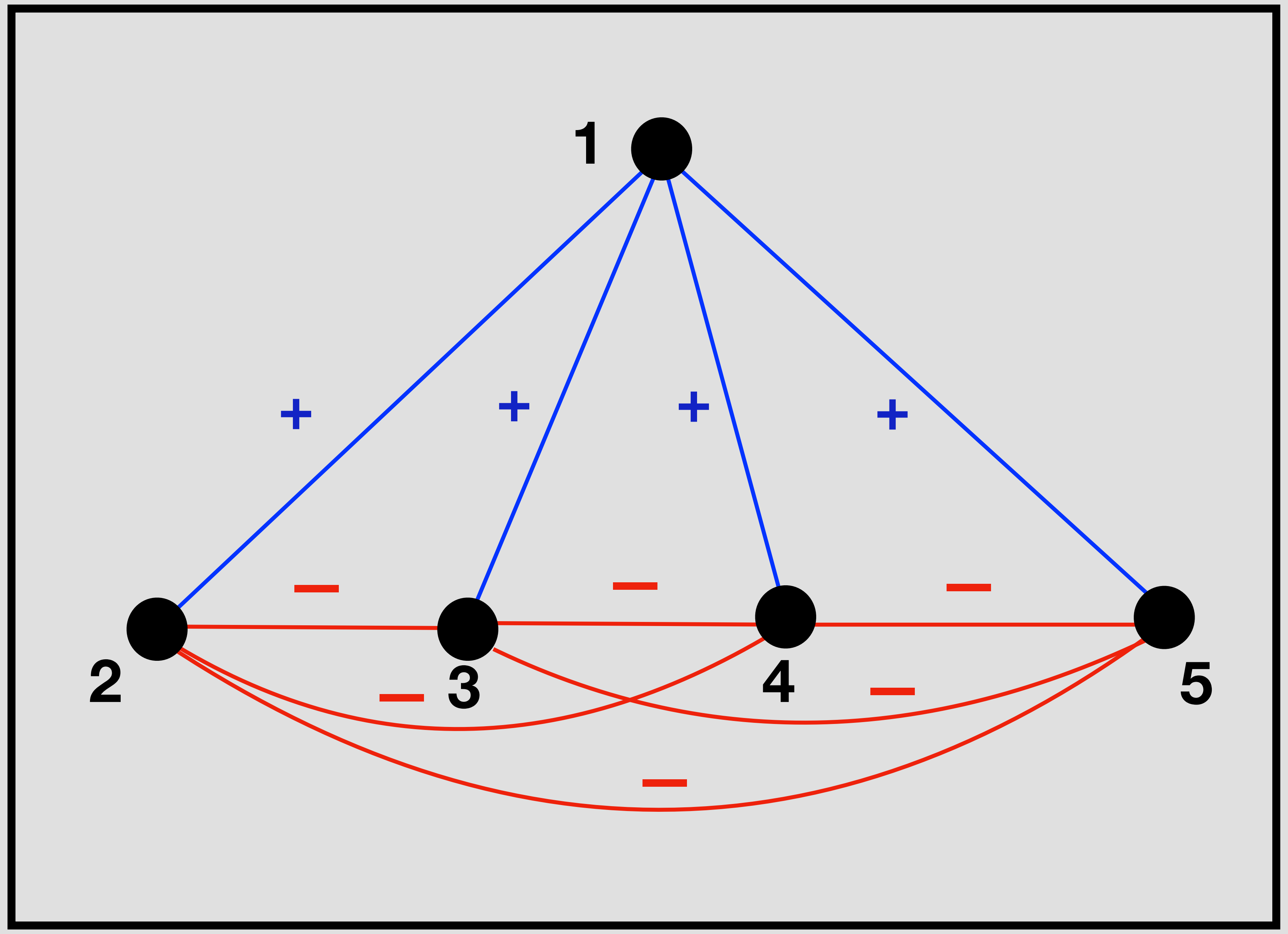


Correlation clustering

Model:

- Cluster ***similar*** nodes together, separate ***dissimilar*** nodes
- Number / sizes of clusters (of vertices) **not** pre-fixed / pre-specified
- **Complete, unweighted graph**: each edge (+) or (–)

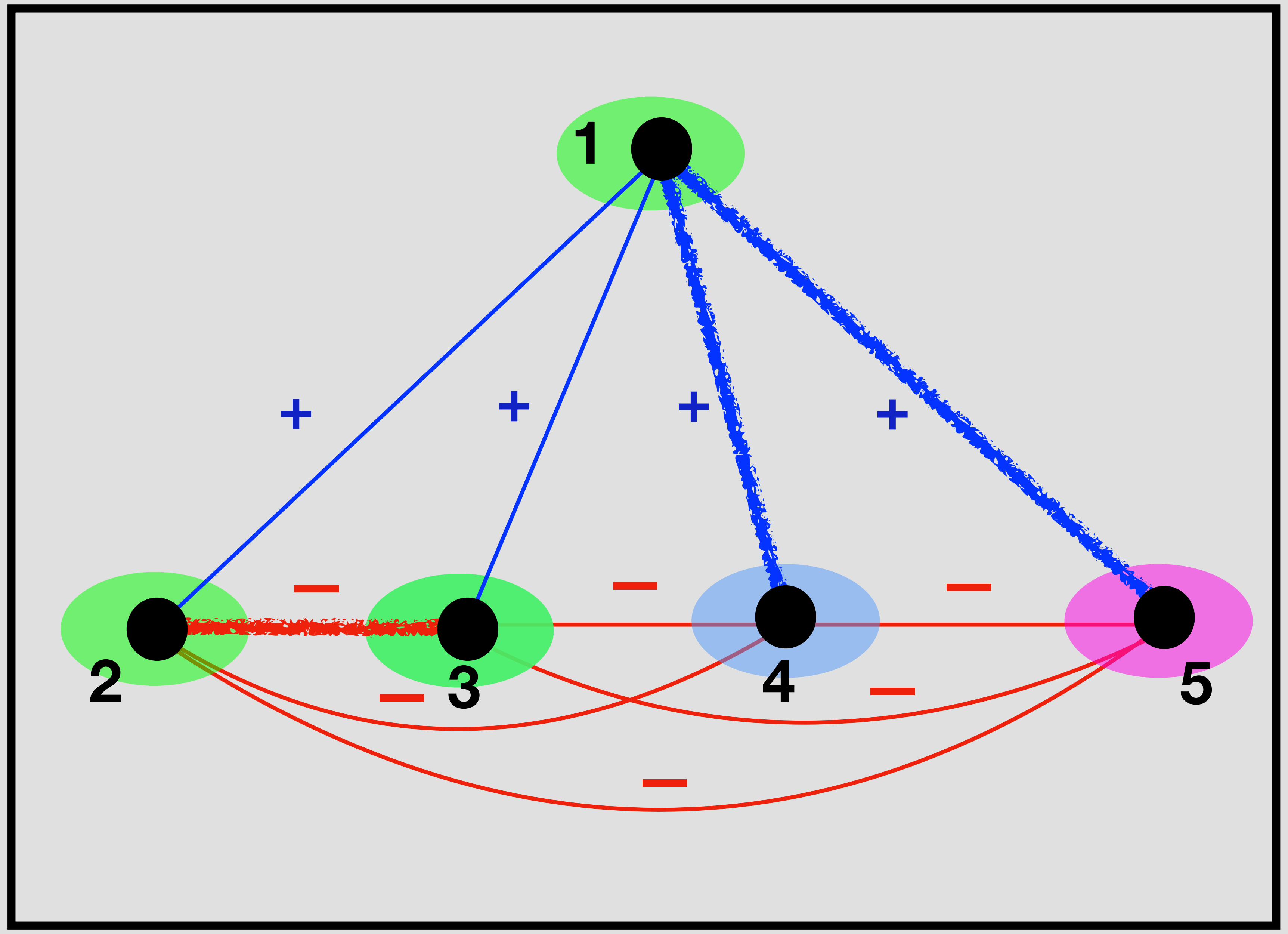




Correlation clustering

Model:

- Cluster ***similar*** nodes together, separate ***dissimilar*** nodes
- Number / sizes of clusters (of vertices) **not** pre-fixed / pre-specified
- **Complete, unweighted graph**: each edge (+) or (–)
- Edge (u,v) in ***disagreement*** w.r.t C if
  - ✦ (+) with u, v different clusters or
  - ✦ (–) with u, v same cluster

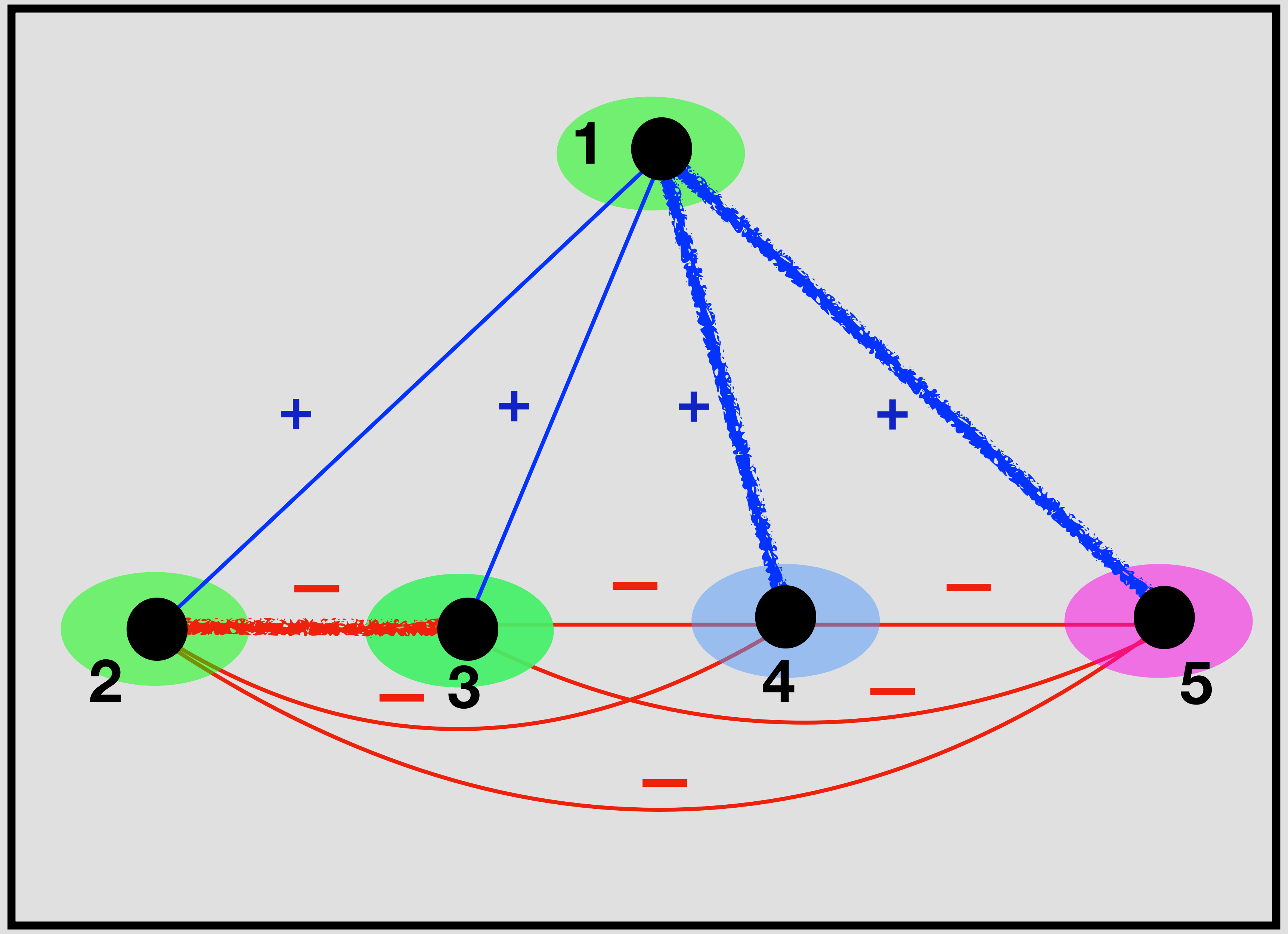


Correlation clustering

Original objective =
minimize # of edges in disagreement

Model:

- Cluster ***similar*** nodes together, separate ***dissimilar*** nodes
- Number / sizes of clusters (of vertices) **not** pre-fixed / pre-specified
- **Complete, unweighted graph**: each edge (+) or (–)
- Edge (u,v) in ***disagreement*** w.r.t C if
  - ✦ (+) with u, v different clusters or
  - ✦ (–) with u, v same cluster

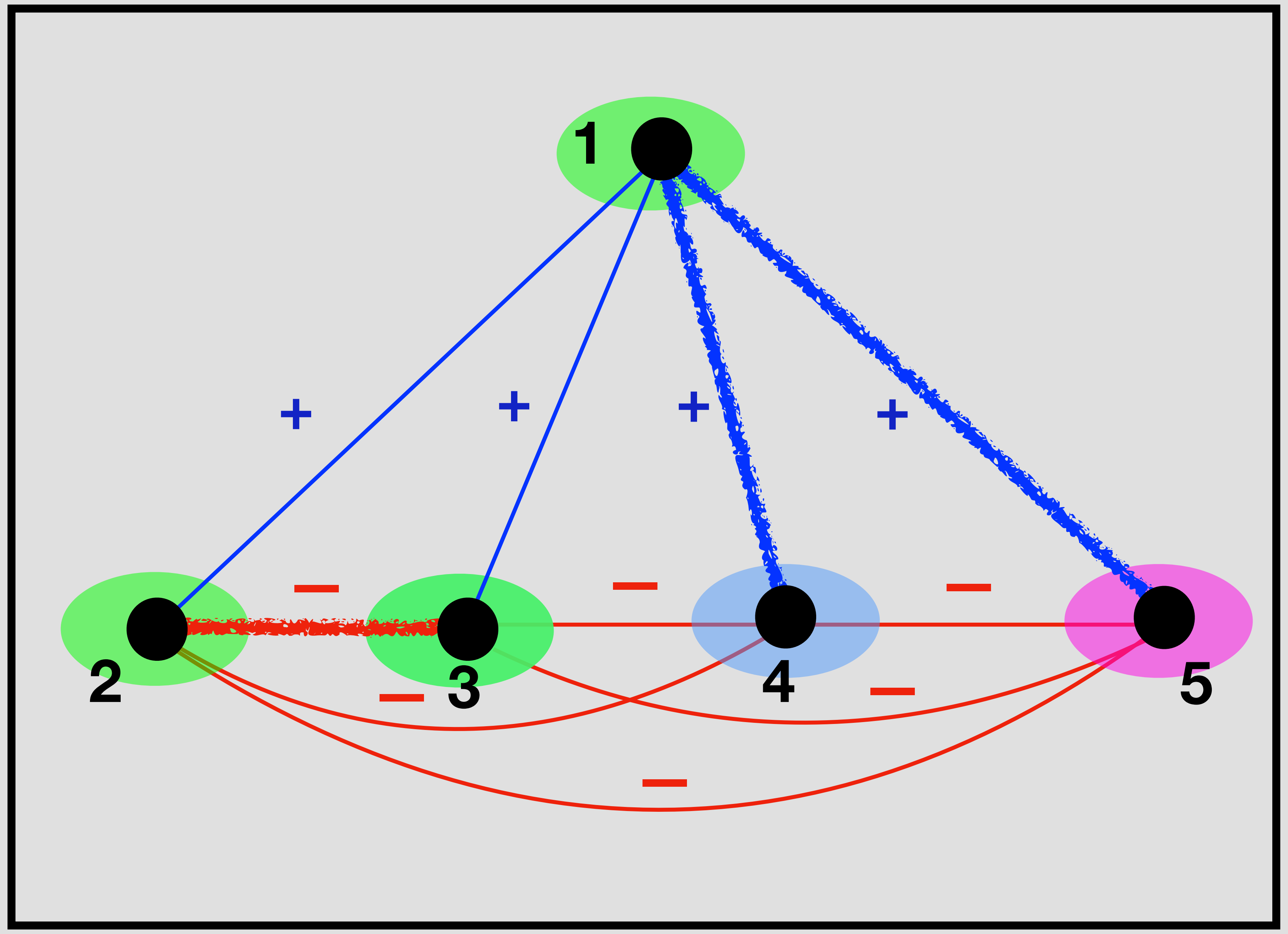


Correlation clustering

Original objective =
minimize # of edges in disagreement

Model:

- Cluster ***similar*** nodes together, separate ***dissimilar*** nodes
- Number / sizes of clusters (of vertices) **not** pre-fixed / pre-specified
- **Complete, unweighted graph**: each edge (+) or (–)
- Edge (u,v) in ***disagreement*** w.r.t C if
  - ✦ (+) with u, v different clusters or
  - ✦ (–) with u, v same cluster



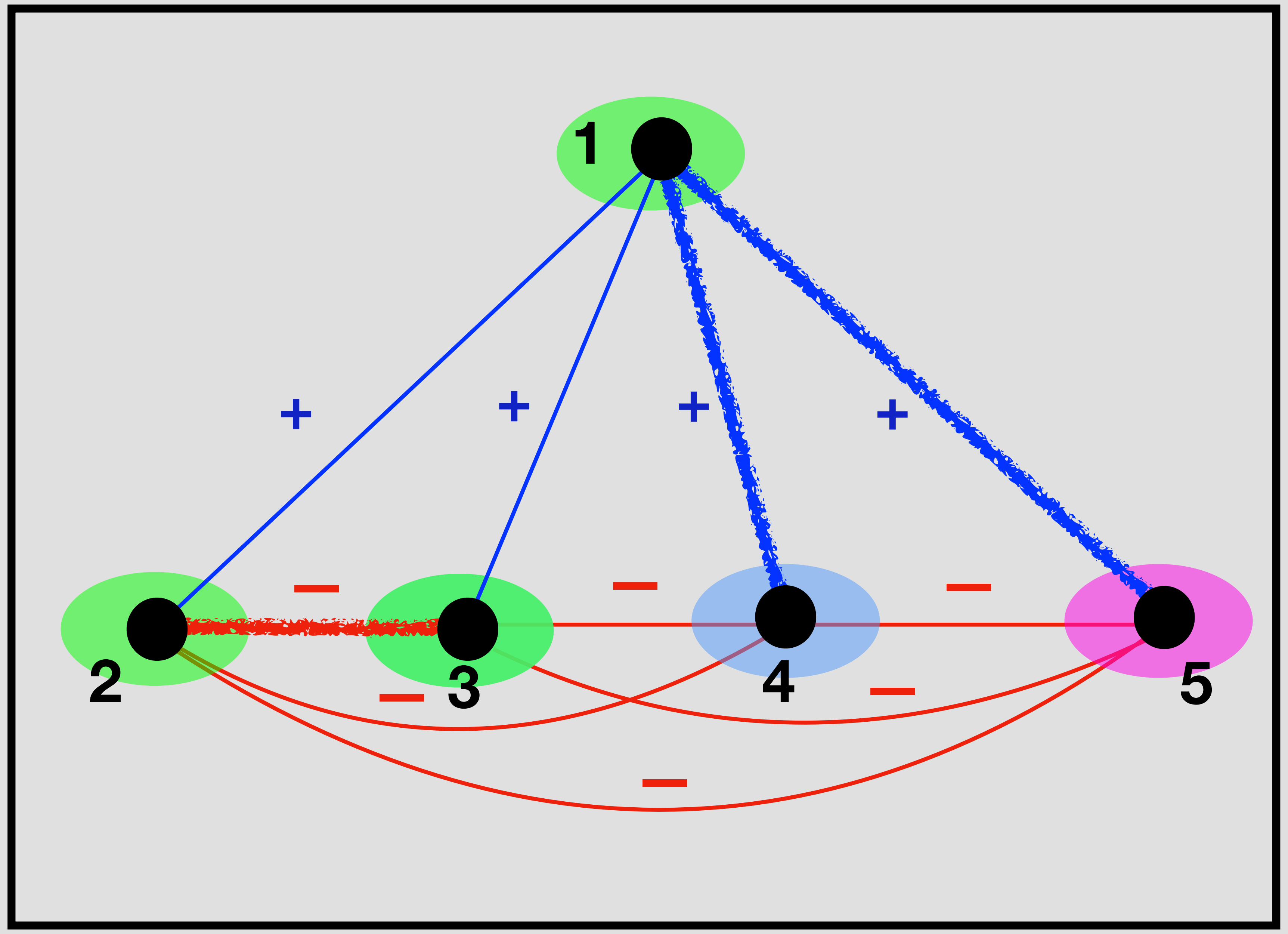
Correlation clustering

Original objective =
minimize # of edges in disagreement

Model:

- Cluster ***similar*** nodes together, separate ***dissimilar*** nodes
- Number / sizes of clusters (of vertices) **not** pre-fixed / pre-specified
- **Complete, unweighted graph**: each edge (+) or (–)
- Edge (u,v) in ***disagreement*** w.r.t C if
  - ✦ (+) with u, v different clusters or
  - ✦ (–) with u, v same cluster

$y_C(v)$ = # disagreements w.r.t. C incident to v



Correlation clustering

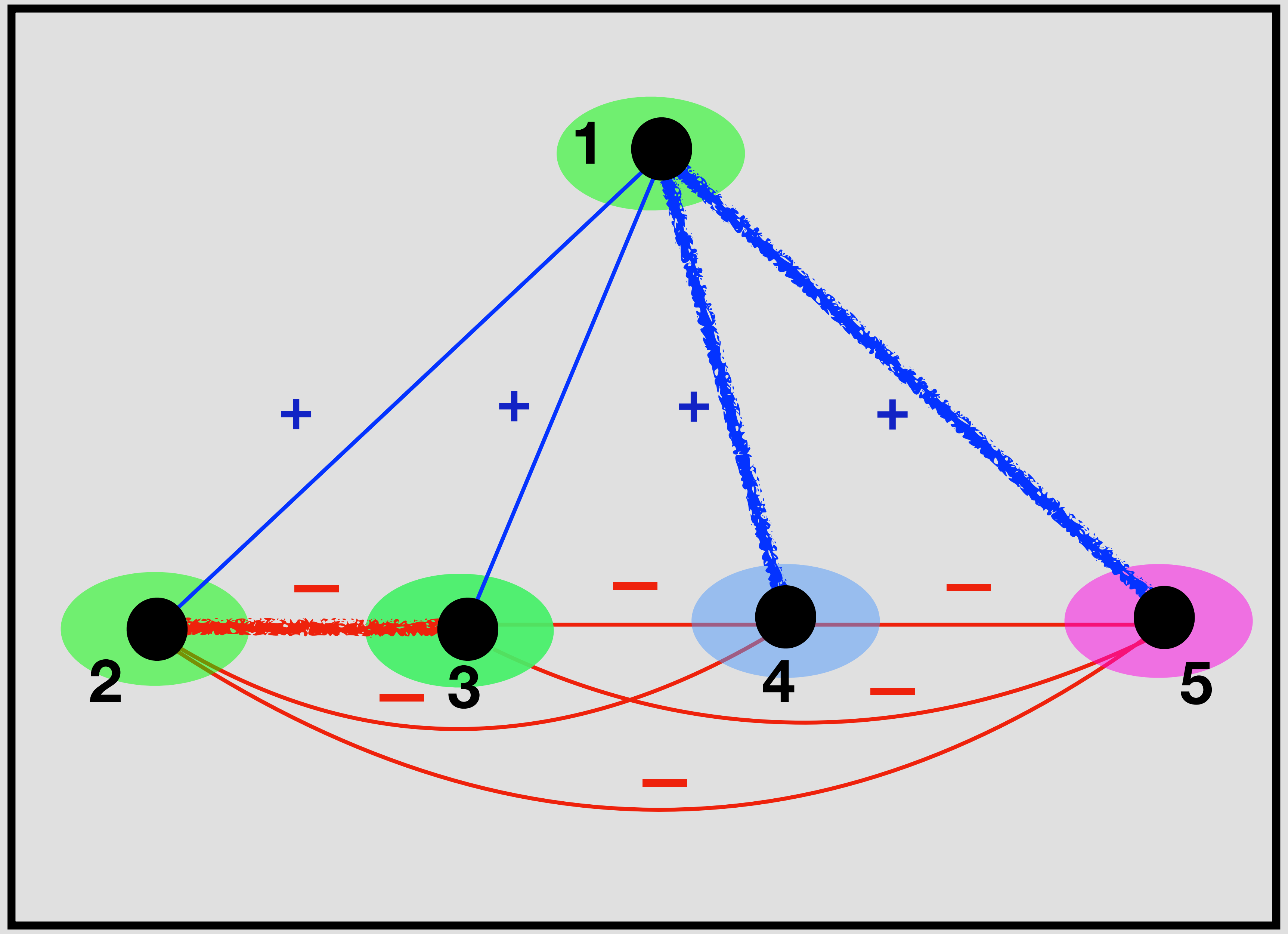
Original objective =
minimize # of edges in disagreement

Model:

- Cluster ***similar*** nodes together, separate ***dissimilar*** nodes
- Number / sizes of clusters (of vertices) **not** pre-fixed / pre-specified
- **Complete, unweighted graph**: each edge (+) or (–)
- Edge (u,v) in ***disagreement*** w.r.t C if
  - (+) with u, v different clusters or
  - (–) with u, v same cluster

$y_C(v)$ = # disagreements w.r.t. C incident to v

Goal: find $\mathbf{argmin}_C \sum_v y_C(v) = \|y_C\|_1$



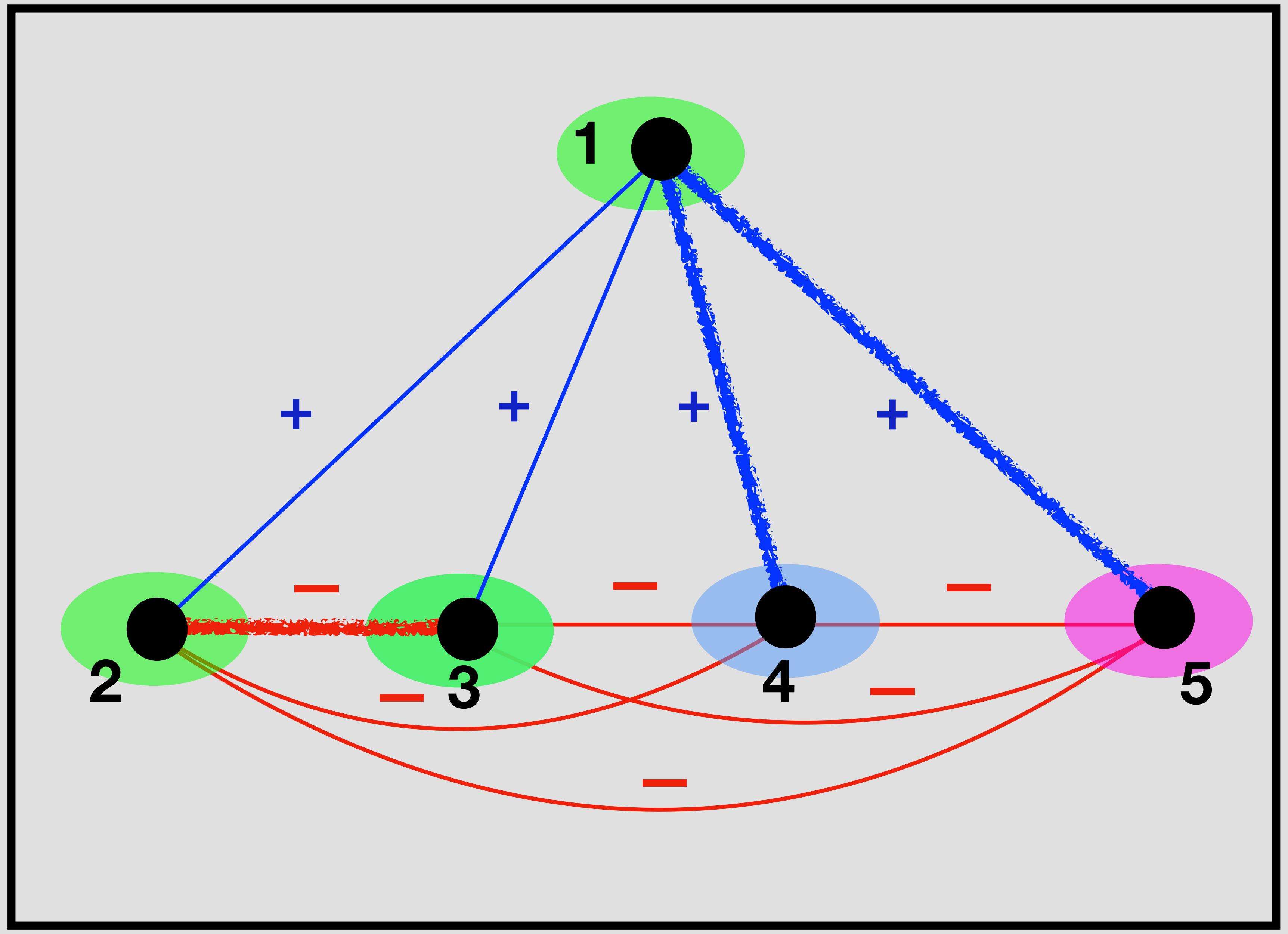
Correlation clustering

Model:

- Cluster ***similar*** nodes together, separate ***dissimilar*** nodes
- Number / sizes of clusters (of vertices) **not** pre-fixed / pre-specified
- **Complete, unweighted graph**: each edge (+) or (–)
- Edge (u,v) in ***disagreement*** w.r.t C if
  - (+) with u, v different clusters or
  - (–) with u, v same cluster

$y_C(v)$ = # disagreements w.r.t. C incident to v

Goal: find $\mathbf{argmin}_C \sum_v y_C(v) = \|y_C\|_1$



Correlation clustering

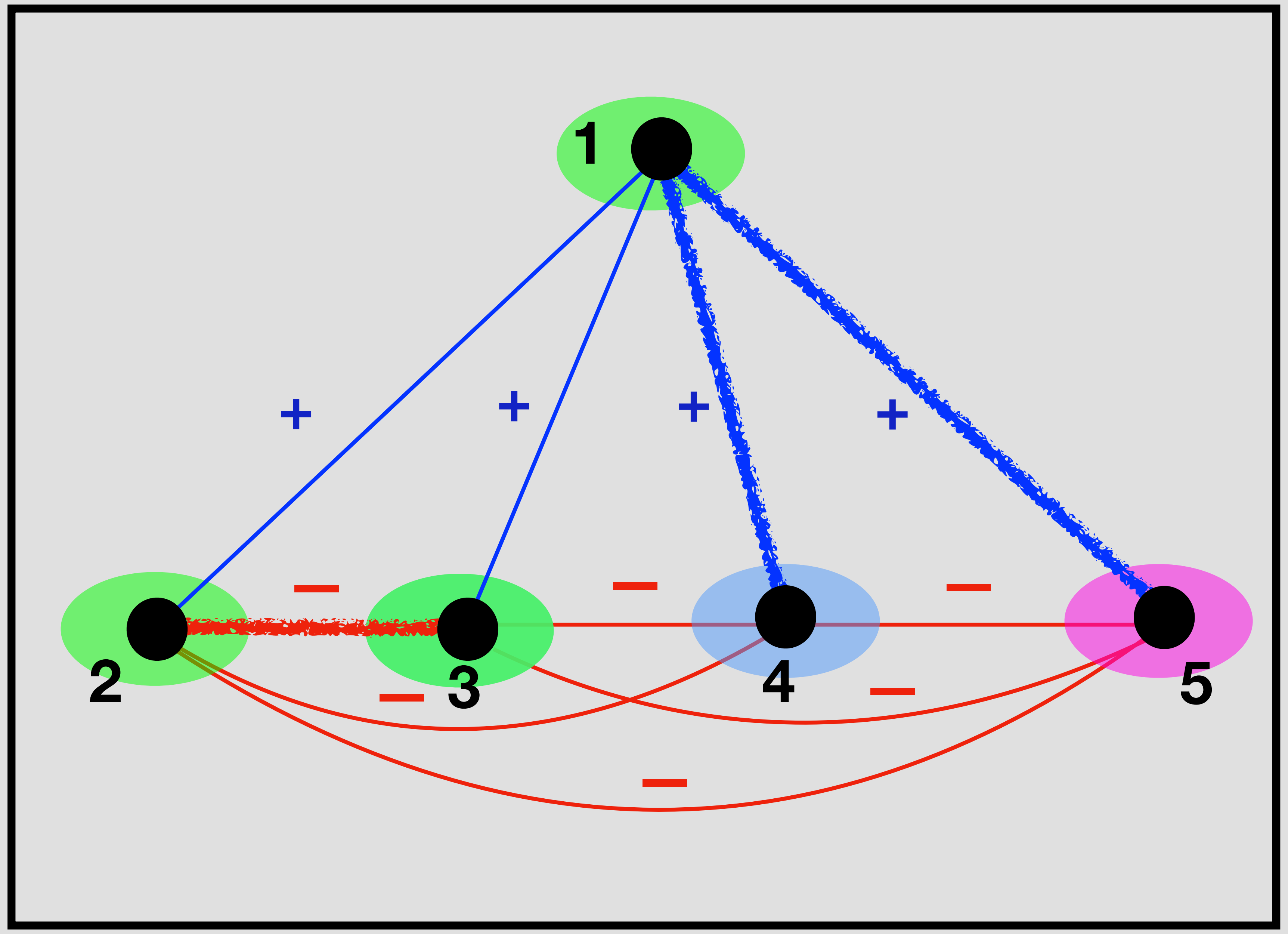
Model:

- Cluster ***similar*** nodes together, separate ***dissimilar*** nodes
- Number / sizes of clusters (of vertices) **not** pre-fixed / pre-specified
- **Complete, unweighted graph**: each edge (+) or (–)
- Edge (u,v) in ***disagreement*** w.r.t C if
  - (+) with u, v different clusters or
  - (–) with u, v same cluster

$y_C(v)$ = # disagreements w.r.t. C incident to v

Goal: find $\textbf{argmin}_C \sum_v y_C(v) = \|y_C\|_1$

$\textbf{argmin}_C \|y_C\|_p$



Correlation clustering

Model:

- Cluster ***similar*** nodes together, separate ***dissimilar*** nodes
- Number / sizes of clusters (of vertices) **not** pre-fixed / pre-specified
- **Complete, unweighted graph**: each edge (+) or (–)
- Edge (u,v) in ***disagreement*** w.r.t C if
  - (+) with u, v different clusters or
  - (–) with u, v same cluster

$y_C(v)$ = # disagreements w.r.t. C incident to v

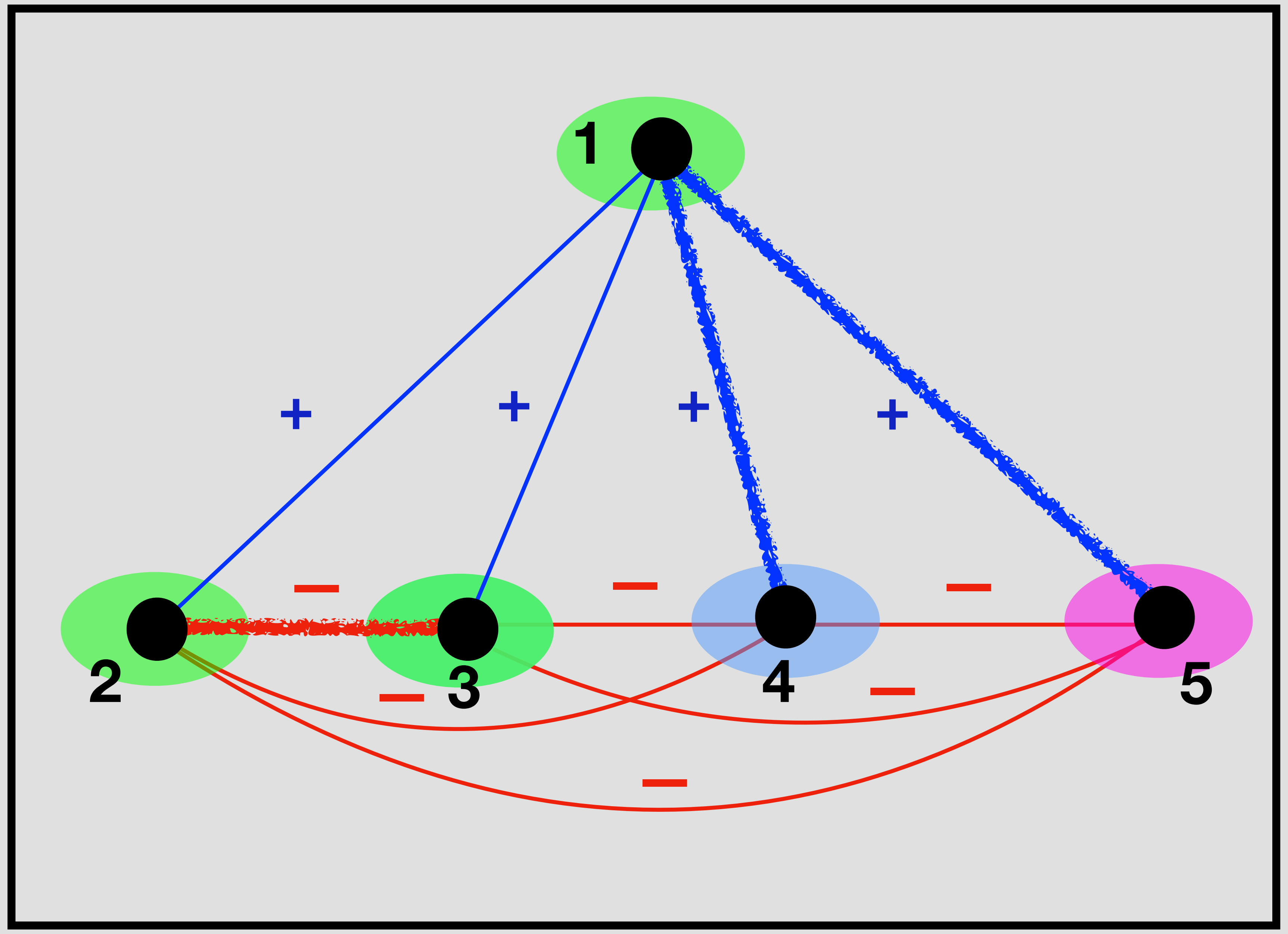
Goal: find ~~$\text{argmin}_C \sum_v y_C(v)$~~ $= \|y_C\|_1$

$p \geq 1$

$\text{argmin}_C \|y_C\|_p$

ℓ_1 = original obj

ℓ_∞ = min max norm



Correlation clustering

Model:

- Cluster ***similar*** nodes together, separate ***dissimilar*** nodes
- Number / sizes of clusters (of vertices) **not** pre-fixed / pre-specified
- **Complete, unweighted graph**: each edge (+) or (–)

p small = global obj ↔ p large = local/fair obj

- (+) with u, v different clusters or
- (–) with u, v same cluster

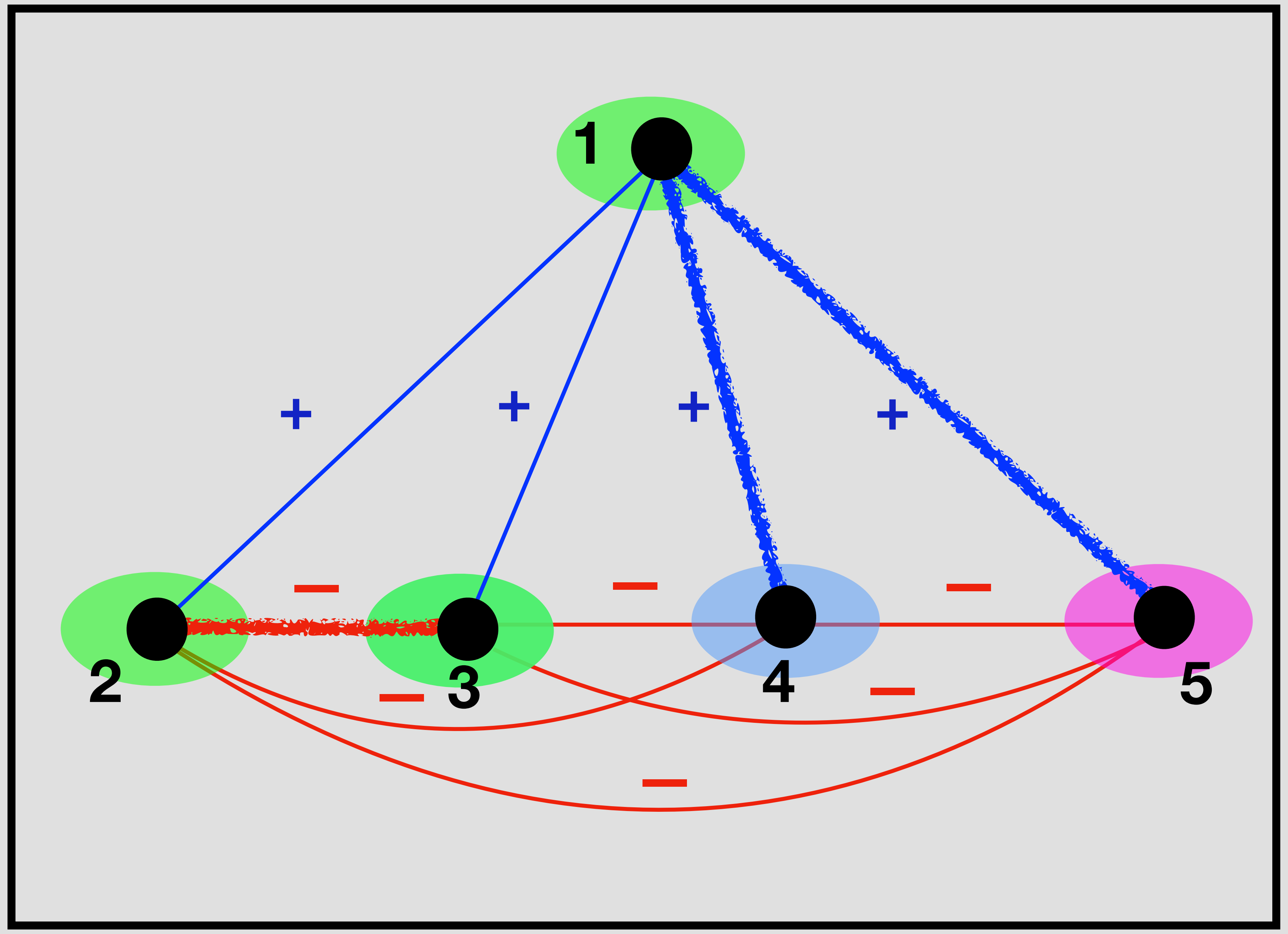
$y_C(v)$ = # disagreements w.r.t. C incident to v

Goal: find $\text{argmin}_C \sum_v y_C(v) = \|y_C\|_1$

$p \geq 1$

$\text{argmin}_C \|y_C\|_p$

ℓ_1 = original obj
ℓ_∞ = min max norm



Previous work

Previous work

For ℓ_1-norm (original) objective:

- Introduced by [Bansal, Blum, Chawla '04]

Previous work

For ℓ_1-norm (original) objective:

- Introduced by [Bansal, Blum, Chawla '04]
- Linear time Pivot algorithm gives 3-apx
  [Ailon, Charikar, Newman JACM08] [Chierichetti, Dalvi, Kumar KDD14]
- APX-hard
  [Charikar, Guruswami, Wirth JCSS05]

Previous work

For ℓ_1-norm (original) objective:

- Introduced by [Bansal, Blum, Chawla '04]
- Linear time Pivot algorithm gives 3-apx
  [Ailon, Charikar, Newman JACM08] [Chierichetti, Dalvi, Kumar KDD14]
- APX-hard
  [Charikar, Guruswami, Wirth JCSS05]
- Many other active threads of research!
  [Ahmadi, Khuller, Saha IPCO19] [Veldt ICML22] [Cohen-Addad, Lee, Li, Newman FOCS23]

Previous work

For ℓ_1-norm (original) objective:

- Introduced by [Bansal, Blum, Chawla '04]
- Linear time Pivot algorithm gives 3-apx
  [Ailon, Charikar, Newman JACM08] [Chierichetti, Dalvi, Kumar KDD14]
- APX-hard
  [Charikar, Guruswami, Wirth JCSS05]
- Many other active threads of research!
  [Ahmadi, Khuller, Saha IPCO19] [Veldt ICML22] [Cohen-Addad, Lee, Li, Newman FOCS23]

For general ℓ_p-norm objectives:

- 5-approximation algorithm; NP-hard (even for $p = \infty$!)
  [Puleo, Milenkovic ICML16], [Charikar, Gupta, Schwartz IPCO17], [Kalhan, Makarychev, Zhou ICML19]
- All previous techniques round solution to a convex program

Trouble with using convex program

Trouble with using convex program

ℓ_p-norm correlation clustering algs solve a convex program

Trouble with using convex program

ℓ_p-norm correlation clustering algs solve a convex program

Solving ***metric constrained*** programs on large networks is slow!

Work on solving CC programs fast only scales to graphs with few thousand vertices!

[Ruggles et al. '20], [Sonthalia & Gilbert '20], [Veldt '22]

Trouble with using convex program

ℓ_p-norm correlation clustering algs solve a convex program

Solving ***metric constrained*** programs on large networks is slow!

Not very amenable to online, streaming, etc.

Work on solving CC programs fast only scales to graphs with few thousand vertices!

[Ruggles et al. '20], [Sonthalia & Gilbert '20], [Veldt '22]

Trouble with using convex program

ℓ_p-norm correlation clustering algs solve a convex program

Solving ***metric constrained*** programs on large networks is slow!

Not very amenable to online, streaming, etc.

Solution specific to ***one fixed*** ℓ_p-norm

Work on solving CC programs fast only scales to graphs with few thousand vertices!

[Ruggles et al. '20], [Sonthalia & Gilbert '20], [Veldt '22]

Trouble with using convex program

ℓ_p-norm correlation clustering algs solve a convex program

Solving ***metric constrained*** programs on large networks is slow!

Work on solving CC programs fast only scales to graphs with few thousand vertices!

[Ruggles et al. '20], [Sonthalia & Gilbert '20], [Veldt '22]

Not very amenable to online, streaming, etc.

Solution specific to ***one fixed*** ℓ_p-norm

Trouble with using convex program

ℓ_p-norm correlation clustering algs solve a convex program

Solving ***metric constrained*** programs on large networks is slow!

Work on solving CC programs fast only scales to graphs with few thousand vertices!

[Ruggles et al. '20], [Sonthalia & Gilbert '20], [Veldt '22]

Not very amenable to online, streaming, etc.

Solution specific to ***one fixed*** ℓ_p-norm

All-norms objective

Trouble with using convex program

ℓ_p-norm correlation clustering algs solve a convex program

Solving ***metric constrained*** programs on large networks is slow!

Work on solving CC programs fast only scales to graphs with few thousand vertices!

[Ruggles et al. '20], [Sonthalia & Gilbert '20], [Veldt '22]

Not very amenable to online, streaming, etc.

Solution specific to ***one fixed*** ℓ_p-norm

All-norms objective

✦**Seek:** *single* clustering that well-approximates all ℓ_p-norms *simultaneously*

Trouble with using convex program

ℓ_p-norm correlation clustering algs solve a convex program

Solving ***metric constrained*** programs on large networks is slow!

Not very amenable to online, streaming, etc.

Solution specific to ***one fixed*** ℓ_p-norm

Work on solving CC programs fast only scales to graphs with few thousand vertices!

[Ruggles et al. '20], [Sonthalia & Gilbert '20], [Veldt '22]

All-norms objective

- **Seek:** *single* clustering that well-approximates all ℓ_p-norms *simultaneously*
- Introduced by [Azar, Epstein, Richter, Woeginger '04] for load balancing

Trouble with using convex program

ℓ_p-norm correlation clustering algs solve a convex program

Solving ***metric constrained*** programs on large networks is slow!

Work on solving CC programs fast only scales to graphs with few thousand vertices!

[Ruggles et al. '20], [Sonthalia & Gilbert '20], [Veldt '22]

Not very amenable to online, streaming, etc.

Solution specific to ***one fixed*** ℓ_p-norm

All-norms objective

- ✦**Seek:** *single* clustering that well-approximates all ℓ_p-norms *simultaneously*
- ✦Introduced by [Azar, Epstein, Richter, Woeginger '04] for load balancing
- ✦ℓ_p set cover, flow time in scheduling, and more

Can we do better?

ℓ_p-norm correlation clustering algs solve a convex program

Solving ***metric constrained*** LPs on large networks is slow!

Not very amenable to online, streaming, etc.

Solution specific to ***one fixed*** ℓ_p-norm

All-norms objective

- **Seek:** *single* clustering that well-approximates all ℓ_p-norms

Can we do better?

ℓ_p-norm correlation clustering algs solve a convex program

Solving ***metric constrained*** LPs on large networks is slow!

Not very amenable to online, streaming, etc.

Solution specific to ***one fixed*** ℓ_p-norm

All-norms objective

✦ **Seek:** *single* clustering that well-approximates all ℓ_p-norms

Today:

Can we do better?

ℓ_p-norm correlation clustering algs solve a convex program

Solving ***metric constrained*** LPs on large networks is slow!

Not very amenable to online, streaming, etc.

Solution specific to ***one fixed*** ℓ_p-norm

All-norms objective

✦ **Seek:** *single* clustering that well-approximates all ℓ_p-norms

Today:

- *Does there exist an **all-norms** solution for CC?*

Can we do better?

ℓ_p-norm correlation clustering algs solve a convex program

Solving ***metric constrained*** LPs on large networks is slow!

Not very amenable to online, streaming, etc.

Solution specific to ***one fixed*** ℓ_p-norm

All-norms objective

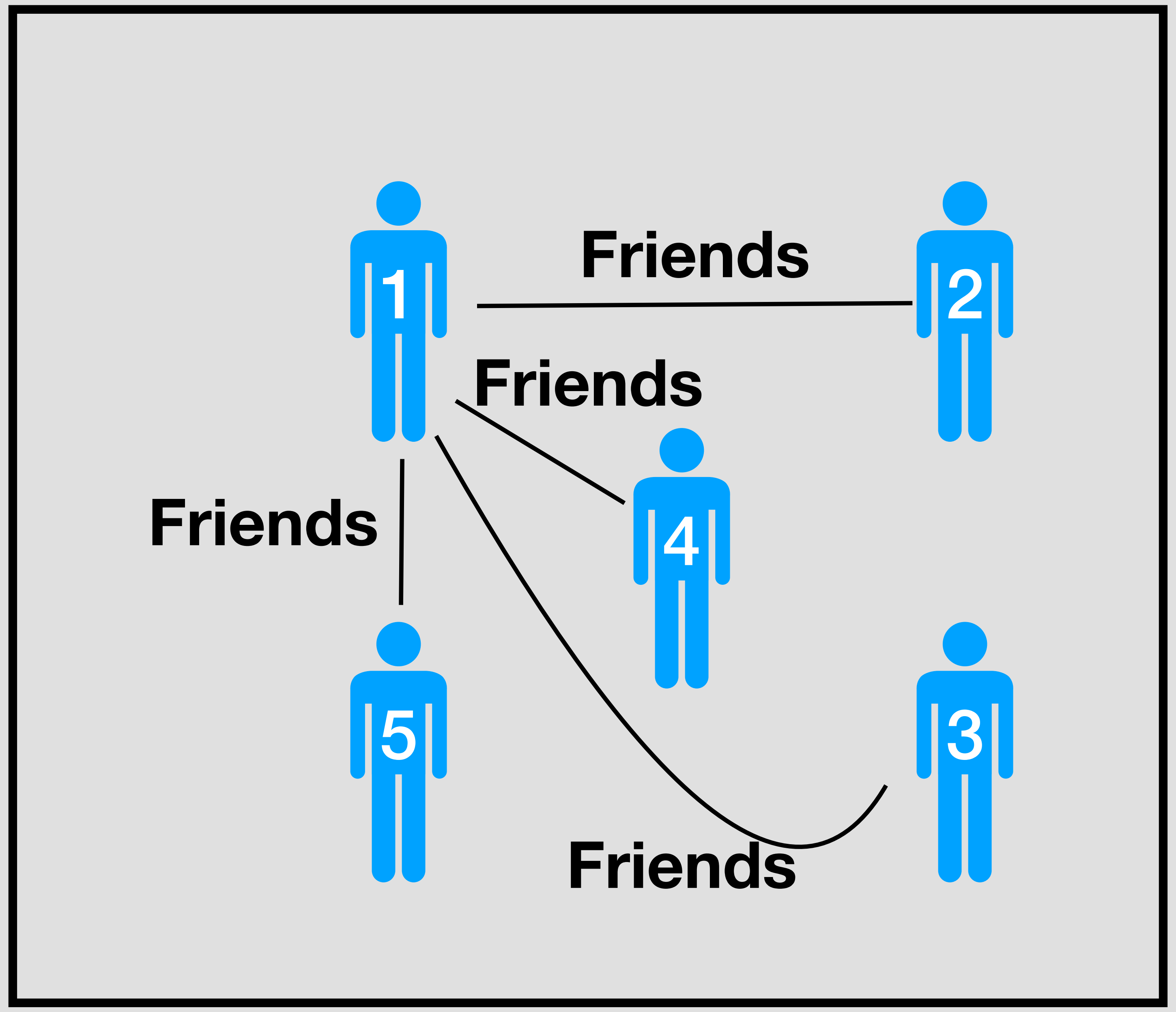
- **Seek:** *single* clustering that well-approximates all ℓ_p-norms

Today:

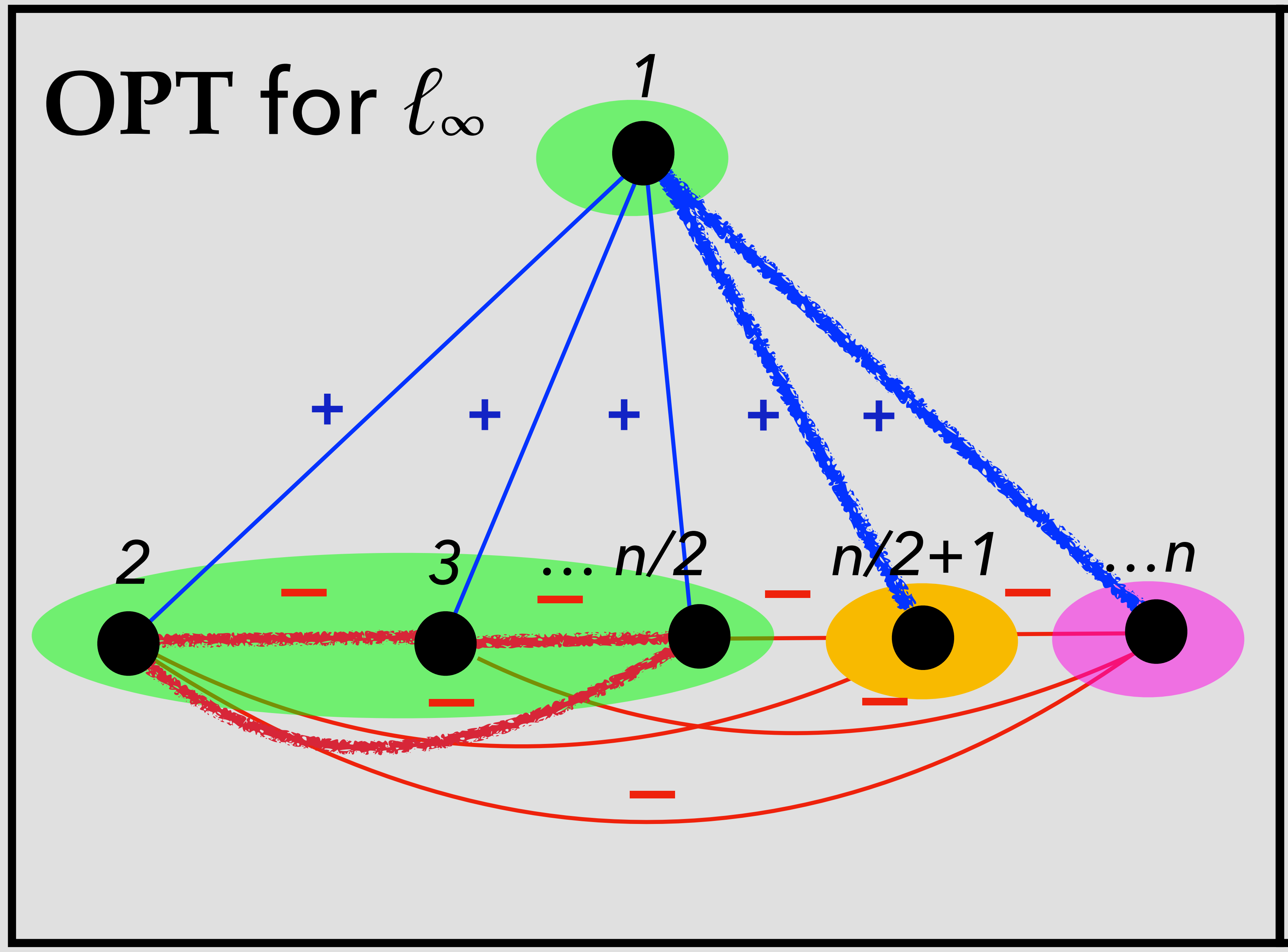
- *Does there exist an* ***all-norms*** *solution for CC?*
- *Can it be found through a* ***fast, combinatorial*** *algorithm?*

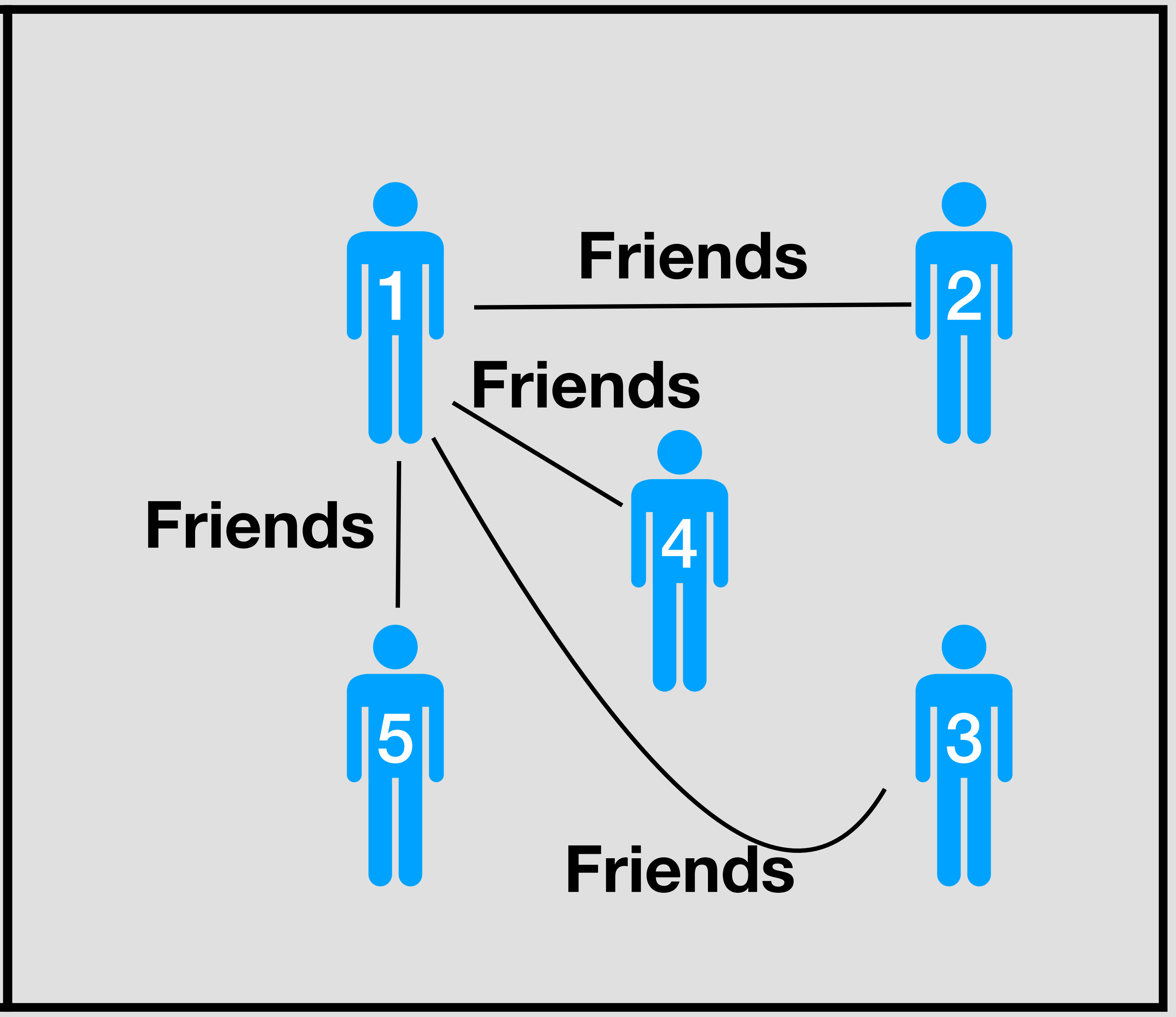
Observe: naively optimizing one ℓ_p-norm can be very sub-optimal for others

Observe: naively optimizing one ℓ_p-norm can be very sub-optimal for others

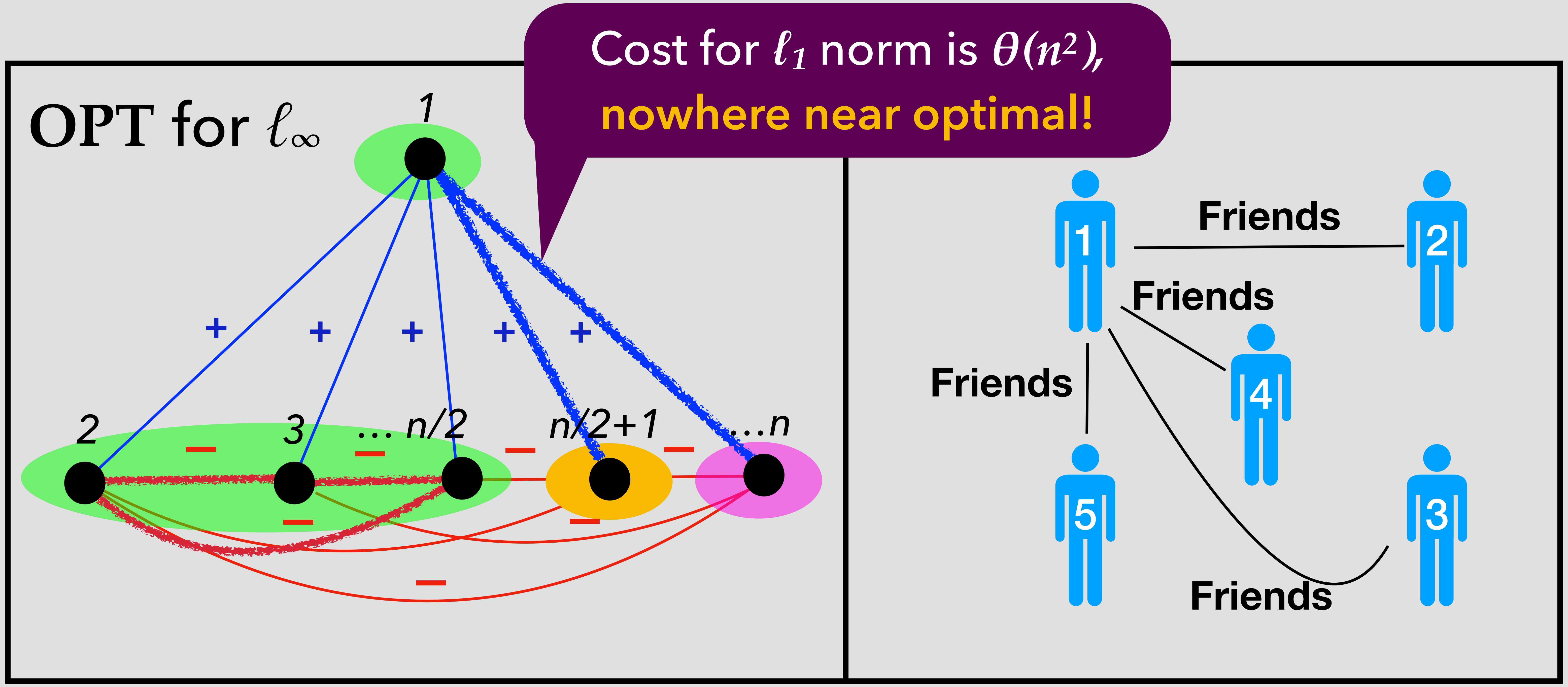


Observe: naively optimizing one ℓ_p-norm can be very sub-optimal for others

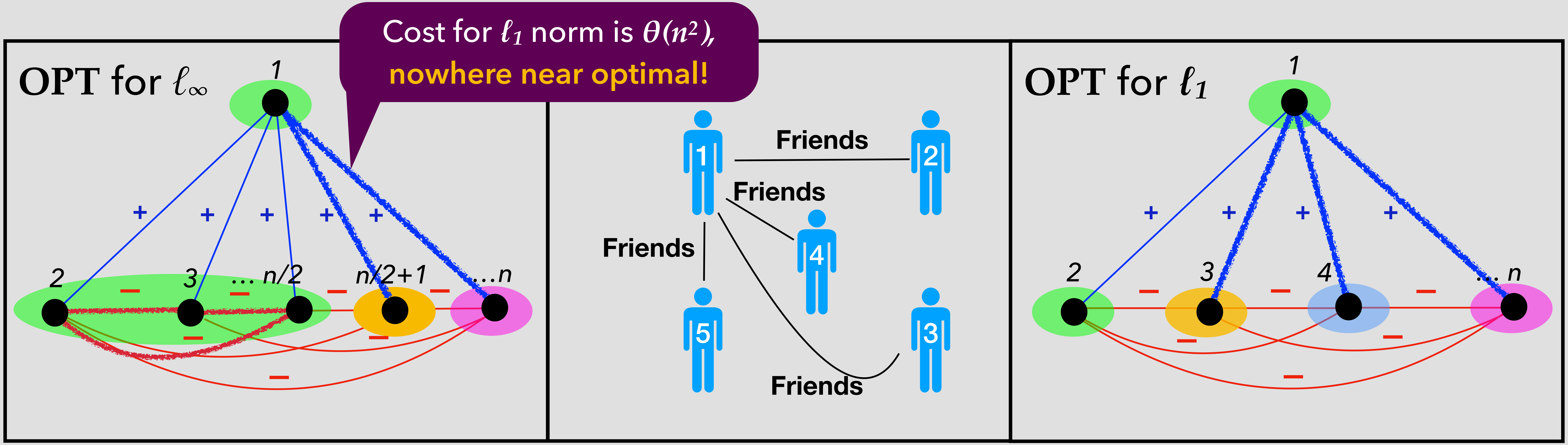




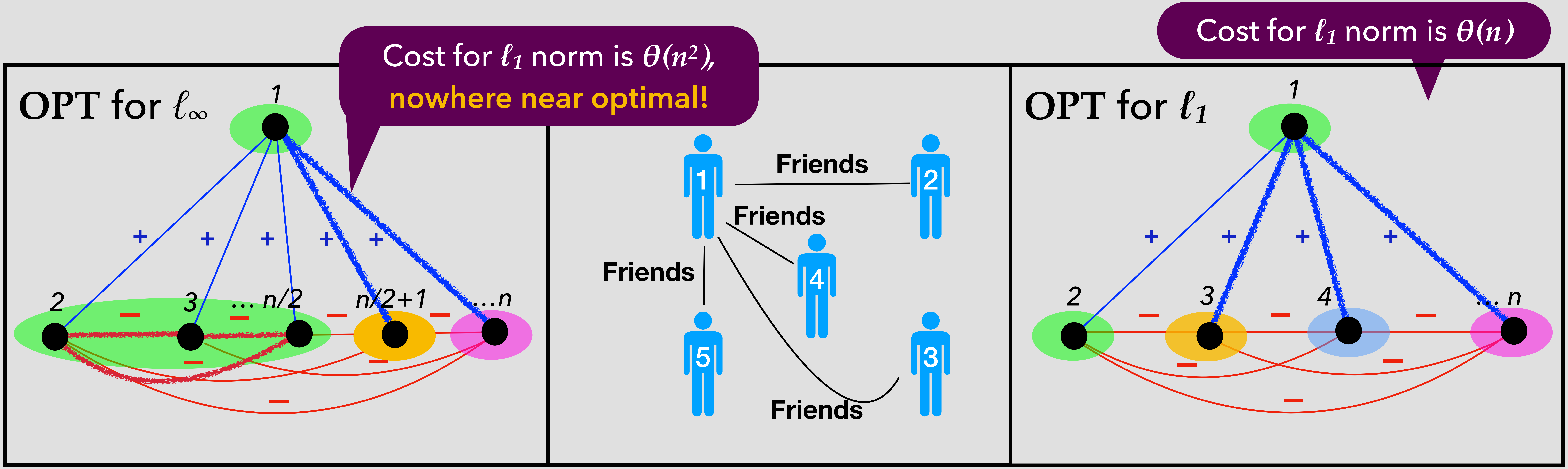
Observe: naively optimizing one ℓ_p-norm can be very sub-optimal for others



Observe: naively optimizing one ℓ_p-norm can be very sub-optimal for others



Observe: naively optimizing one ℓ_p-norm can be very sub-optimal for others

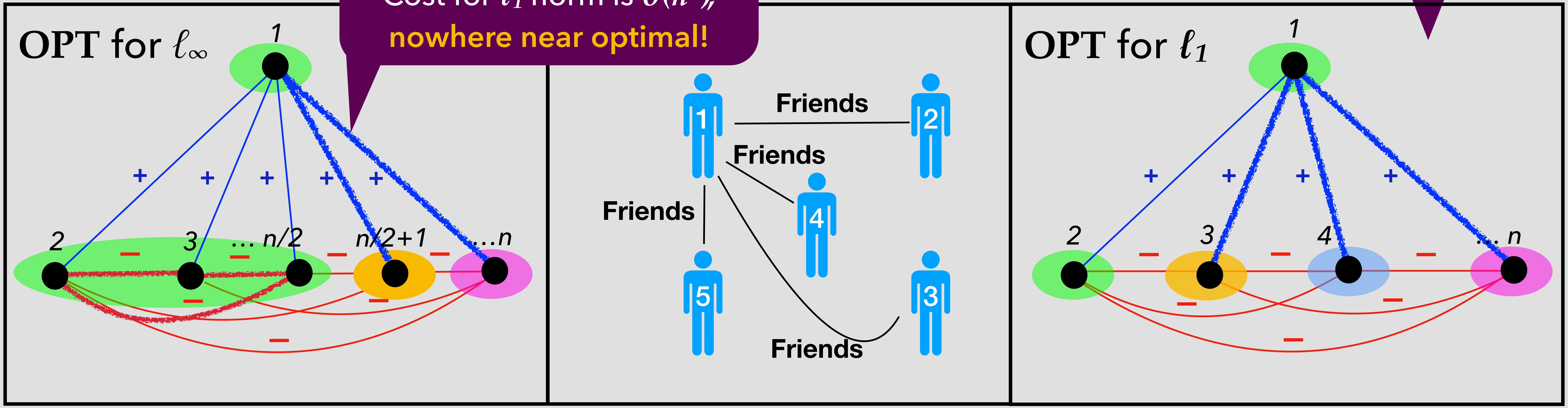


Observe: naively optimizing one ℓ_p-norm can be very sub-optimal for others

Does there exist a "pretty good" (say, $O(1)$-apx) solution for all ℓ_p-norms?

Cost for ℓ_1 norm is $\theta(n^2)$, **nowhere near optimal!**

Cost for ℓ_1 norm is $\theta(n)$



Today:

- *Does there exist an* ***all-norms*** *solution for CC?*
- *Can it be found through a* ***fast, combinatorial*** *algorithm?*

Today:

- *Does there exist an* ***all-norms*** *solution for CC?*
- *Can it be found through a* ***fast, combinatorial*** *algorithm?*

Yes!

Today:

- *Does there exist an* ***all-norms*** *solution for CC?*
- *Can it be found through a* ***fast, combinatorial*** *algorithm?*

Yes!

(1) $O(1)$-apx for min-max CC ($p = \infty$)

Today:

- *Does there exist an **all-norms** solution for CC?*
- *Can it be found through a **fast, combinatorial** algorithm?*

Yes!

(1) $O(1)$-apx for min-max CC ($p = \infty$)

↪ completely combinatorial (first for $p = \infty$)

Today:

- *Does there exist an **all-norms** solution for CC?*
- *Can it be found through a **fast, combinatorial** algorithm?*

Yes!

(1) $O(1)$-apx for min-max CC ($p = \infty$)

↪ completely combinatorial (first for $p = \infty$)

↪ $O(n^\omega)$ time, near-linear for small max (+) degree

Today:

- *Does there exist an **all-norms** solution for CC?*
- *Can it be found through a **fast, combinatorial** algorithm?*

Yes!

(1) $O(1)$-apx for min-max CC ($p = \infty$)

↪ completely combinatorial (first for $p = \infty$)

↪ $O(n^\omega)$ time, near-linear for small max (+) degree

↓ "tweak"

Today:

- *Does there exist an* ***all-norms*** *solution for CC?*
- *Can it be found through a* ***fast, combinatorial*** *algorithm?*

Yes!

(1) $O(1)$-apx for min-max CC ($p = \infty$)

↪ completely combinatorial (first for $p = \infty$)

↪ $O(n^{\omega})$ time, near-linear for small max (+) degree

↓ "tweak"

Today:

- *Does there exist an* ***<u>all-norms</u>*** *solution for CC?*
- *Can it be found through a* ***<u>fast, combinatorial</u>*** *algorithm?*

Yes!

(1) $O(1)$-apx for min-max CC ($p = \infty$)

↪ completely combinatorial (first for $p = \infty$)

↪ $O(n^{\omega})$ time, near-linear for small max (+) degree

↓ "tweak"

(2) $O(1)$-apx for **<u>all</u>** $p \in [1,\infty]$, i.e., **<u>all-norms solution</u>**

Today:

- *Does there exist an **all-norms** solution for CC?*
- *Can it be found through a **fast, combinatorial** algorithm?*

Yes!

(1) $O(1)$-apx for min-max CC ($p = \infty$)

↪ completely combinatorial (first for $p = \infty$)

↪ $O(n^{\omega})$ time, near-linear for small max (+) degree

↓ "tweak"

(2) $O(1)$-apx for **all** $p \in [1,\infty]$, i.e., **all-norms solution**

↪ completely combinatorial (first for $p > 1$)

Today:

- *Does there exist an **all-norms** solution for CC?*
- *Can it be found through a **fast, combinatorial** algorithm?*

Yes!

(1) $O(1)$-apx for min-max CC ($p = \infty$)

↪ completely combinatorial (first for $p = \infty$)

↪ $O(n^\omega)$ time, near-linear for small max (+) degree

↓ "tweak"

(2) $O(1)$-apx for **all** $p \in [1,\infty]$, i.e., **all-norms solution**

↪ completely combinatorial (first for $p > 1$)

↪ $O(n^\omega)$ time, near-linear for small max (+) degree

Today:

- *Does there exist an* ***all-norms*** *solution for CC?*
- *Can it be found through a* ***fast, combinatorial*** *algorithm?*

Yes!

(1) $O(1)$-apx for min-max CC ($p = \infty$)

↪ completely combinatorial (first for $p = \infty$)

↪ $O(n^{\omega})$ time, near-linear for small max (+) degree

↓ "tweak"

(2) $O(1)$-apx for **all** $p \in [1,\infty]$, i.e., **all-norms solution**

↪ completely combinatorial (first for $p > 1$)

↪ $O(n^{\omega})$ time, near-linear for small max (+) degree

Not possible for k-center & k-median [Alamdari & Shmoys WAOA17]

Today:

- *Does there exist an **all-norms** solution for CC?*
- *Can it be found through a **fast, combinatorial** algorithm?*

Yes!

"Fast Combinatorial Algorithms for Min Max Correlation Clustering" *(ICML 23)*

(1) $O(1)$-apx for min-max CC ($p = \infty$)

↪ completely combinatorial (first for $p = \infty$)

↪ $O(n^{\omega})$ time, near-linear for small max (+) degree

↓ "tweak"

(2) $O(1)$-apx for **all** $p \in [1,\infty]$, i.e., **all-norms solution**

↪ completely combinatorial (first for $p > 1$)

↪ $O(n^{\omega})$ time, near-linear for small max (+) degree

Not possible for k-center & k-median [Alamdari & Shmoys WAOA17]

Previous techniques

Previous techniques

Convex program relaxation

Can be solved "efficiently"

Previous techniques

Convex program relaxation

Can be solved "efficiently"

$$\min \|y\|_p$$

$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$

Previous techniques

$x_{uv} = 0$ then u, v same cluster
$x_{uv} = 1$ then u, v different clusters

Convex program relaxation

Can be solved "efficiently"

$$\min \|y\|_p$$

$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$

Previous techniques

Convex program relaxation

Can be solved "efficiently"

$x_{uv} = 0$ then u, v same cluster

$x_{uv} = 1$ then u, v different clusters

$$\min \|y\|_p$$

$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$

triangle inequality only constraints

↪ feasible solutions = (semi-)metrics

Previous techniques

Convex program relaxation

Can be solved "efficiently"

$x_{uv} = 0$ then u, v same cluster
$x_{uv} = 1$ then u, v different clusters

$$\min \|y\|_p$$

$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$

triangle inequality only constraints
↪ feasible solutions = (semi-)metrics

"probability" u, v separated

Previous techniques

Convex program relaxation

Can be solved "efficiently"

$x_{uv} = 0$ then u, v same cluster
$x_{uv} = 1$ then u, v different clusters

$$\min \|y\|_p$$

$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$

triangle inequality only constraints
↪ feasible solutions = (semi-)metrics

"probability" u, v separated

Previous techniques

Convex program relaxation

Can be solved "efficiently"

$x_{uv} = 0$ then u, v same cluster

$x_{uv} = 1$ then u, v different clusters

$$\min \|y\|_p$$

$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$

triangle inequality only constraints

↪ feasible solutions = (semi-)metrics

"probability" u, v separated

Previous techniques

Past approaches

Step 1: Solve convex program

Step 2: "Round" fractional solution to integral one

Convex program relaxation

Can be solved "efficiently"

$x_{uv} = 0$ then u, v same cluster

$x_{uv} = 1$ then u, v different clusters

$$\min \|y\|_p$$

$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$

triangle inequality only constraints

↪ feasible solutions = (semi-)metrics

"probability" u, v separated

Previous techniques

Convex program relaxation

Can be solved "efficiently"

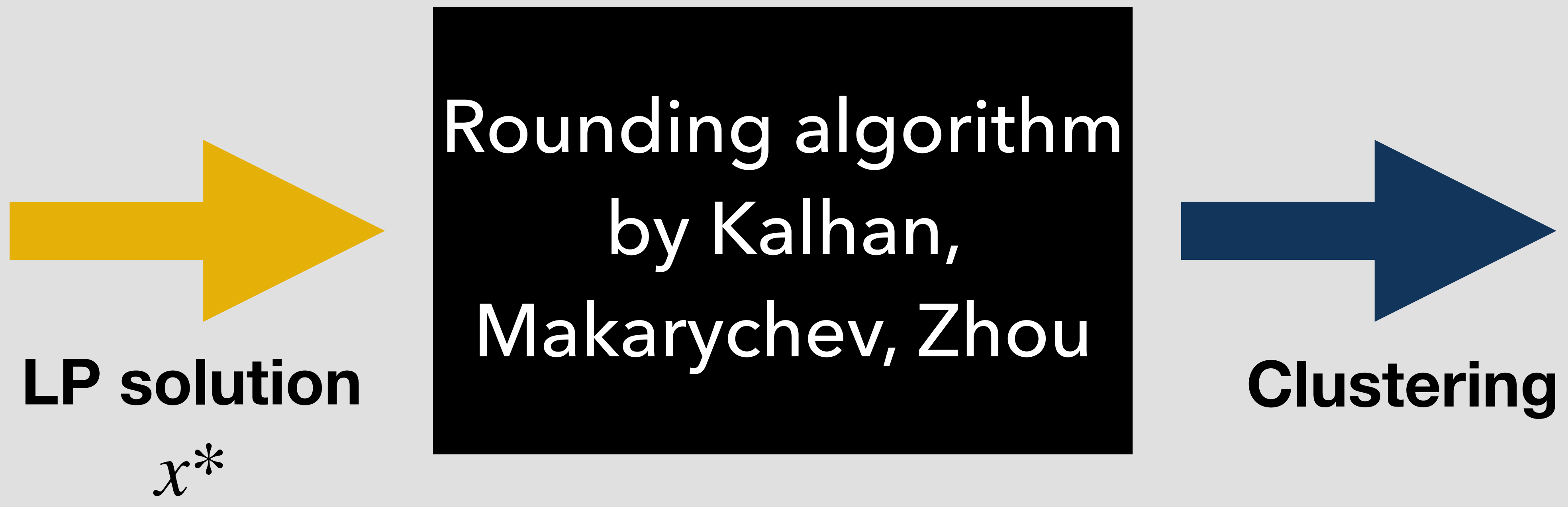
$$\min \|y\|_p$$

$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$

Previous techniques



Convex program relaxation

Can be solved "efficiently"

$$\min \|y\|_p$$

$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$

Today

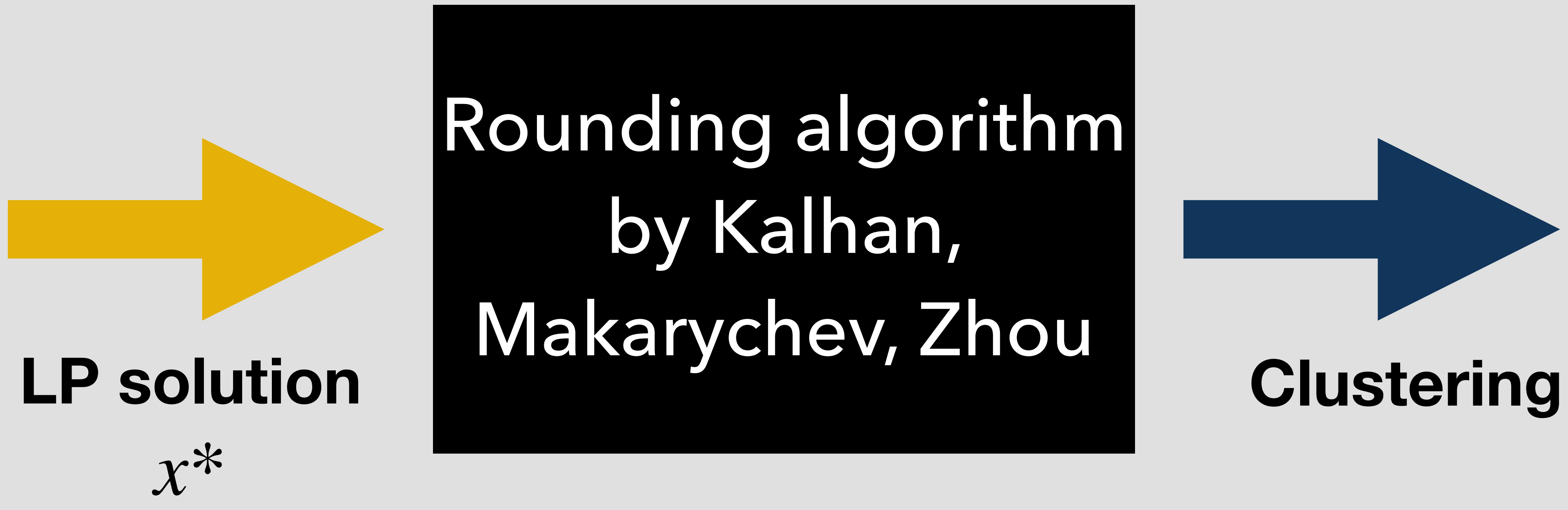
- ✦The **correlation metric** **(constructing a "guess" for the optimal solution to convex relaxation)**
- ✦Tweaking correlation metric for **all** ℓ_p-norms
- ✦Open questions

Today

based solely on combinatorial properties

- ✦The **correlation metric** (constructing a "guess" for the optimal solution to convex relaxation)
- ✦Tweaking correlation metric for **all** ℓ_p-norms
- ✦Open questions

Our technique



Convex program relaxation

$$\min \|y\|_p$$

$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \le x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \le x_{uv} \le 1 \qquad \forall u, v \in V$$

Our technique

Convex ~~program~~ relaxation

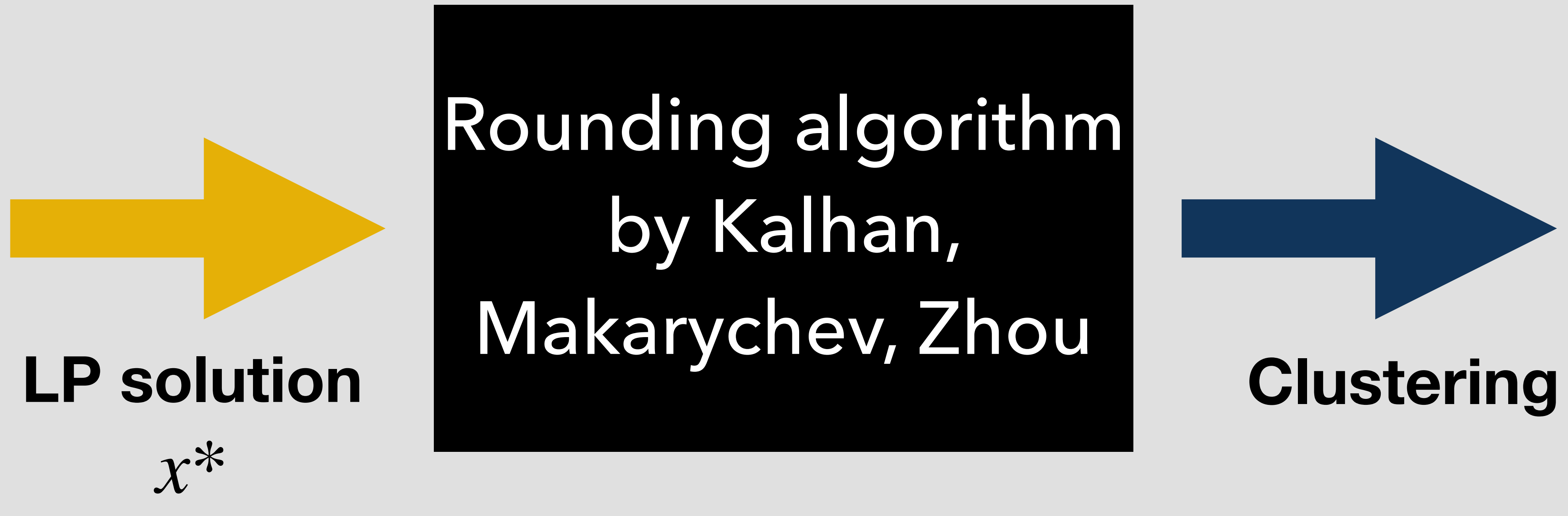
Don't solve this!

$$\min \|y\|_p$$

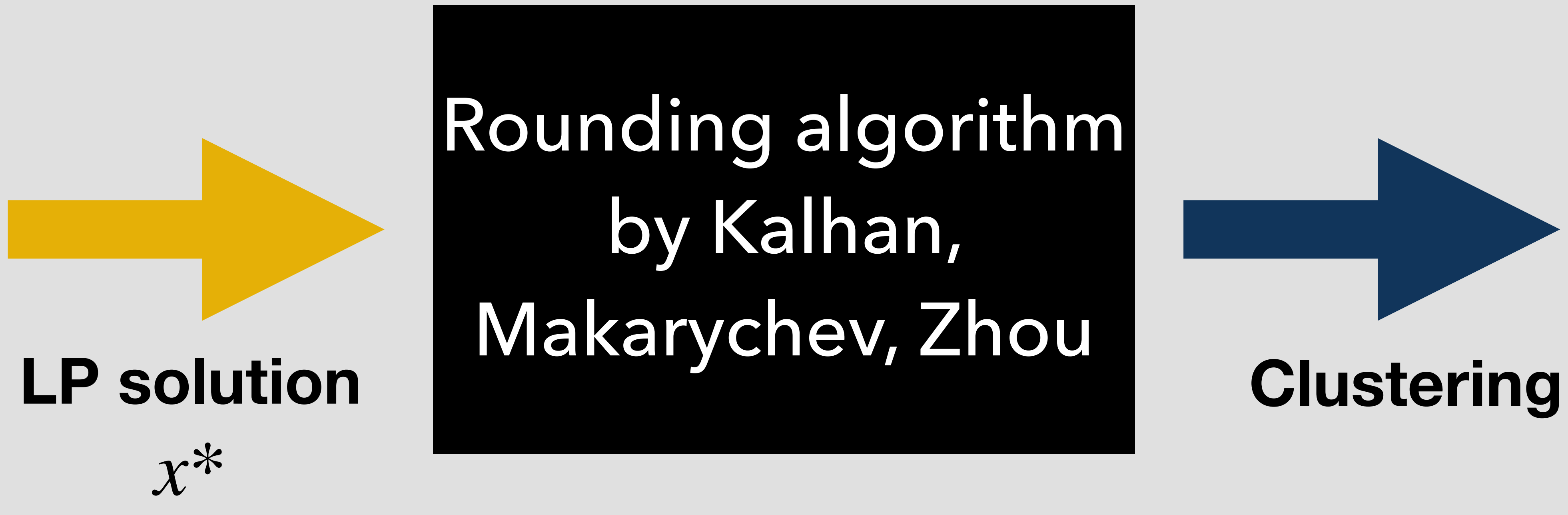
$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$



Our technique



Convex program relaxation

Don't solve this!

Quantify probabilities x_{uv} combinatorially

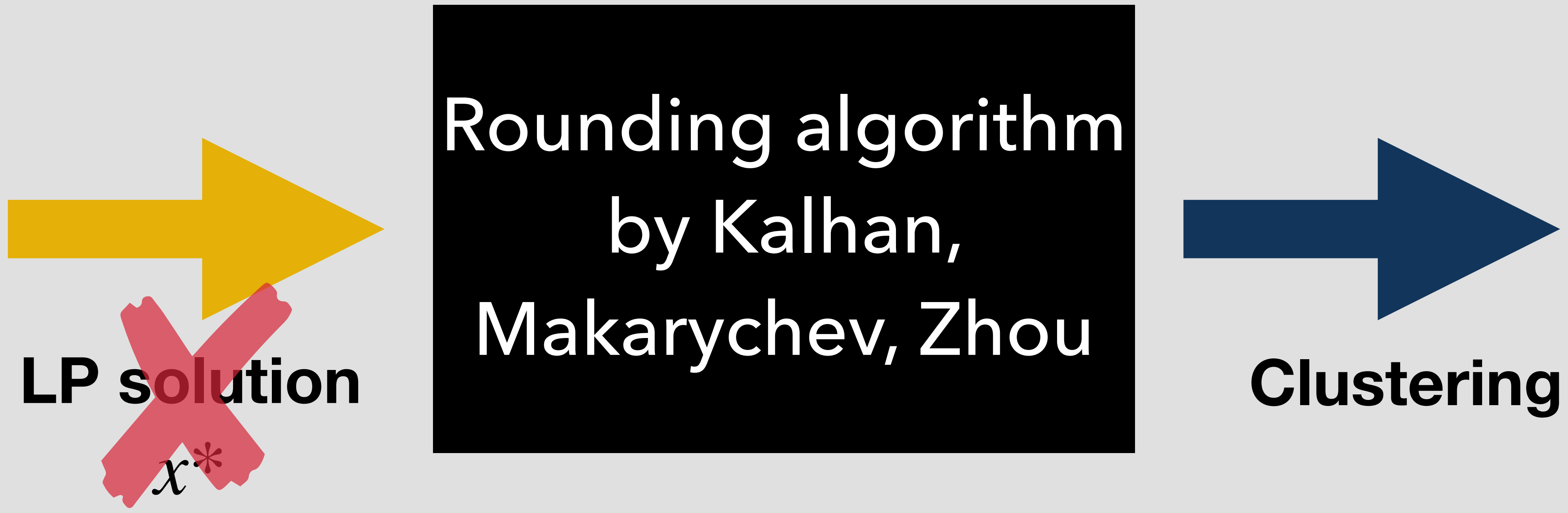
$$\min \|y\|_p$$

$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$

Our technique



Convex program relaxation

Don't solve this!

Quantify probabilities x_{uv} combinatorially

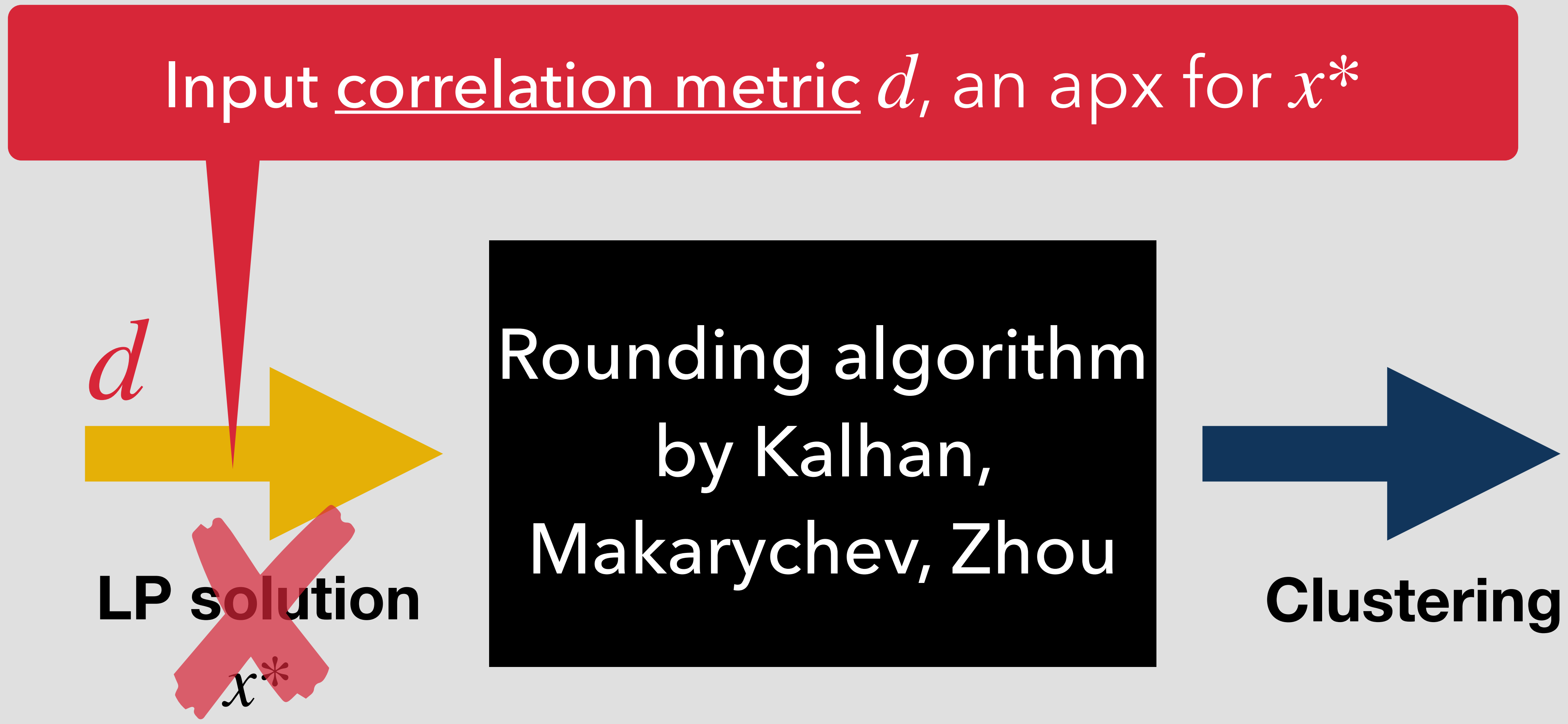
$$\min \|y\|_p$$

$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \le x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \le x_{uv} \le 1 \qquad \forall u, v \in V$$

Our technique



Convex program relaxation

Don't solve this!

Quantify probabilities x_{uv} combinatorially

$$\min \|y\|_p$$

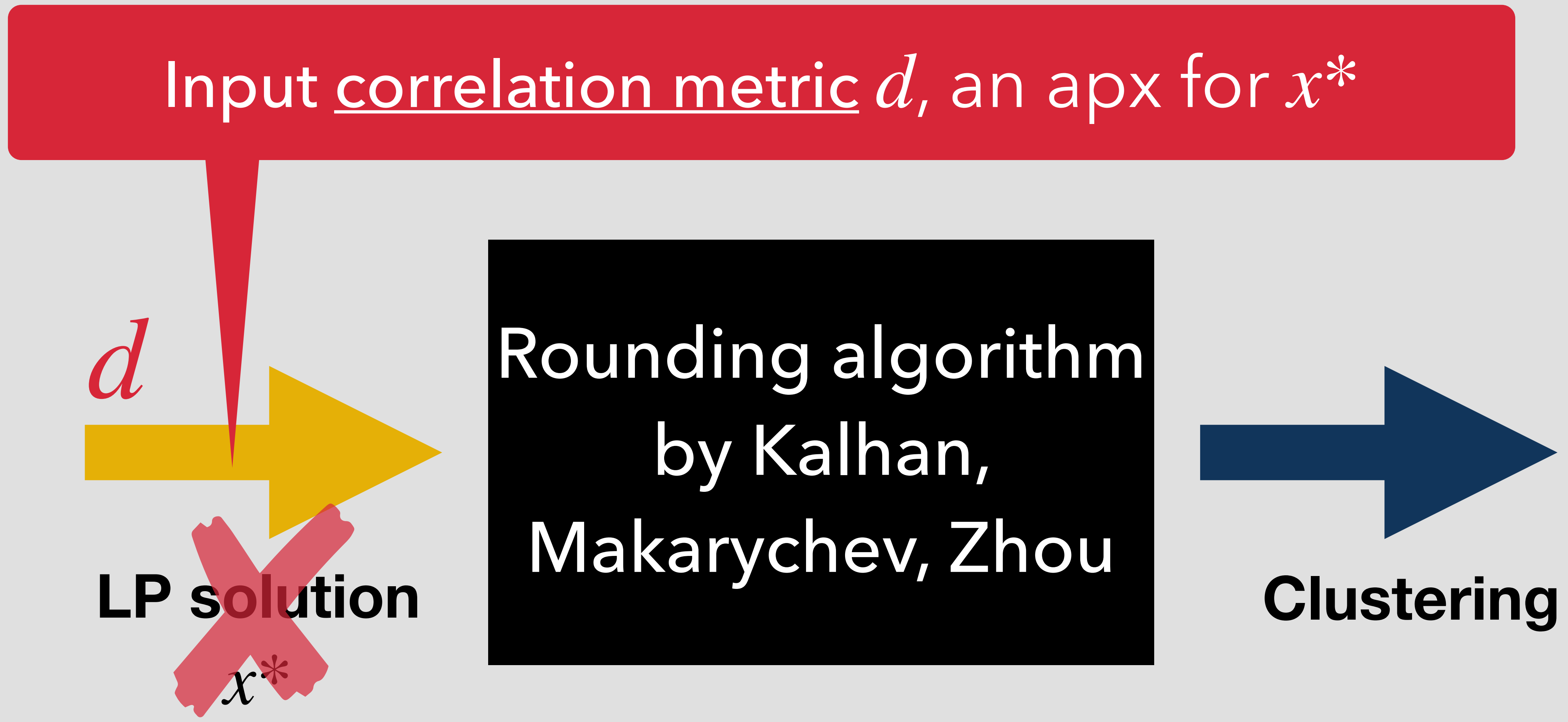
$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$

Our technique

Desiderata for correlation metric d for ℓ_p:



Convex program relaxation

Don't solve this!

Quantify probabilities x_{uv} combinatorially

$$\min \|y\|_p$$

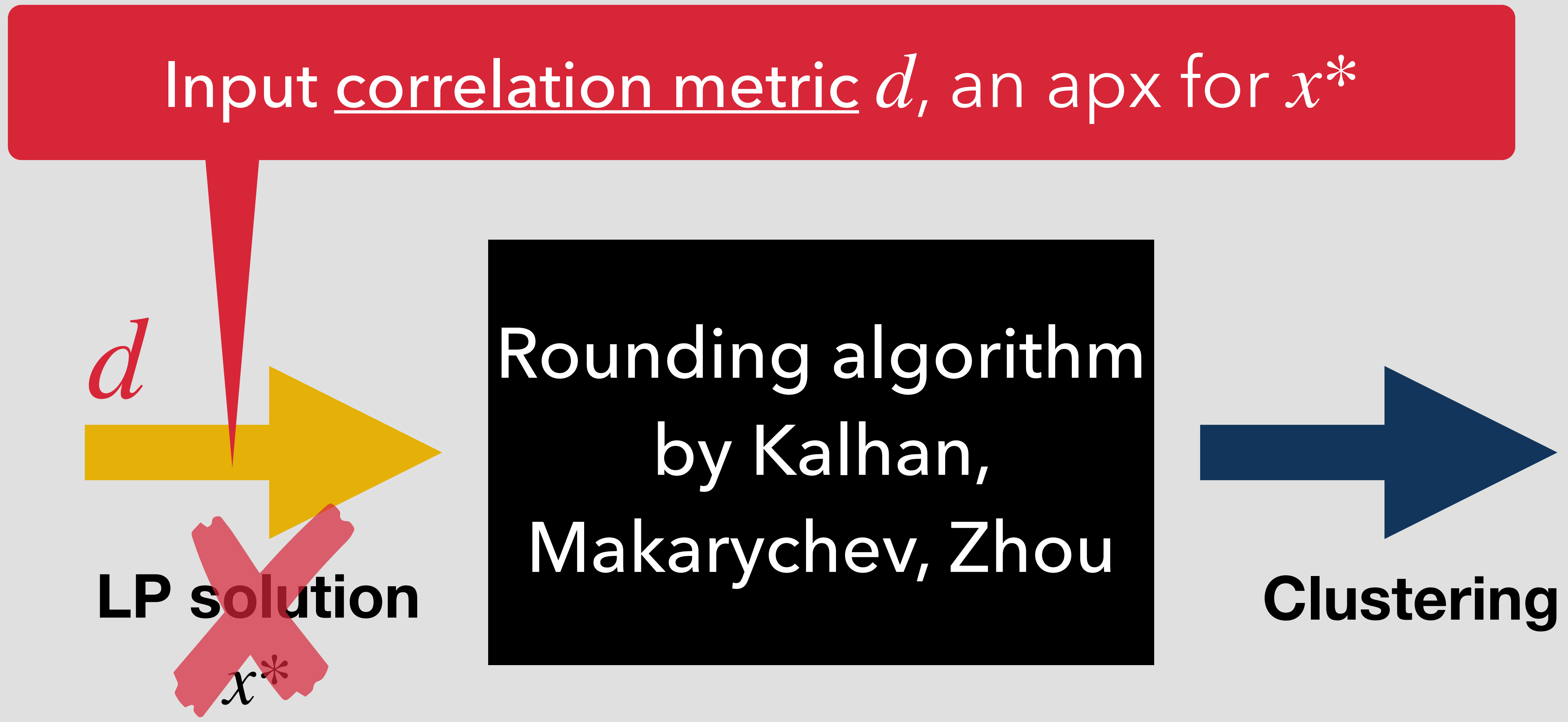
$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$

Our technique

Desiderata for correlation metric d for ℓ_p:



Convex program relaxation

Don't solve this!

Quantify probabilities x_{uv} combinatorially

$$\min \|y\|_p$$

$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

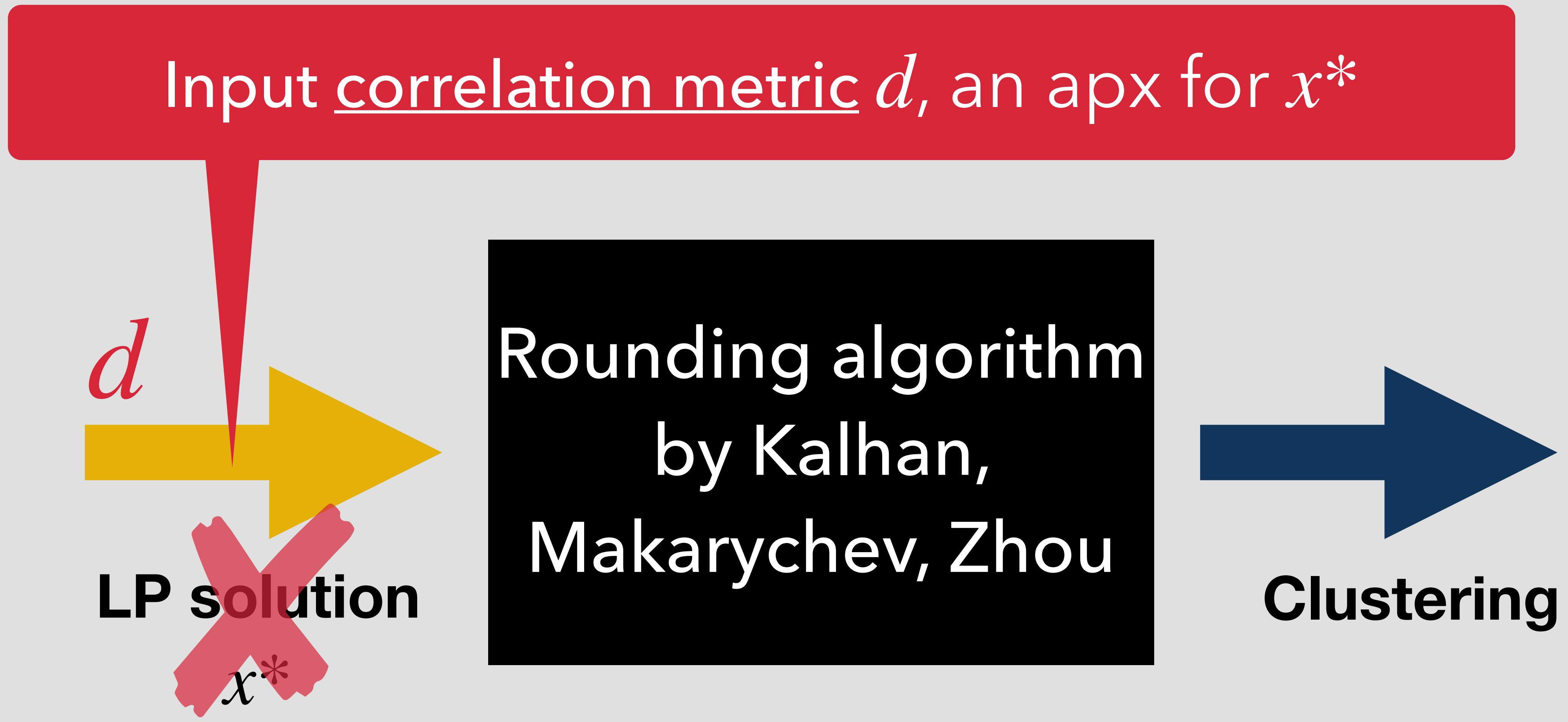
$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$

Our technique

Desiderata for correlation metric d for ℓ_p:

(1) satisfies triangle inequality



Convex ~~program~~ relaxation

Don't solve this!

Quantify probabilities x_{uv} combinatorially

$$\min \|y\|_p$$

$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

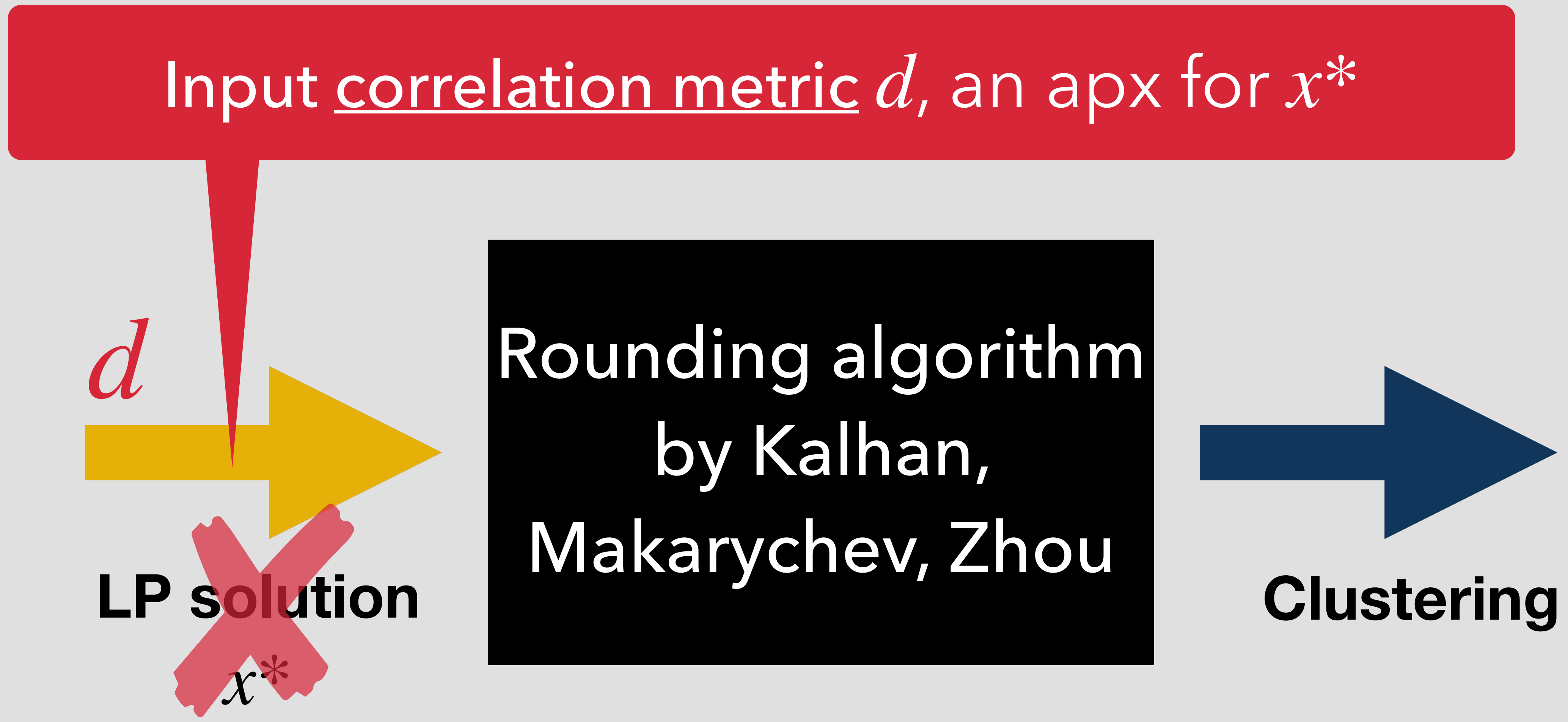
$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$

Our technique

Desiderata for correlation metric d for ℓ_p:

(1) satisfies triangle inequality



Convex program relaxation

Don't solve this!

Quantify probabilities x_{uv} combinatorially

$$\min \|y\|_p$$

Lower bound on (integral) OPT

$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

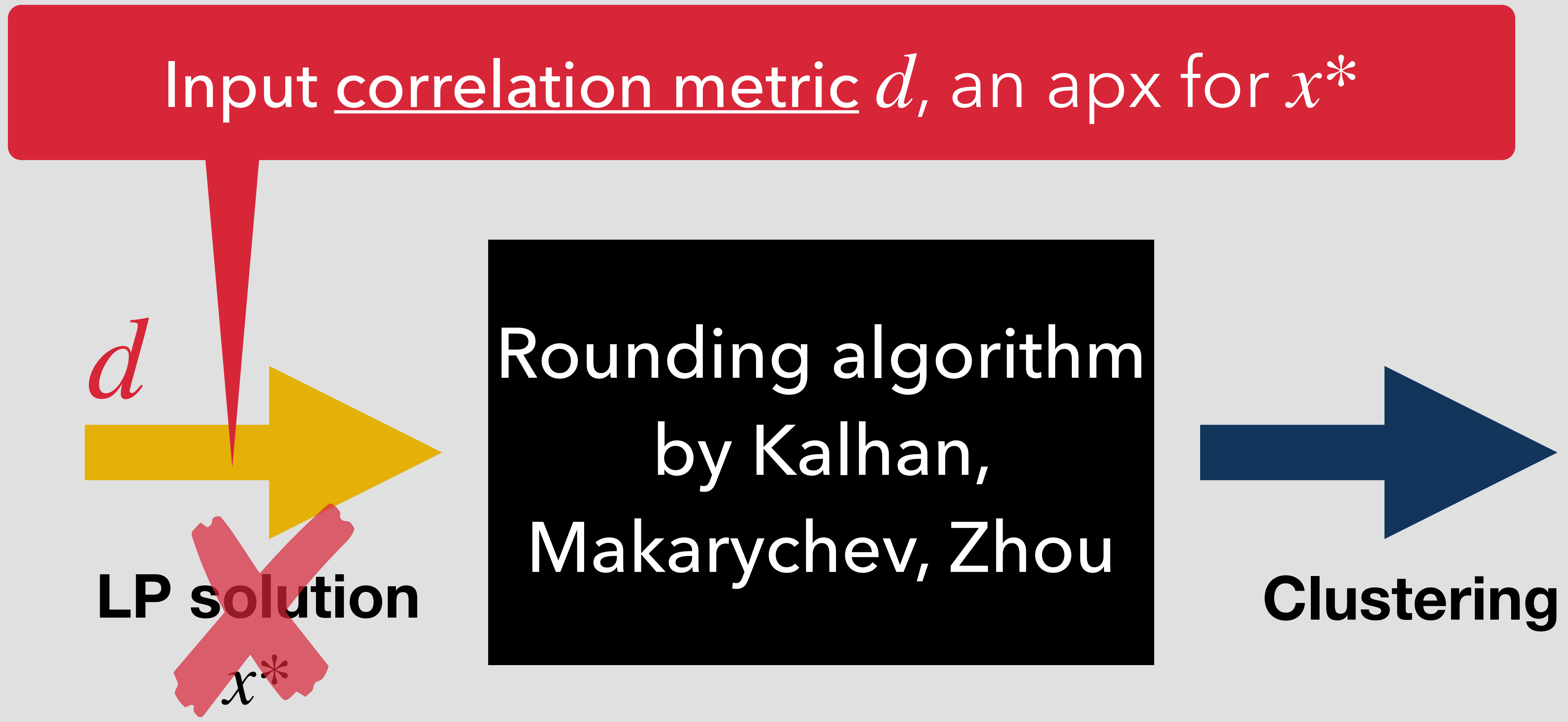
$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$

Our technique

Desiderata for correlation metric d for ℓ_p:

(1) satisfies triangle inequality

(2) $\sum_{u \in V} \left(\sum_{v \in N_u^+} d_{uv} + \sum_{v \in N_u^-} (1 - d_{uv}) \right)^p \leq O(1) \cdot \mathrm{OPT}_p$



Convex ~~program~~ relaxation

Don't solve this!

Quantify probabilities x_{uv} combinatorially

$\min \|y\|_p$ — Lower bound on (integral) OPT

$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$

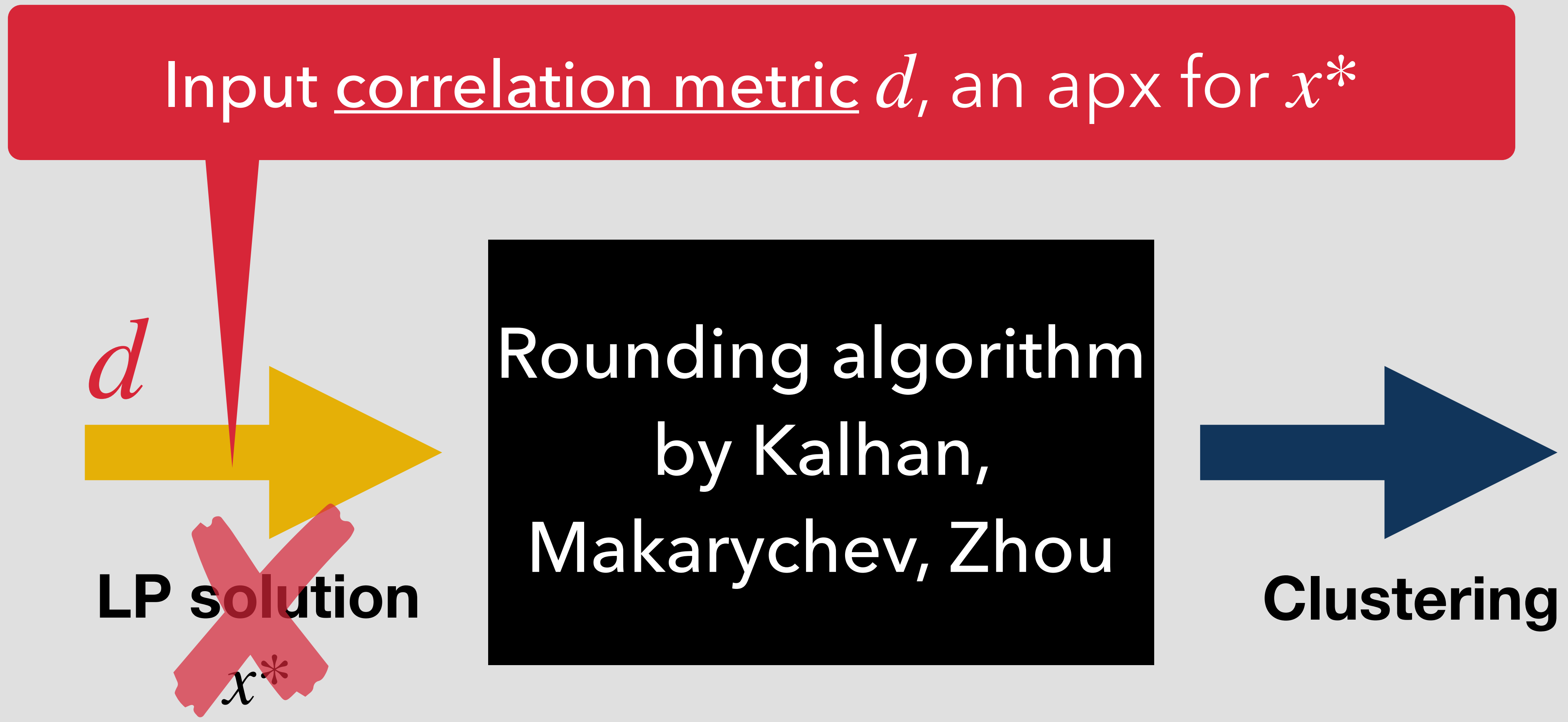
Our technique

Desiderata for correlation metric d for ℓ_p:

(1) satisfies triangle inequality

(2) $\sum_{u \in V} \left(\sum_{v \in N_u^+} d_{uv} + \sum_{v \in N_u^-} (1 - d_{uv}) \right)^p \leq O(1) \cdot \mathrm{OPT}_p$

frac. cost of d = ℓ_p-norm of (fractional) disagreements



Convex program relaxation

Don't solve this!

Quantify probabilities x_{uv} combinatorially

$\min \|y\|_p$ — Lower bound on (integral) OPT

$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$

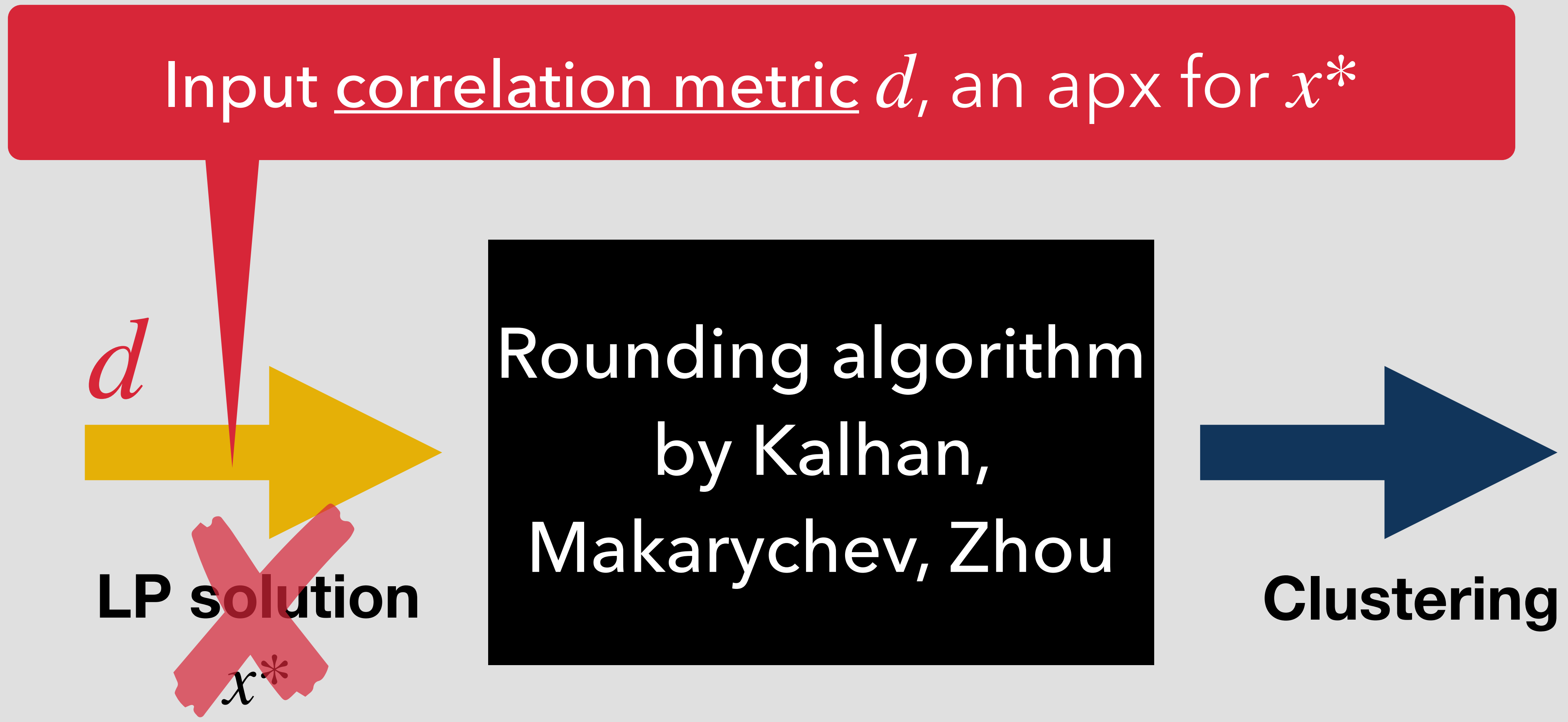
Our technique

Desiderata for correlation metric d for ℓ_p:

(1) satisfies triangle inequality

(2) $\sum_{u \in V} \left(\sum_{v \in N_u^+} d_{uv} + \sum_{v \in N_u^-} (1 - d_{uv}) \right)^p \leq O(1) \cdot \mathrm{OPT}_p$

frac. cost of d = ℓ_p-norm of (fractional) disagreements



Convex ~~program~~ relaxation

Don't solve this!

Quantify probabilities x_{uv} combinatorially

$$\min \|y\|_p$$

Lower bound on (integral) OPT

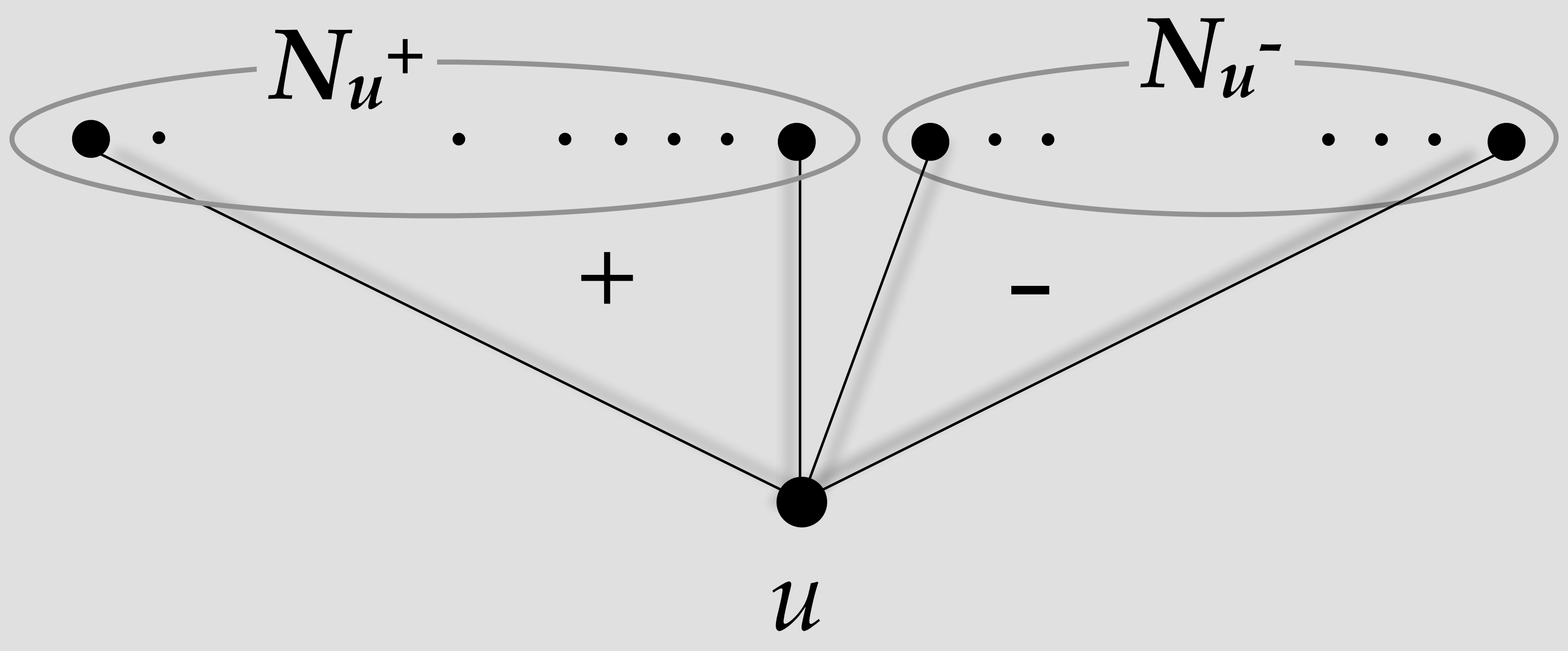
$$y(u) = \sum_{v \in N_u^+} x_{uv} + \sum_{v \in N_u^-} (1 - x_{uv}) \qquad \forall u \in V$$

$$x_{uv} \leq x_{vw} + x_{uw} \qquad \forall u, v, w \in V$$

$$0 \leq x_{uv} \leq 1 \qquad \forall u, v \in V$$

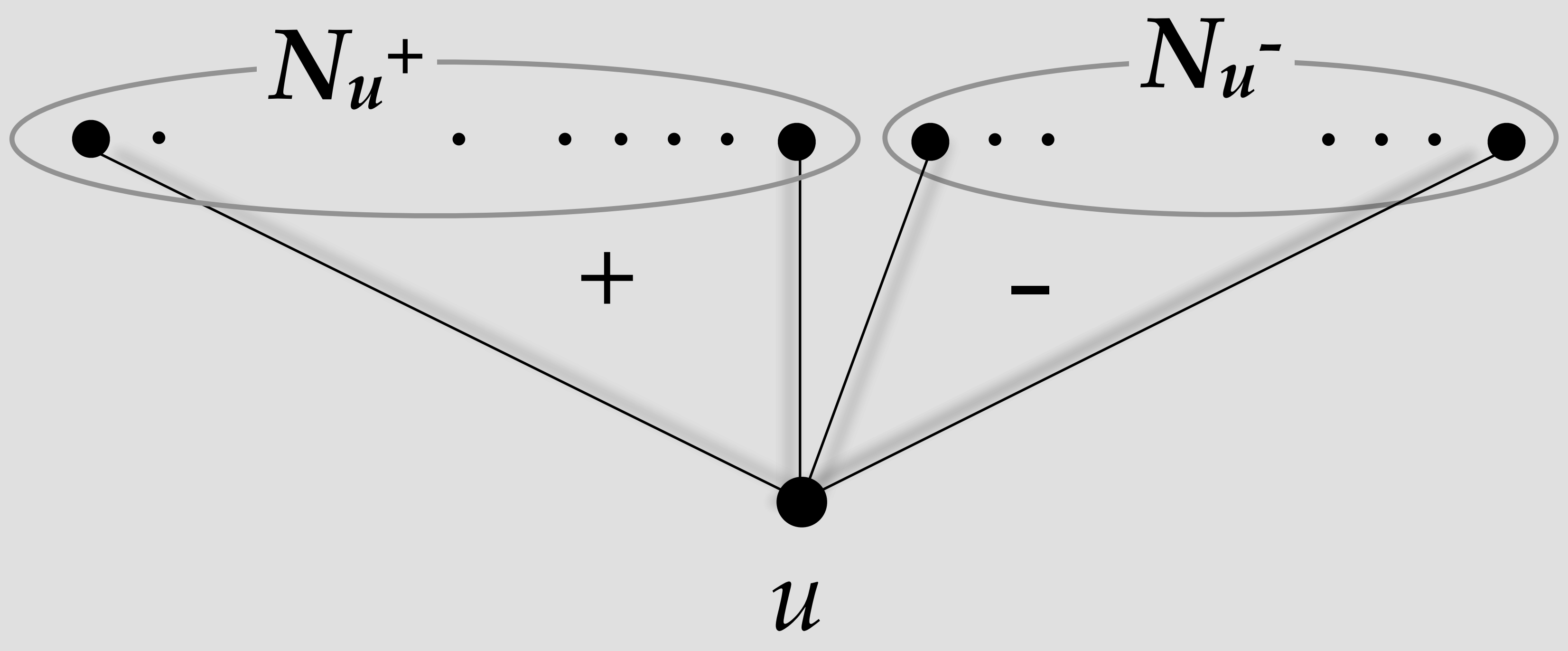
Correlation metric

Correlation metric



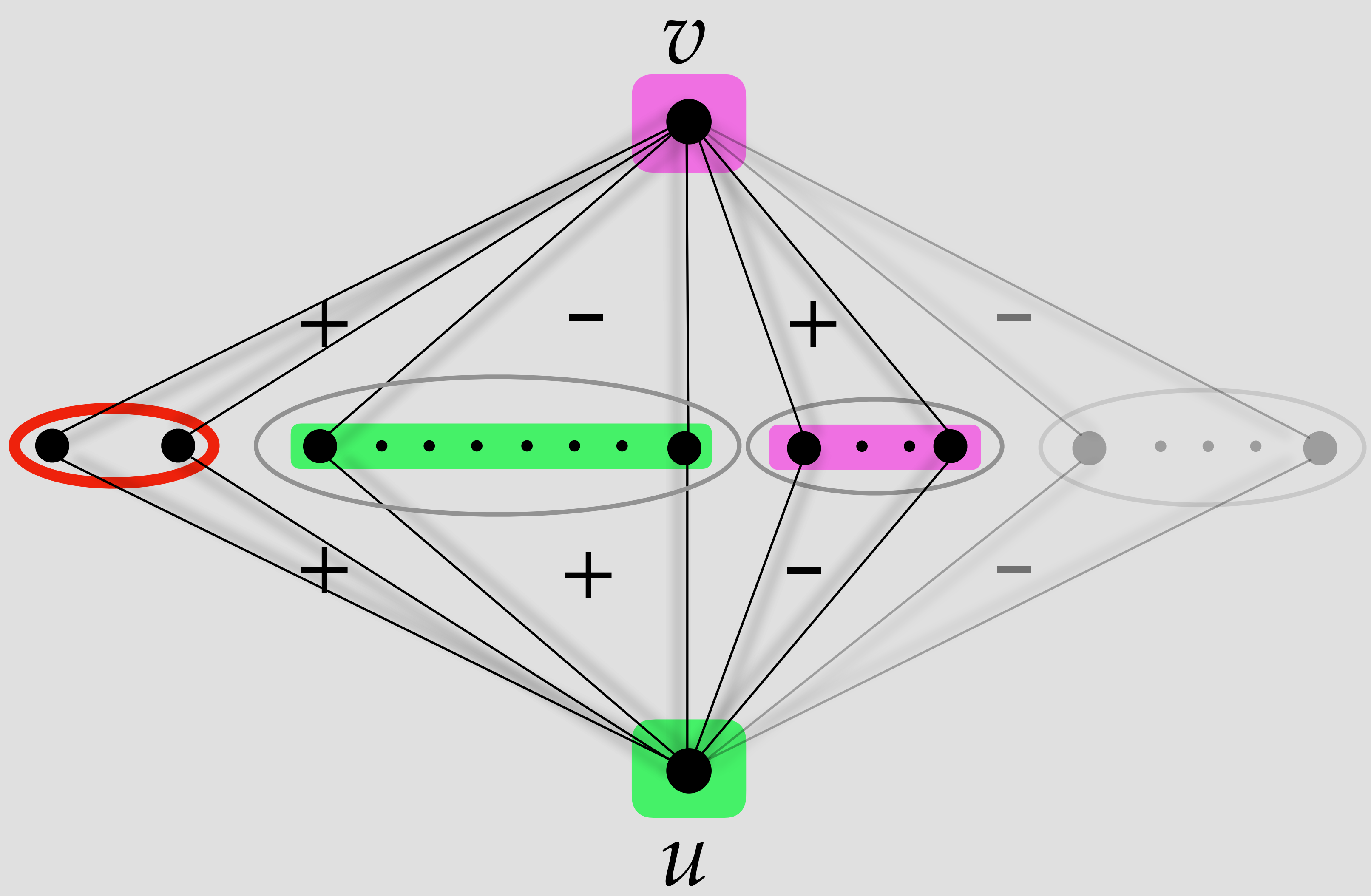
Correlation metric

- N_u^+ = (+) neighbors of u,
- N_u^- = (–) neighbors of u



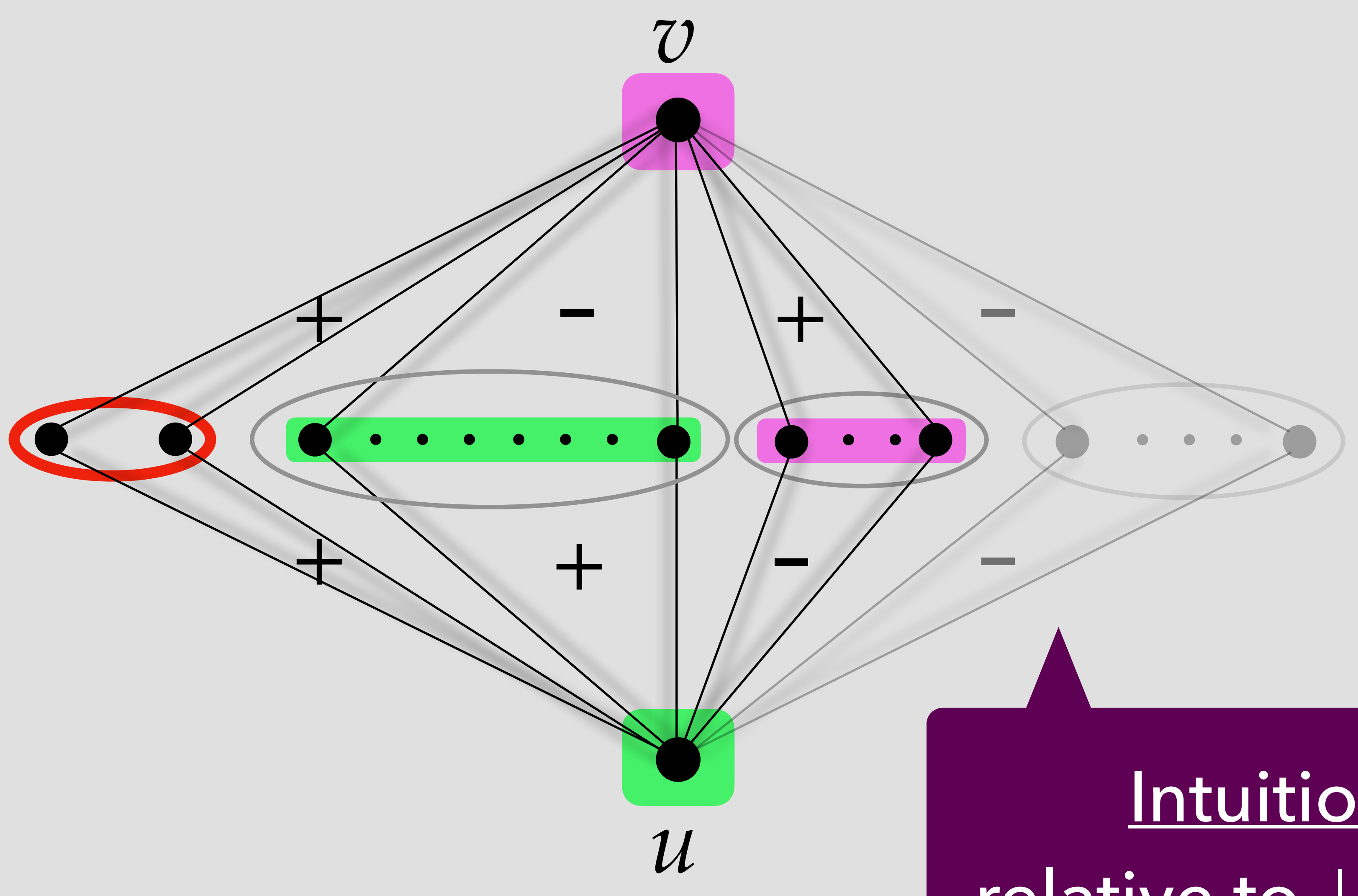
Correlation metric

- N_u^+ = (+) neighbors of u
- N_u^- = (–) neighbors of u



Correlation metric

- N_u^+ = (+) neighbors of u
- N_u^- = (–) neighbors of u

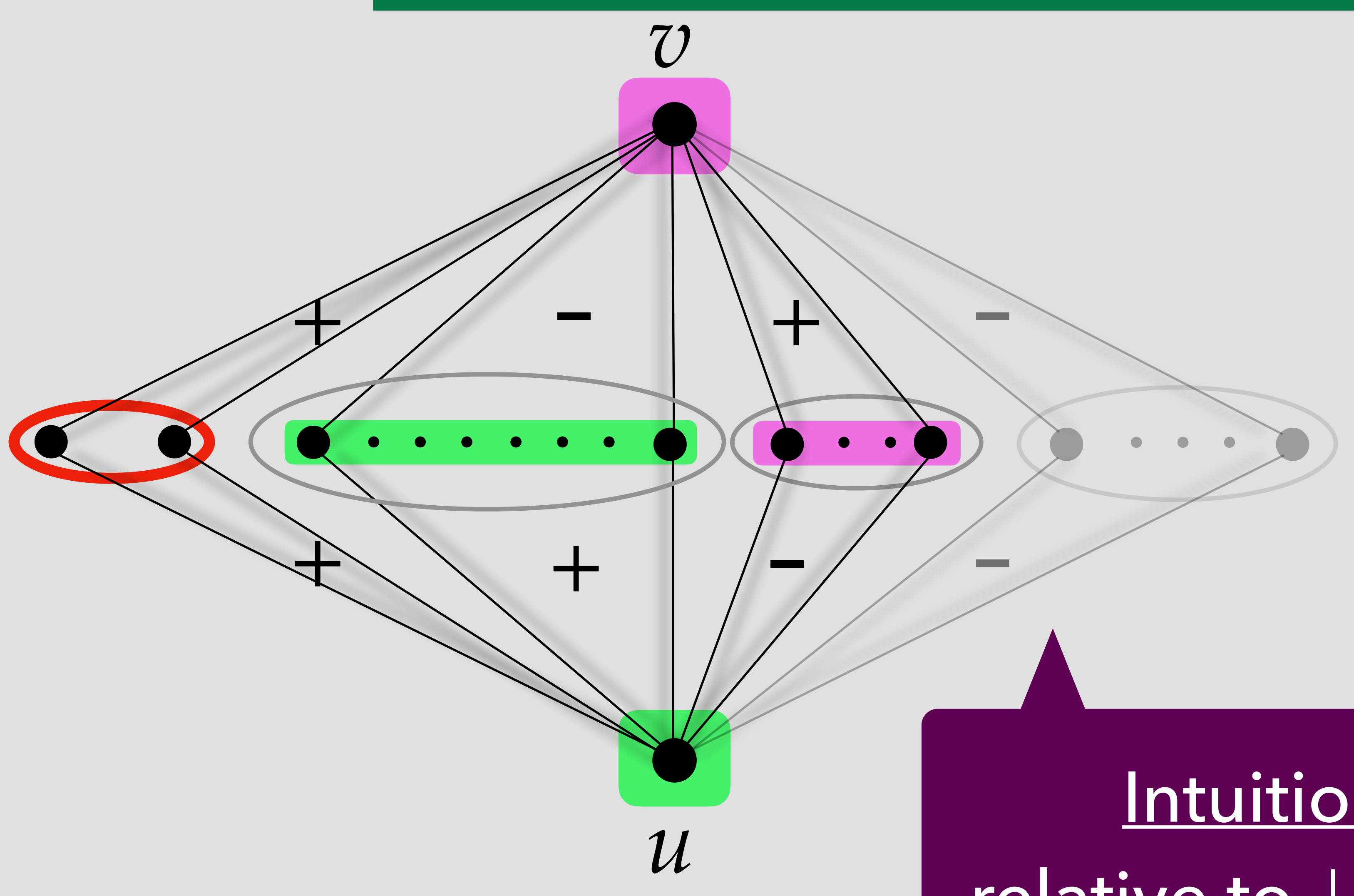


Intuition: if u and v have large mixed nbhds relative to $|N_u^+ \cup N_v^+|$, want them in different clusters

Correlation metric

- N_u^+ = (+) neighbors of u
- N_u^- = (–) neighbors of u

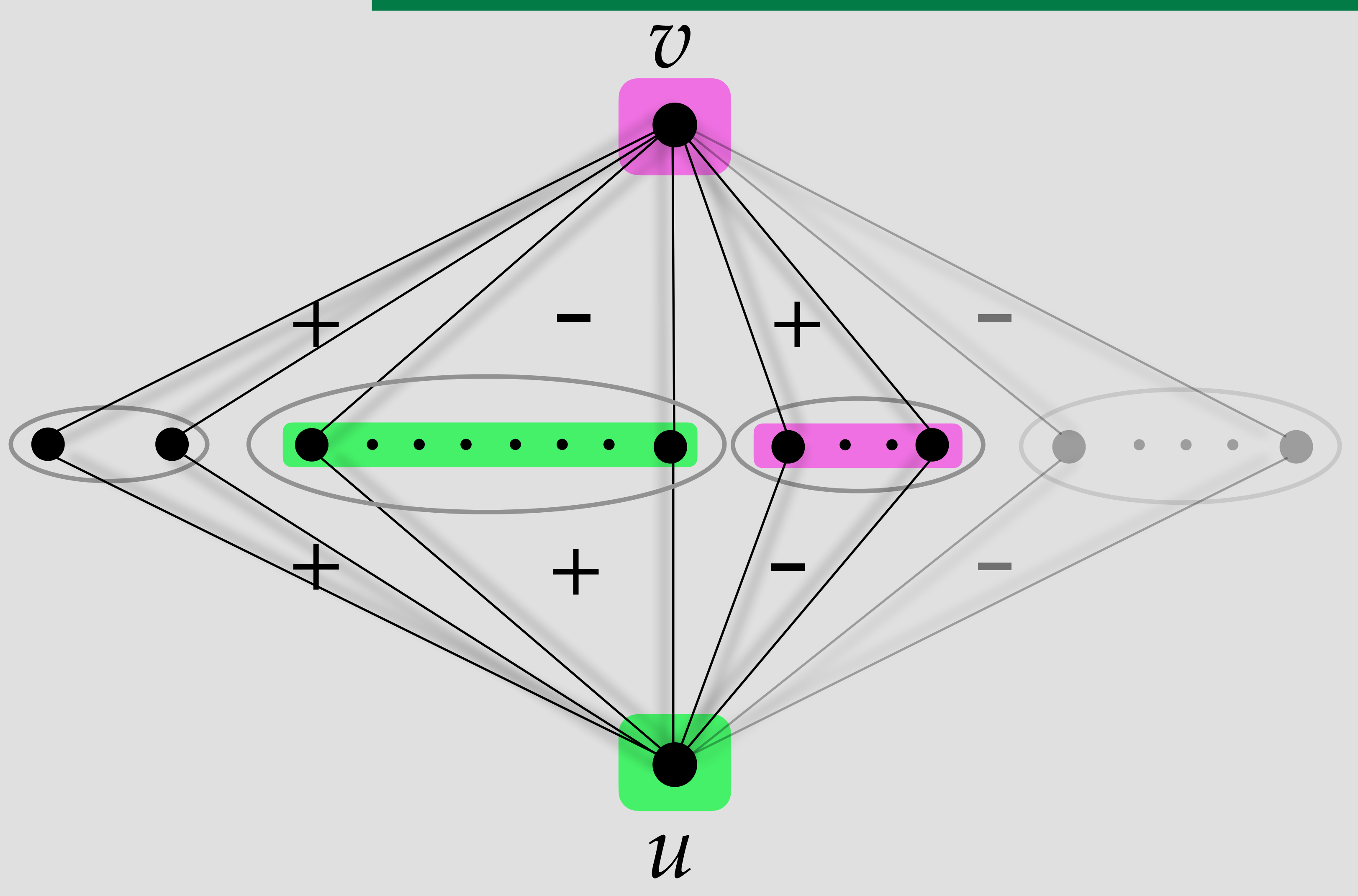
Correlation metric = $d_{uv} = 1 - \frac{|N_u^+ \cap N_v^+|}{|N_u^+ \cup N_v^+|} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$



Intuition: if u and v have large mixed nbhds relative to $|N_u^+ \cup N_v^+|$, want them in different clusters

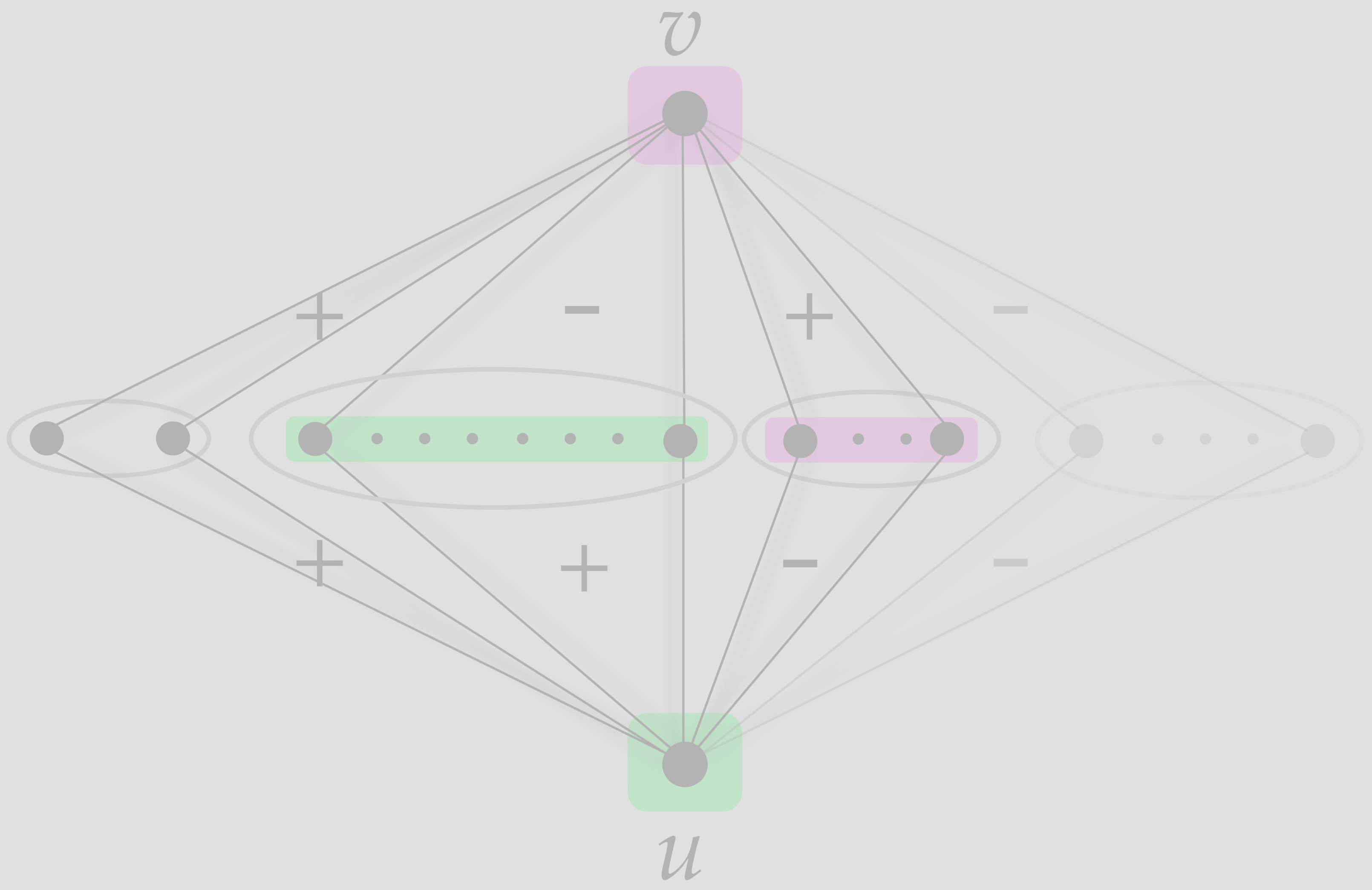
Correlation metric

Correlation metric = $d_{uv} = 1 - \frac{|N_u^+ \cap N_v^+|}{|N_u^+ \cup N_v^+|} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$



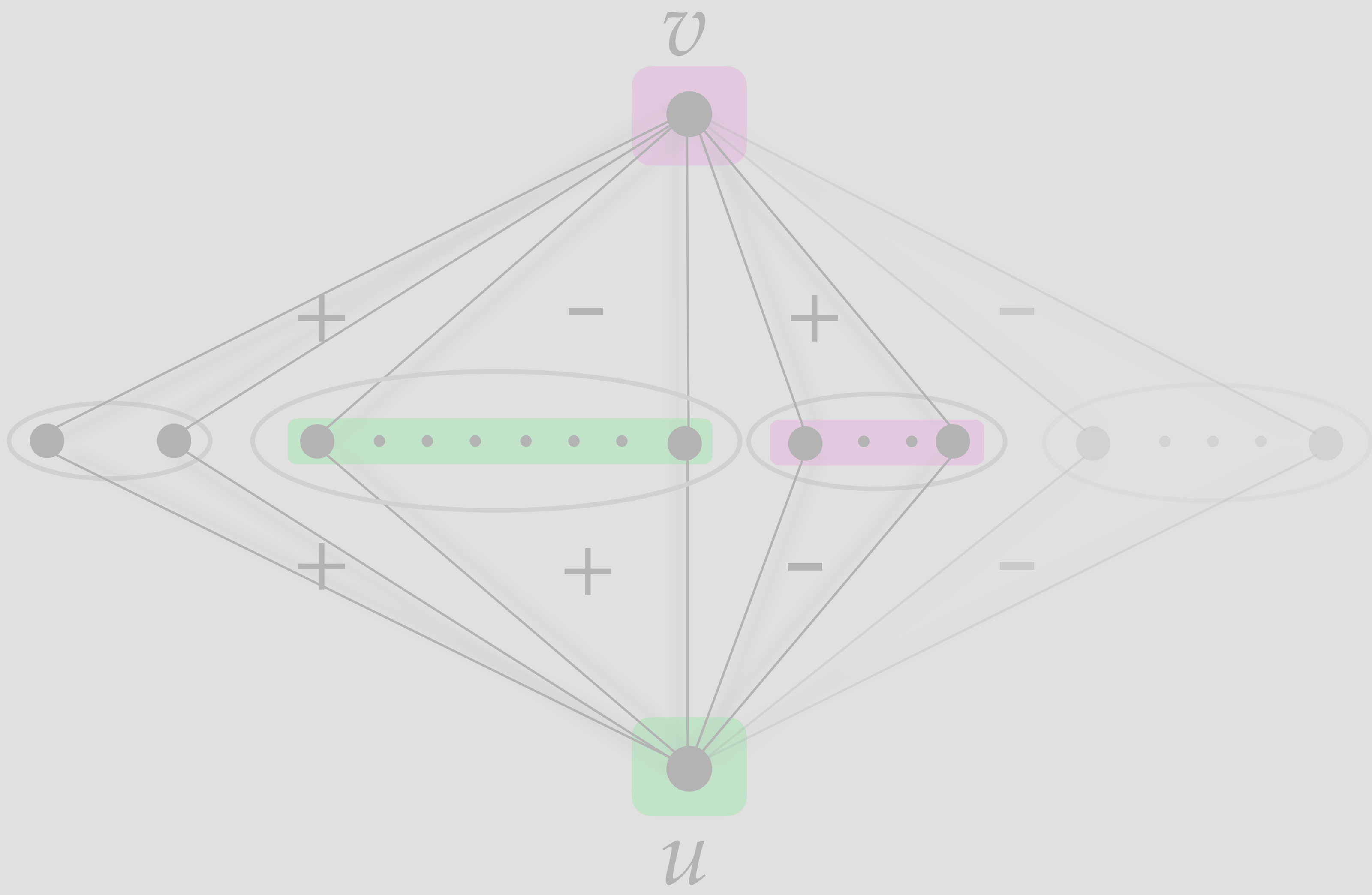
Correlation metric

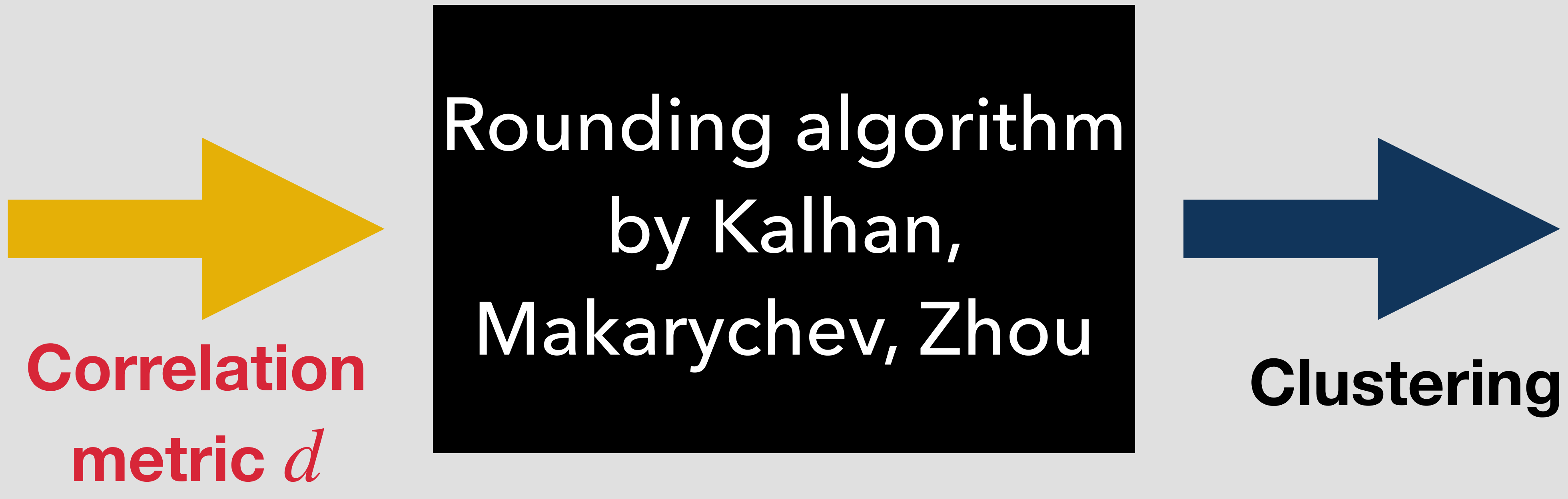
Correlation metric = $d_{uv} = 1 - \frac{|N_u^+ \cap N_v^+|}{|N_u^+ \cup N_v^+|} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$



Correlation metric

Correlation metric = $d_{uv} = 1 - \dfrac{|N_u^+ \cap N_v^+|}{|N_u^+ \cup N_v^+|} = \dfrac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$

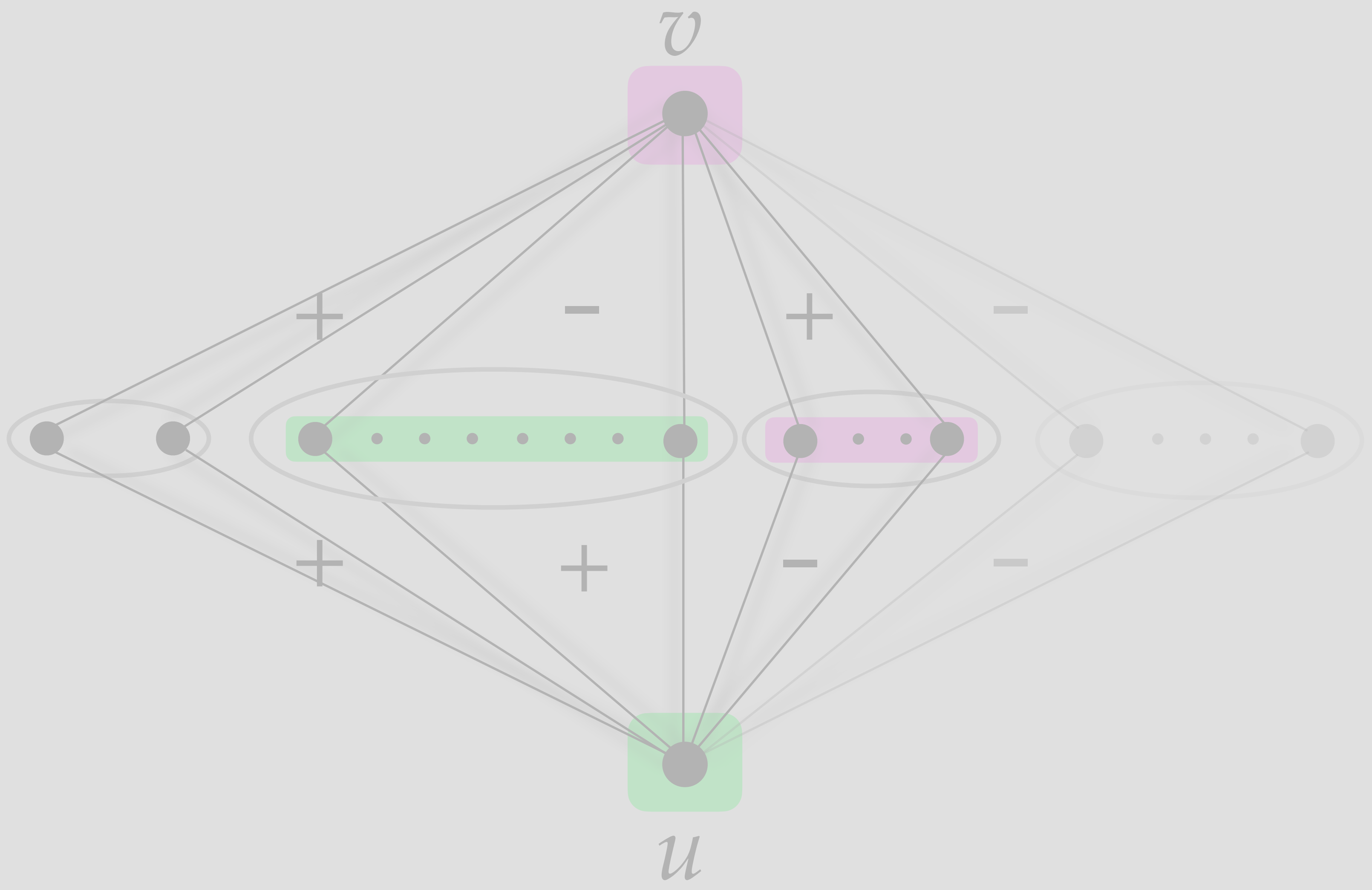


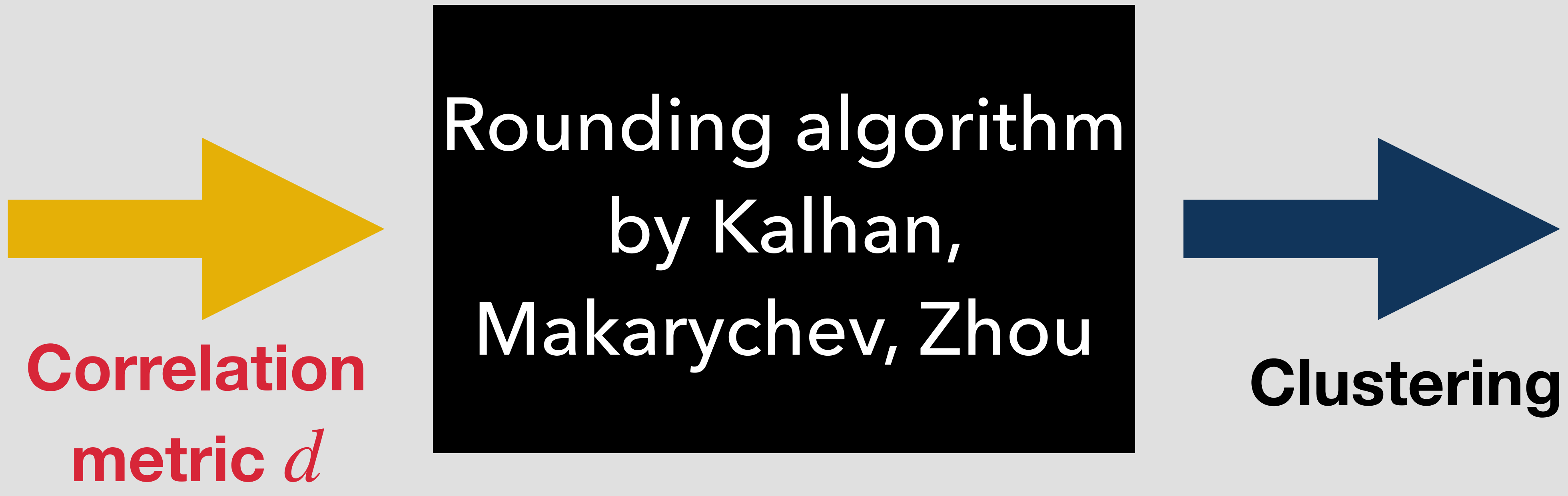


Correlation metric

Works as is for ℓ_∞ norm objective

Correlation metric = $d_{uv} = 1 - \frac{|N_u^+ \cap N_v^+|}{|N_u^+ \cup N_v^+|} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$





Today

- ✦~~The **correlation metric** (constructing a "guess" for the optimal solution to convex relaxation)~~
- ✦Tweaking correlation metric for **all** ℓ_p-norms
- ✦Open questions

Today

- ✦~~The **correlation metric** **(constructing a "guess" for the optimal solution to convex relaxation)**~~
- ✦Tweaking correlation metric for **all** ℓ_p-norms
- ✦Open questions

Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Goal: find single clustering that is $O(1)$-apx for **all** ℓ_p-norms **simultaneously**

Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Goal: find single clustering that is $O(1)$-apx for **all** ℓ_p-norms **simultaneously**

Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Goal: find single clustering that is $O(1)$-apx for **all** ℓ_p-norms **simultaneously**

Key Idea 1: for ***regular graphs***, correlation metric also $O(1)$–apxs all ℓ_p-norms!

Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Goal: find single clustering that is $O(1)$-apx for **all** ℓ_p-norms **simultaneously**

Key Idea 1: for ***regular graphs***, correlation metric also $O(1)$–apxs all ℓ_p-norms!

- Proof via dual fitting when $p = 1$

Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Goal: find single clustering that is $O(1)$-apx for **all** ℓ_p-norms **simultaneously**

Key Idea 1: for ***regular graphs***, correlation metric also $O(1)$–apxs all ℓ_p-norms!

- Proof via dual fitting when $p = 1$
- Problem is when (+) subgraph is far from regular

Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Goal: find single clustering that is $O(1)$-apx for **all** ℓ_p-norms **simultaneously**

Key Idea 1: for ***regular graphs***, correlation metric also $O(1)$–apxs all ℓ_p-norms!

- Proof via dual fitting when $p = 1$
- Problem is when (+) subgraph is far from regular

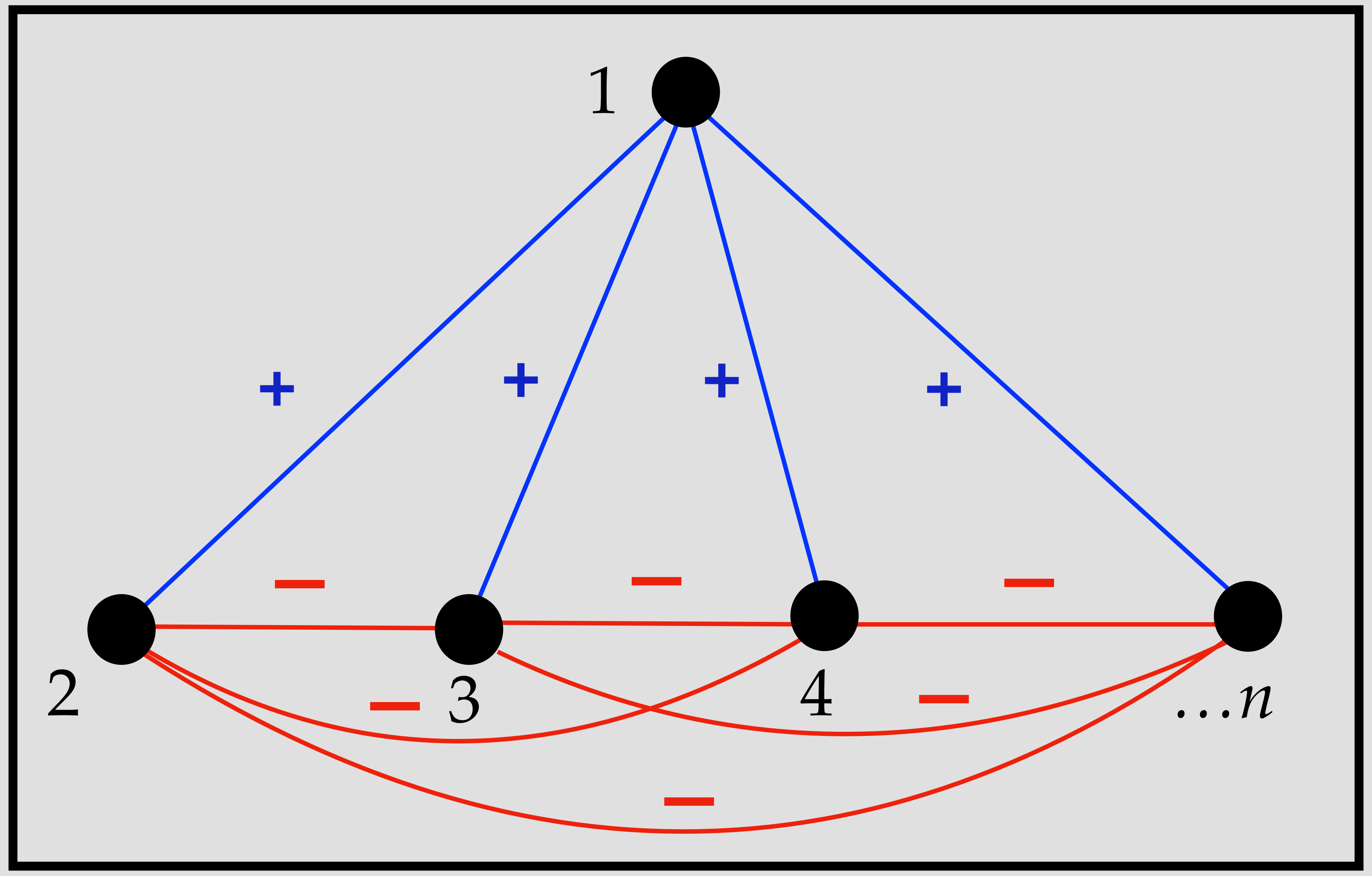
Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Goal: find single clustering that is $O(1)$-apx for **all** ℓ_p-norms **simultaneously**

Key Idea 1: for ***regular graphs***, correlation metric also $O(1)$–apxs all ℓ_p-norms!

- Proof via dual fitting when $p = 1$
- Problem is when (+) subgraph is far from regular



Adjusted correlation metric

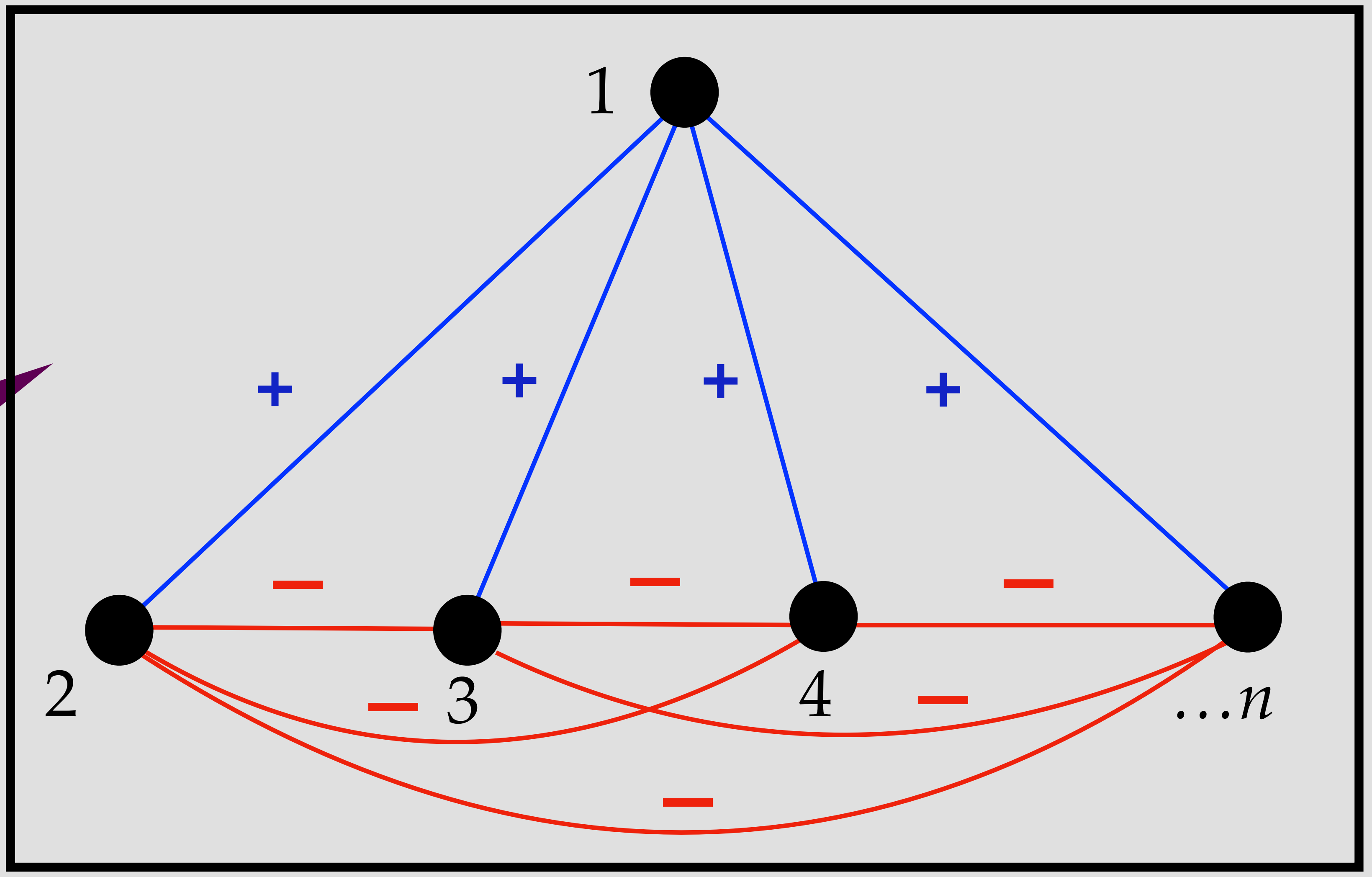
$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Goal: find single clustering that is $O(1)$-apx for **all** ℓ_p-norms **simultaneously**

Key Idea 1: for ***regular graphs***, correlation metric also $O(1)$–apxs all ℓ_p-norms!

- Proof via dual fitting when $p = 1$
- Problem is when (+) subgraph is far from regular

d_{uv} = 2/3 for all u,v in negative clique
frac. cost of d = $\theta(n^2)$



Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

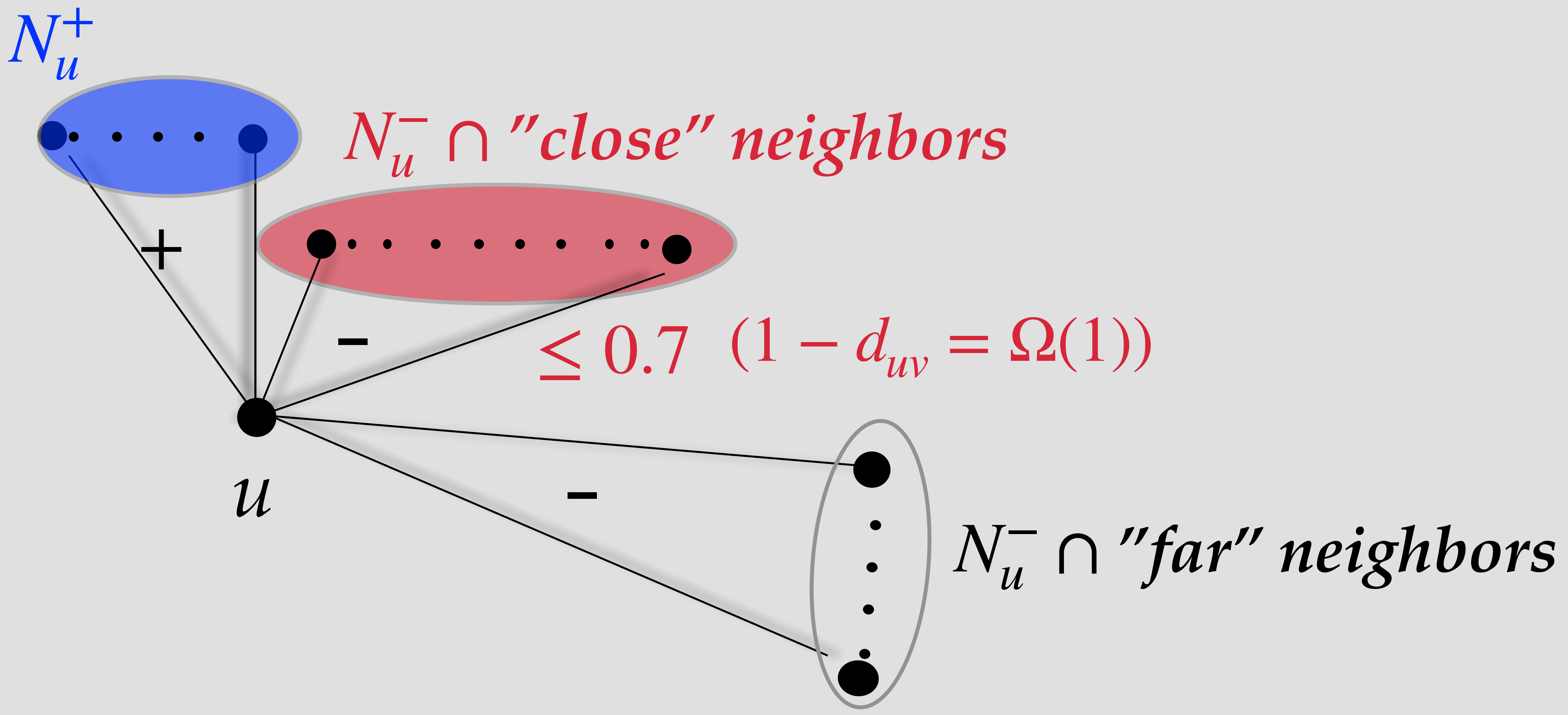
Key Idea 2: for ***positive edges***, correlation metric has ***bounded cost for all p***!

$$\sum_{u \in V} \left(\sum_{v \in N_u^+} d_{uv} + \sum_{v \in N_u^-} (1 - d_{uv}) \right)^p \leq O(1) \cdot \mathrm{OPT}_p$$

Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

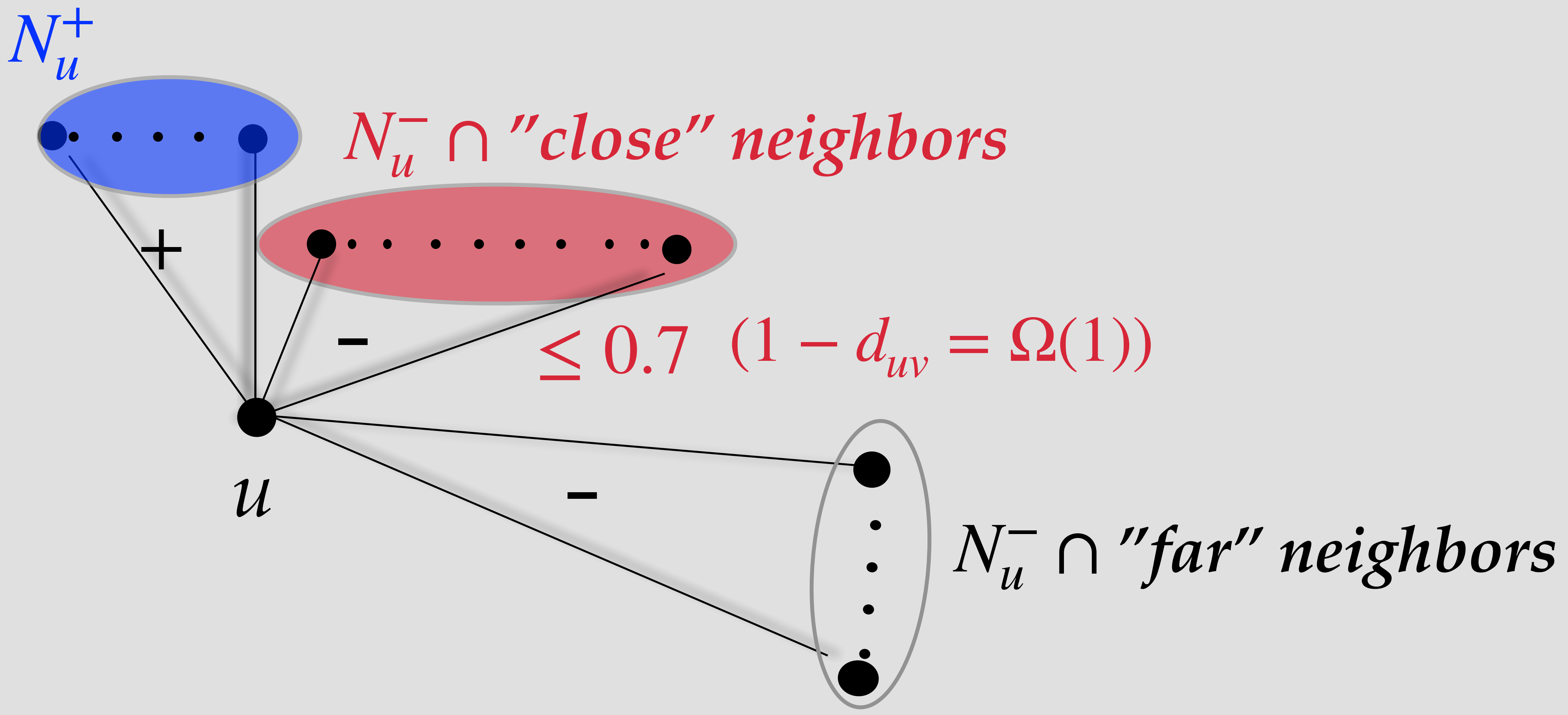
Key Idea 2: for ***positive edges***, correlation metric has *bounded cost for all p*!



Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

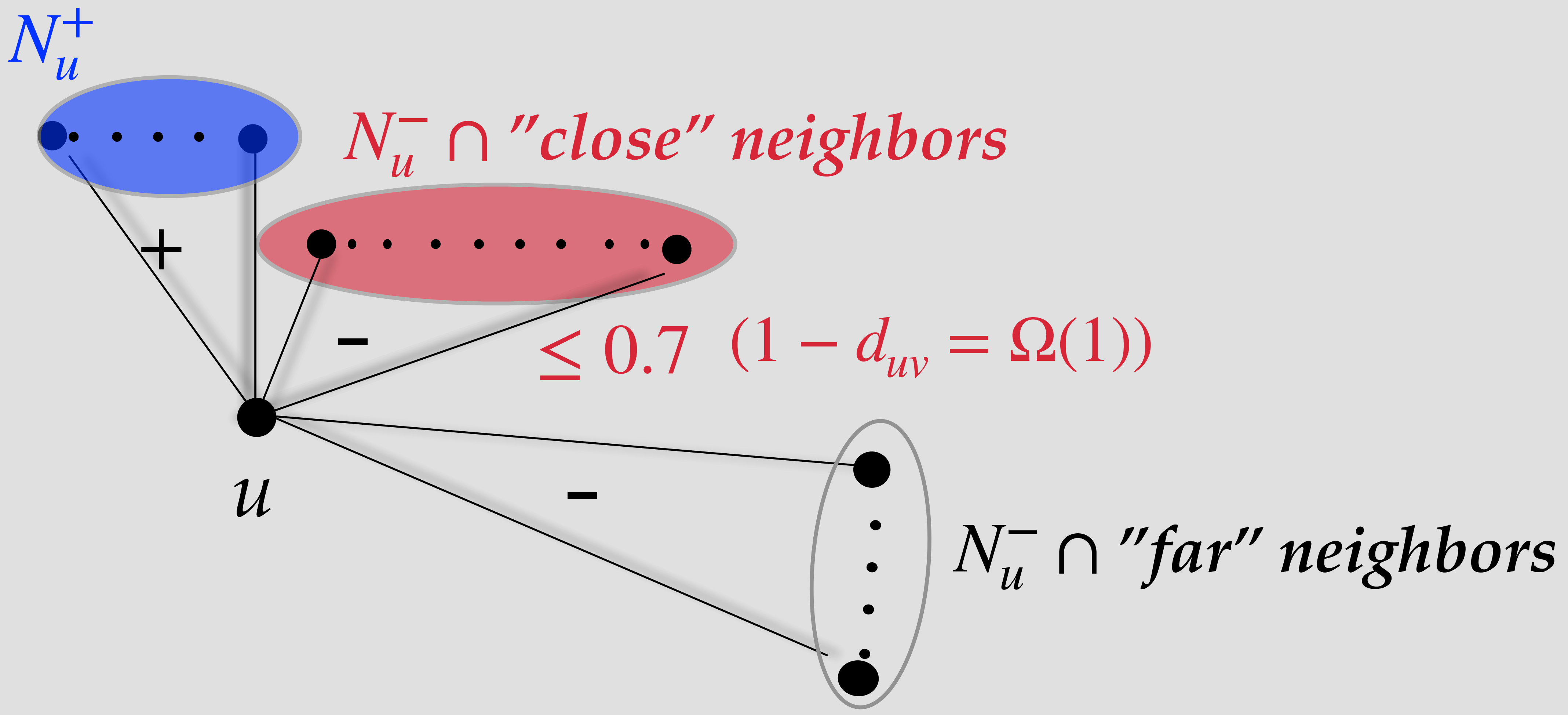
Key Idea 2: for ***positive edges***, correlation metric has ***bounded cost for all p***!



Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Key Idea 2: for ***positive edges***, correlation metric has *bounded cost for all p*!

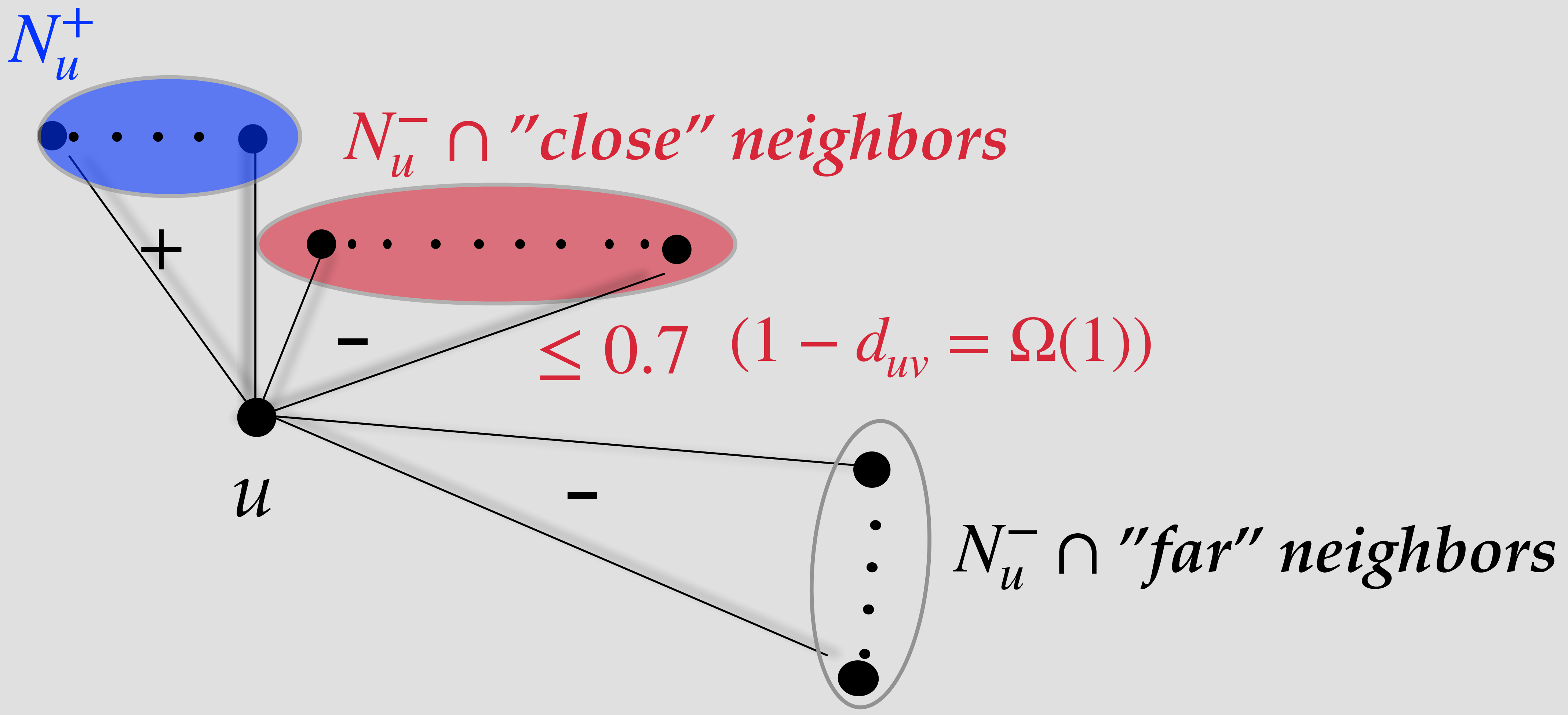


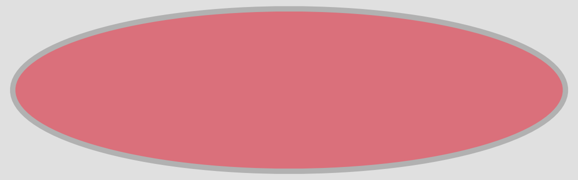
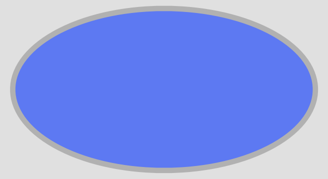
If [red ellipse] >> [blue ellipse] : round all edges out of u to 1 → pay $|N_u^+|^p$ for u

Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Key Idea 2: for ***positive edges***, correlation metric has ***bounded cost for all*** p!

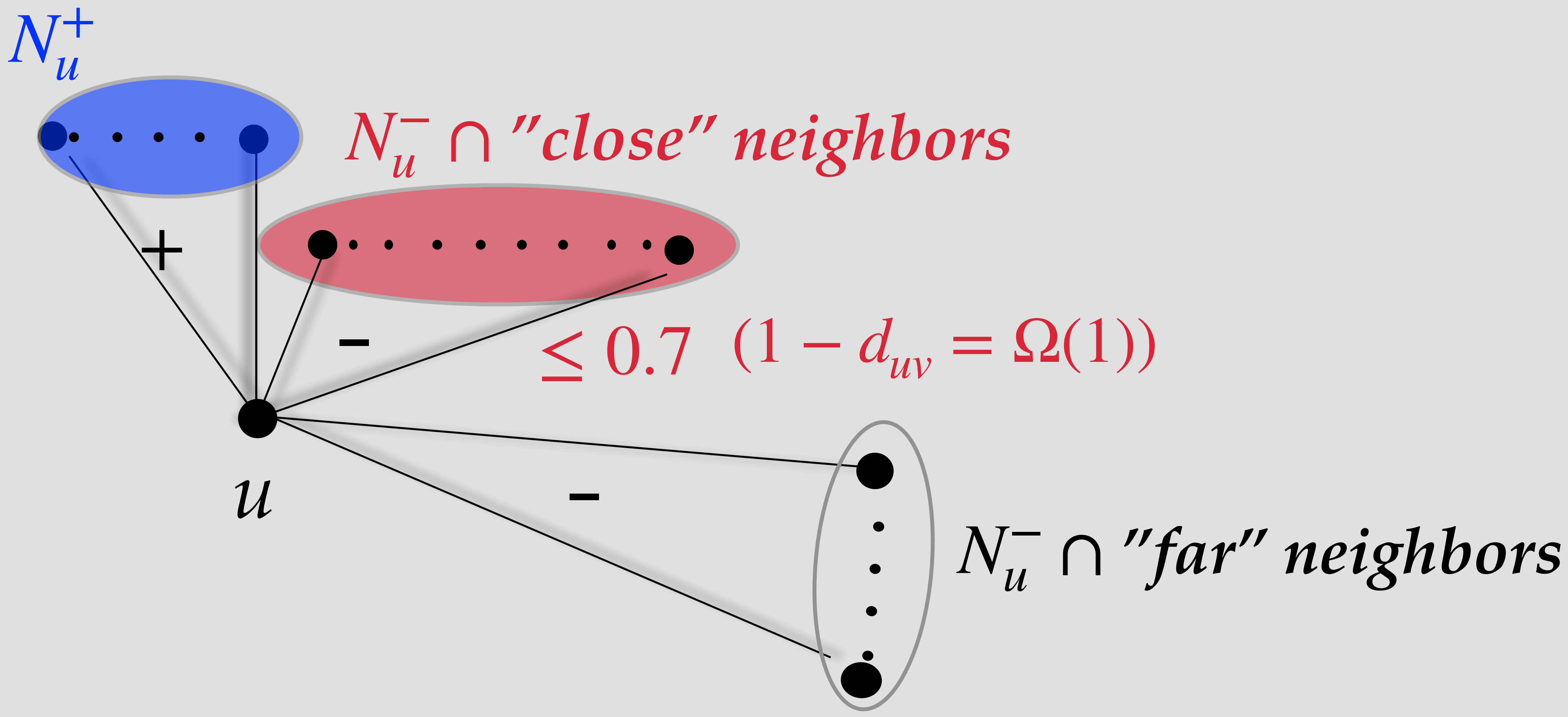


If  >>  : round all edges out of u to 1 → pay $|N_u^+|^p$ for u

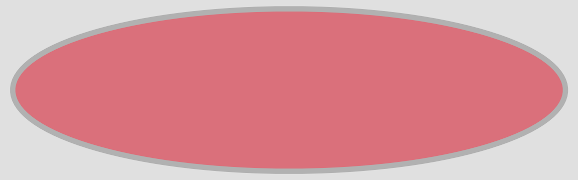
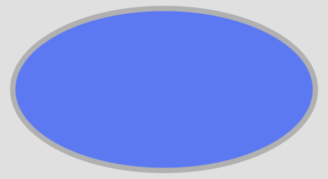
Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Key Idea 2: for ***positive edges***, correlation metric has *bounded cost for all p*!



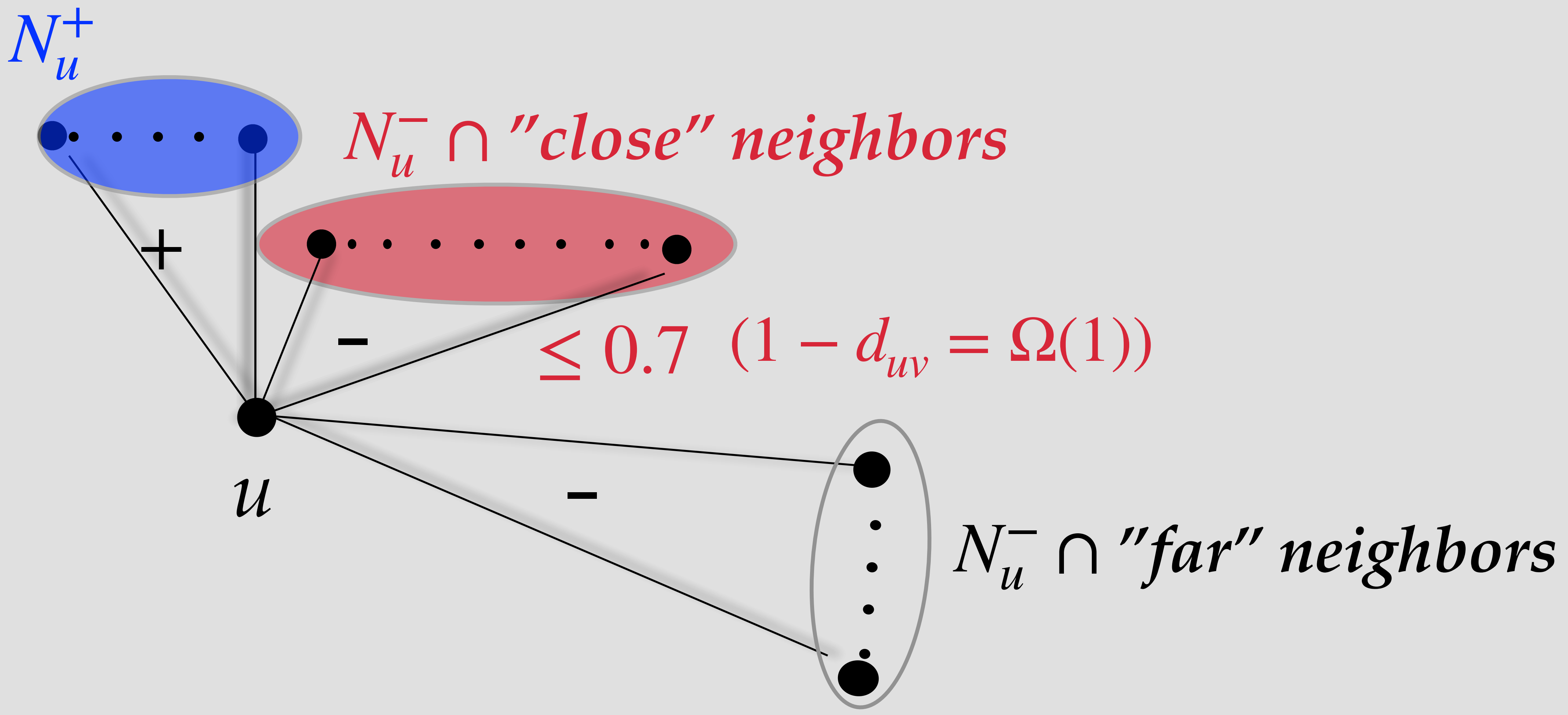
Number of subtleties to showing fractional ℓ_p cost of **adjusted correlation metric** is bounded against OPT_p for any p

If  >>  : round all edges out of u to 1 → pay $|N_u^+|^p$ for u

Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Key Idea 2: for ***positive edges***, correlation metric has *bounded cost for all p*!



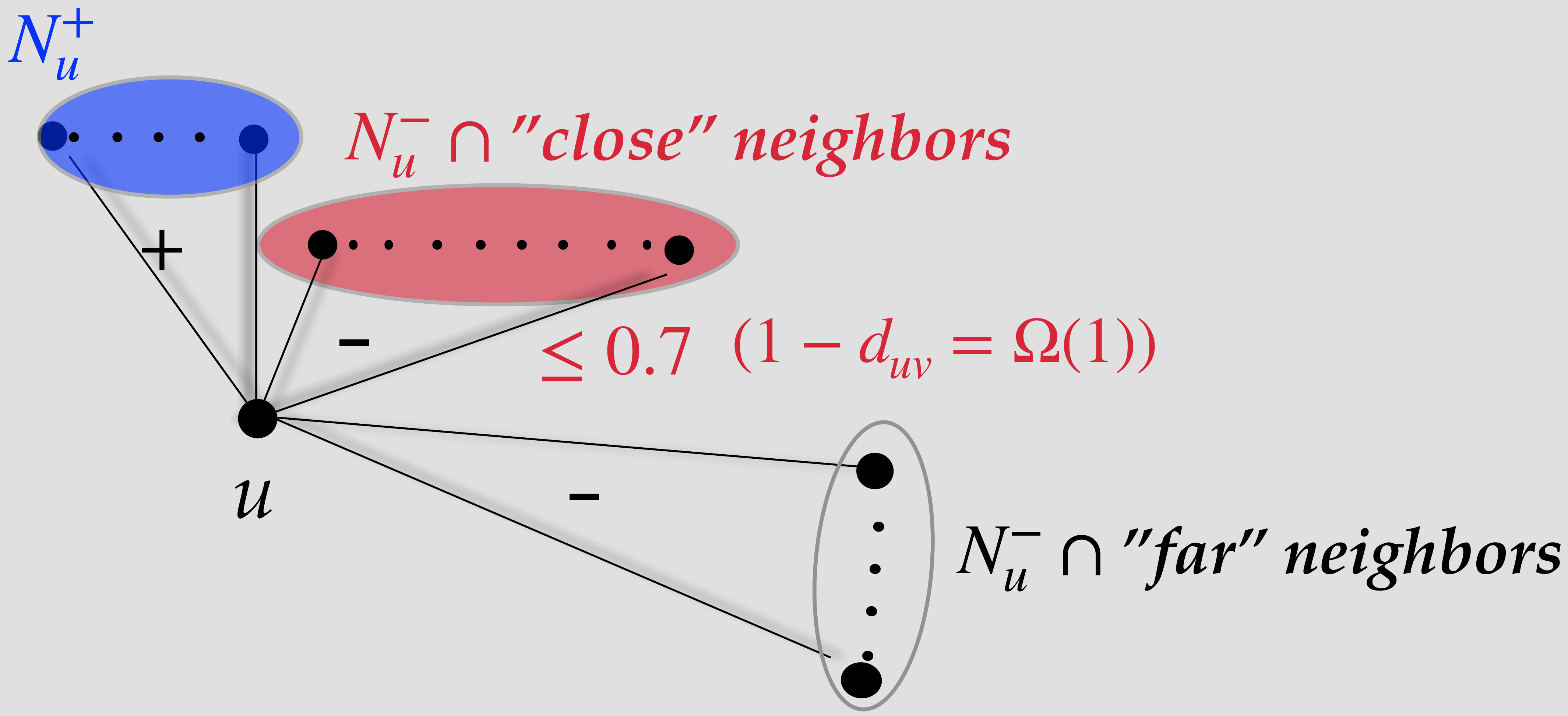
Number of subtleties to showing fractional ℓ_p cost of **adjusted correlation metric** is bounded against OPT_p for any p

If >> : round all edges out of u to 1 → pay $|N_u^+|^p$ for u

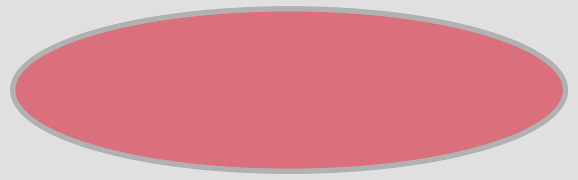
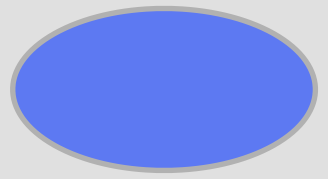
Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Key Idea 2: for ***positive edges***, correlation metric has *bounded cost for all p*!



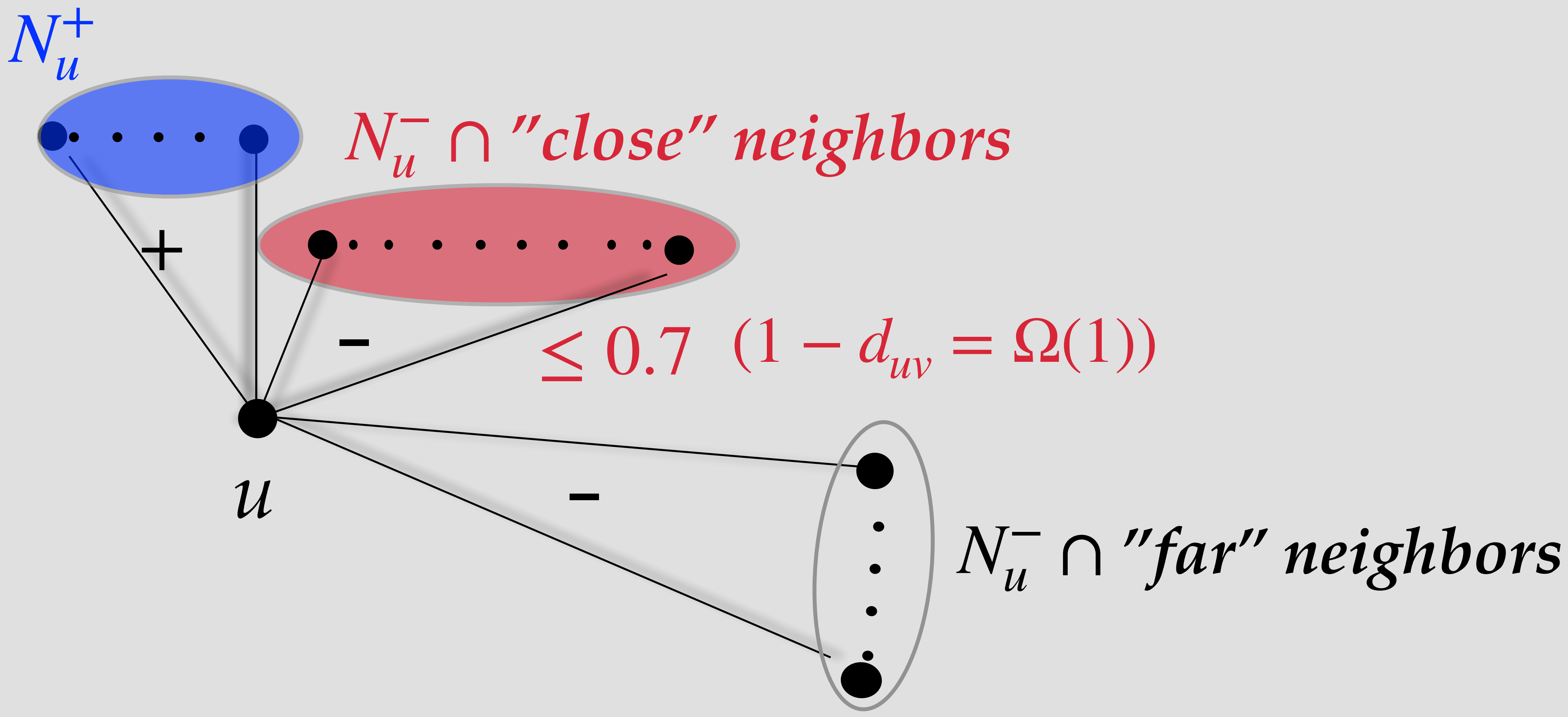
Number of subtleties to showing fractional ℓ_p cost of **adjusted correlation metric** is bounded against OPT_p for any p

If  >>  : round all edges out of u to 1 ➔ pay $|N_u^+|^p$ for u

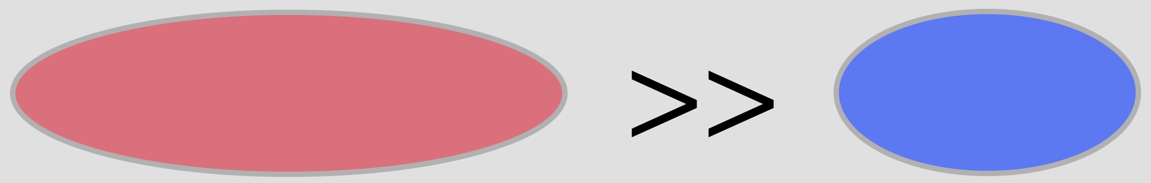
Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Key Idea 2: for ***positive edges***, correlation metric has *bounded cost for all p*!



Number of subtleties to showing fractional ℓ_p cost of **adjusted correlation metric** is bounded against OPT_p for any p

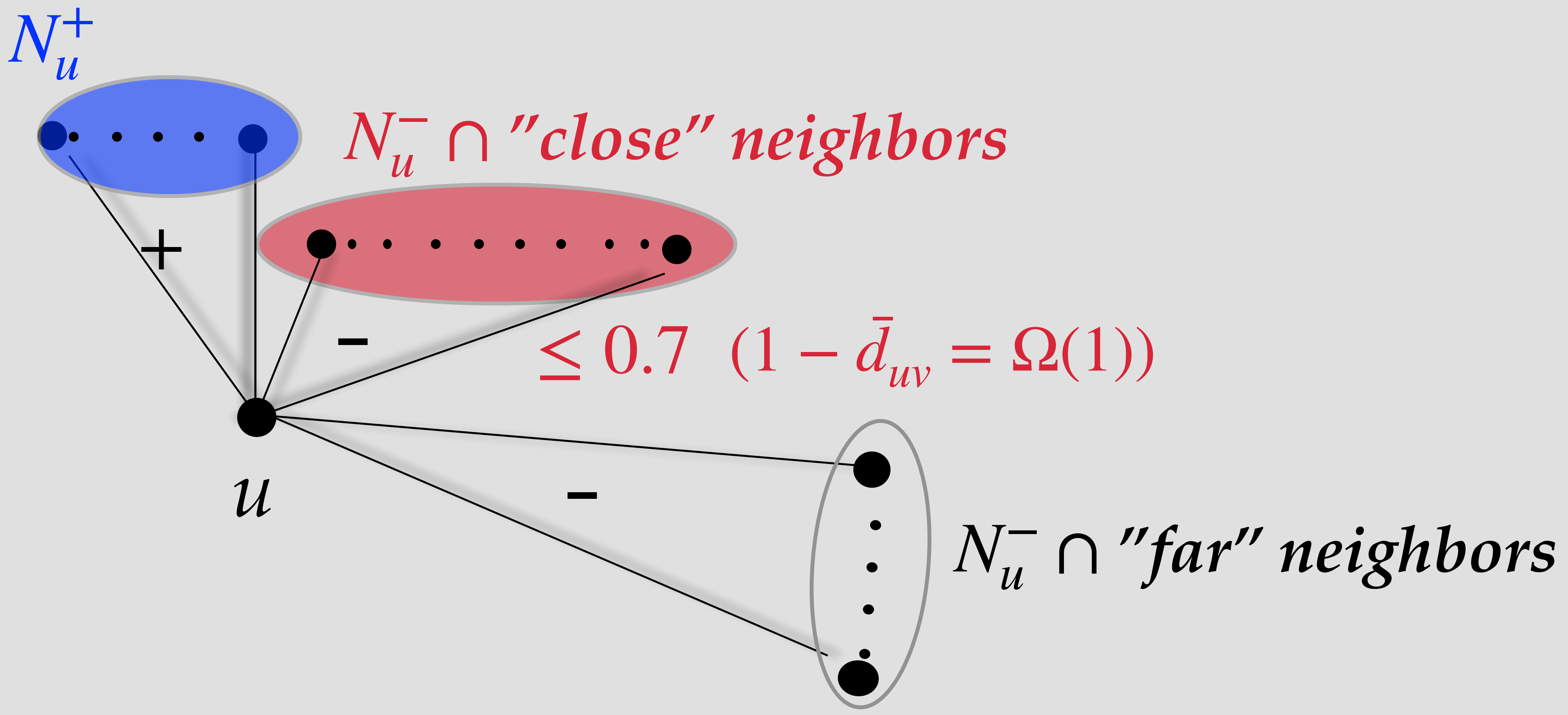


If [red ellipse] >> [blue ellipse] : round all edges out of u to 1 → pay $|N_u^+|^p$ for u

Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Key Idea 2: for ***positive edges***, correlation metric has *bounded cost for all p*!



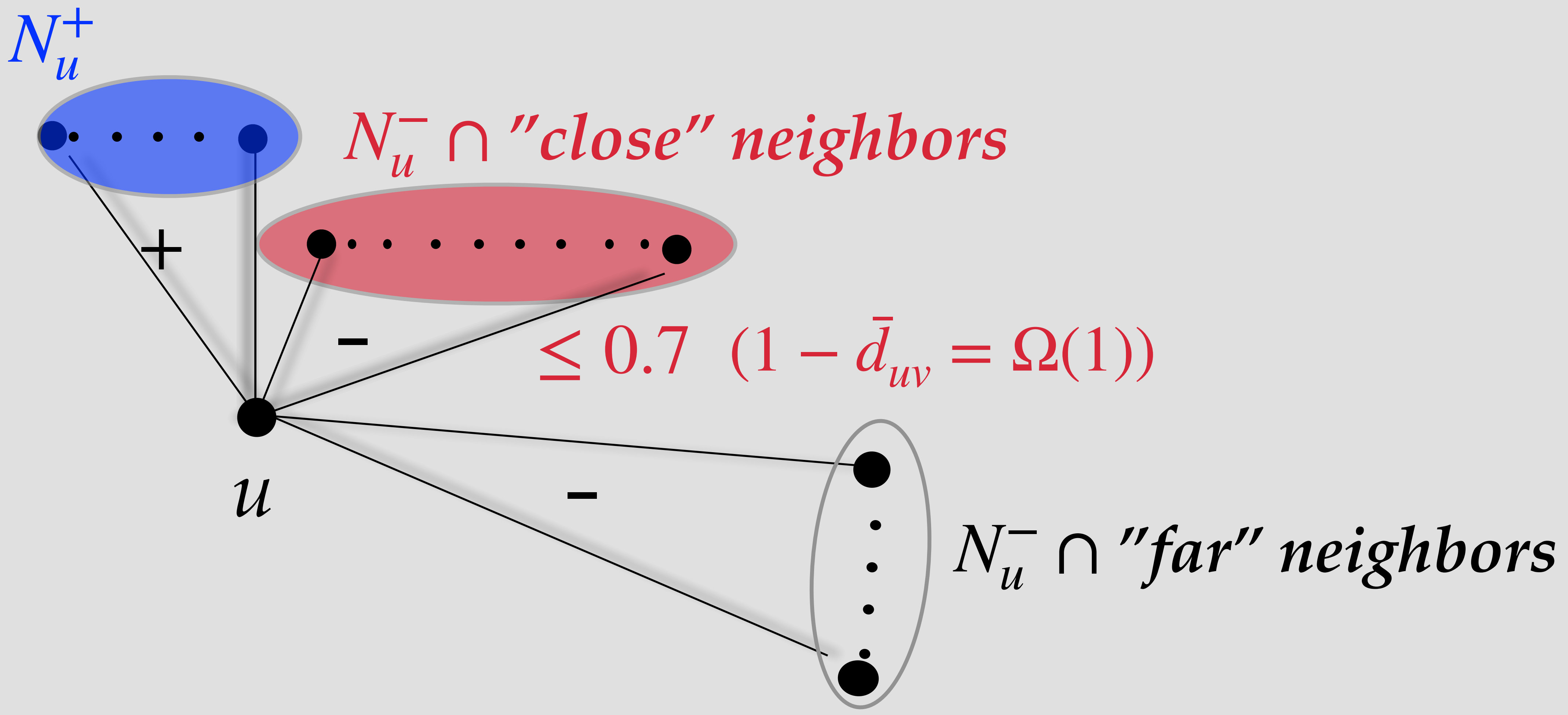
If [red] >> [blue] : round all edges out of u to 1 → pay $|N_u^+|^p$ for u

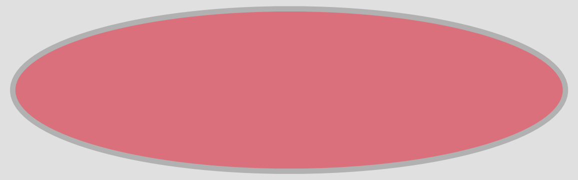
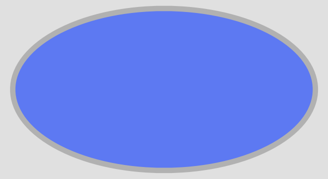
Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Key Idea 2: for ***positive edges***, correlation metric has *bounded cost for all p*!

Q1: when $p \neq 1$, rounding may be good from u's perspective, but what about v's?

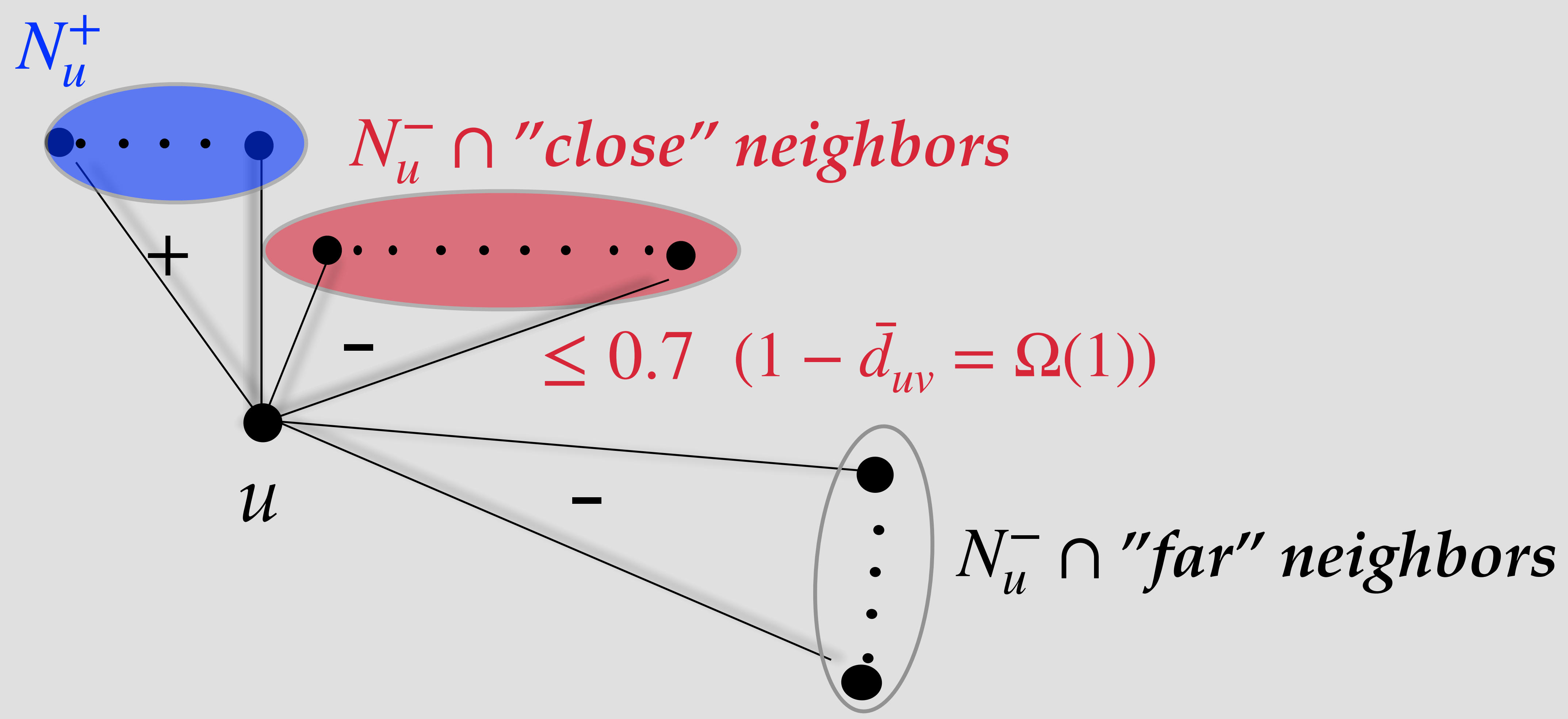


If  >>  : round all edges out of u to 1 → pay $|N_u^+|^p$ for u

Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Key Idea 2: for ***positive edges***, correlation metric has *bounded cost for all p*!



Q1: when $p \neq 1$, rounding may be good from u's perspective, but what about v's?

A1: reduce to regular case when $d_{uv} \leq 1/4$, say

↓

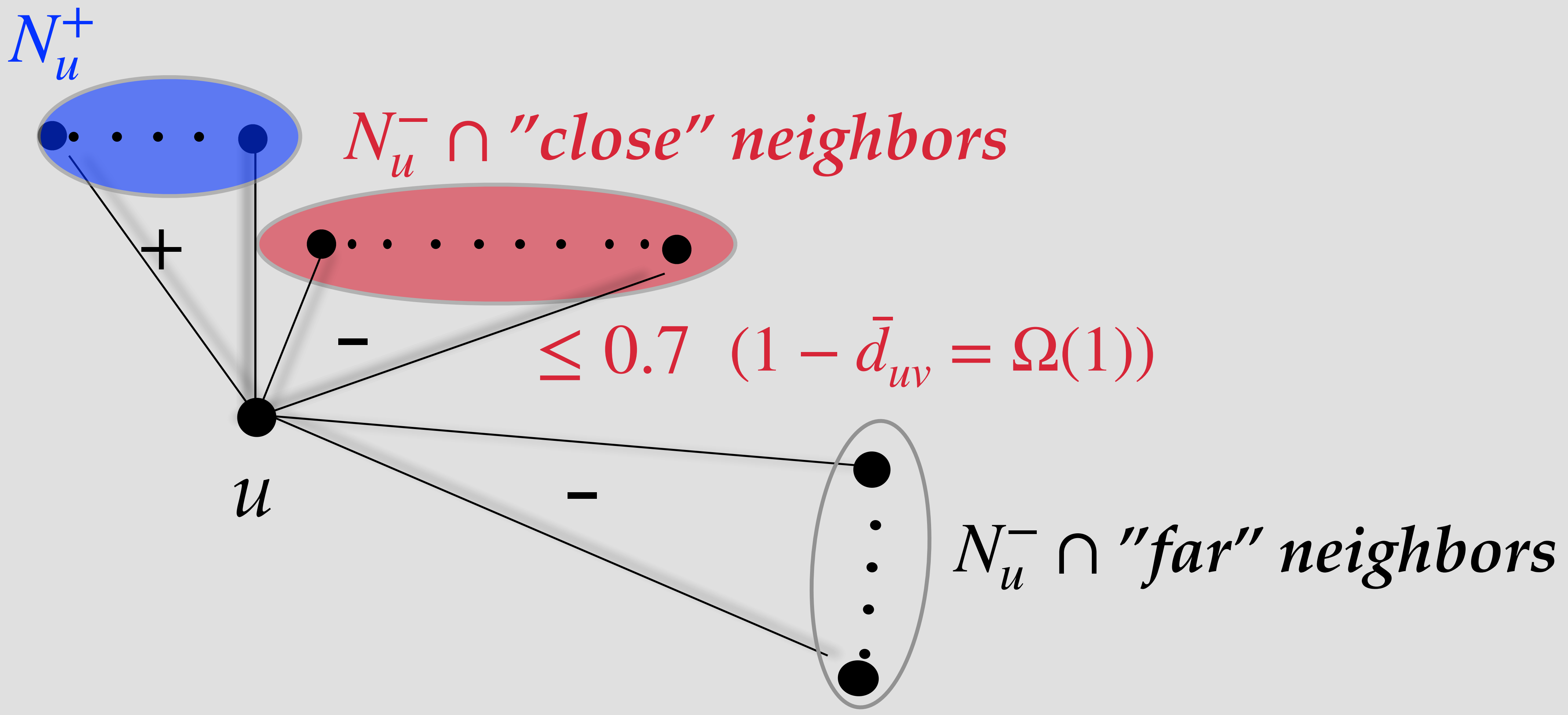
in some average sense, u can charge to v

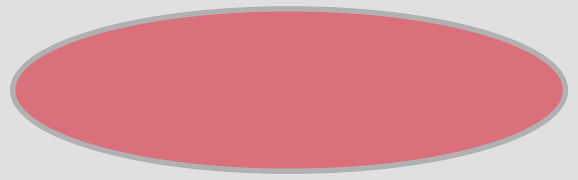
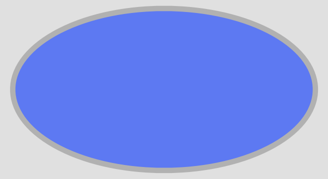
If (red) >> (blue) : round all edges out of u to 1 → pay $|N_u^+|^p$ for u

Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Key Observation 2: for ***positive edges***, correlation metric has *bounded cost for all p*!



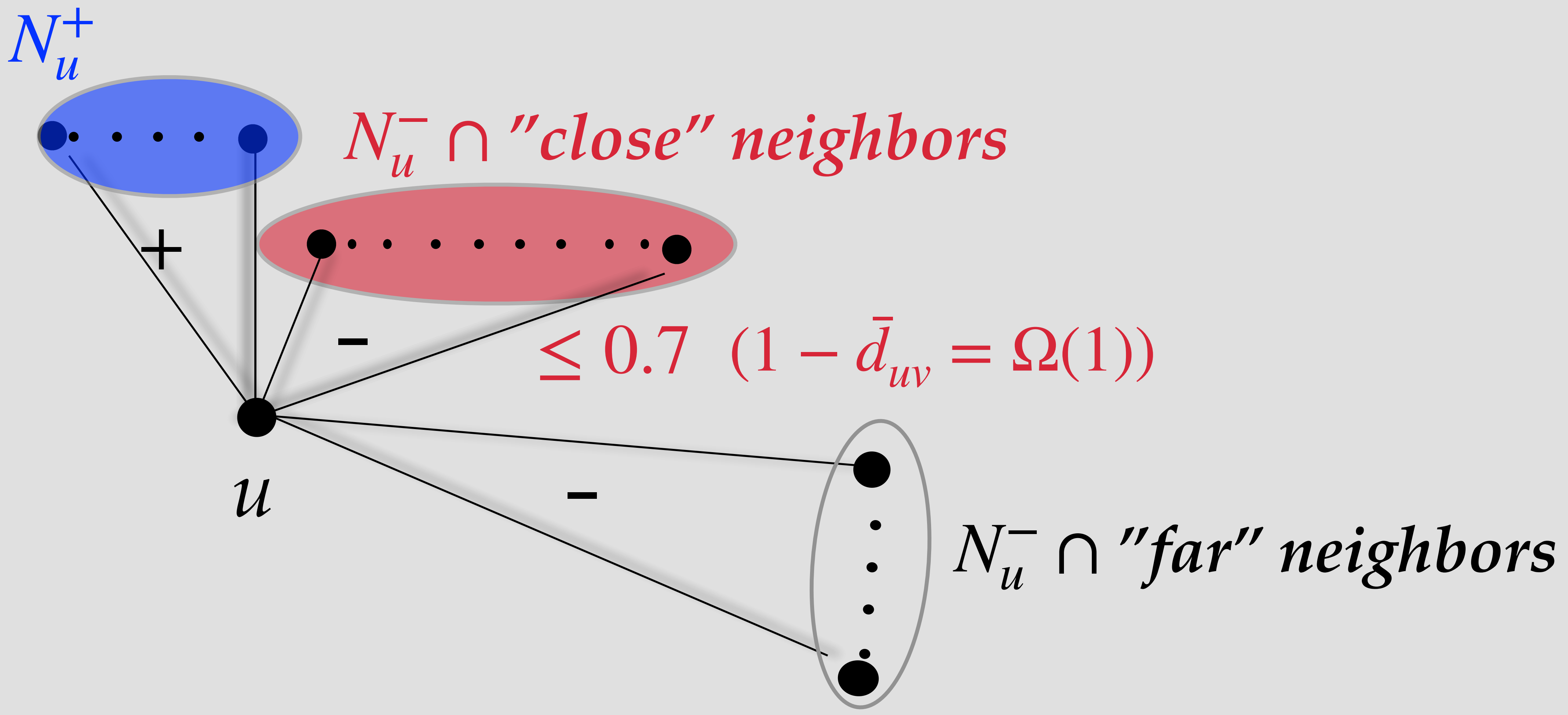
If  >>  : round all edges out of u to 1 ➜ pay $|N_u^+|^p$ for u

Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Key Observation 2: for ***positive edges***, correlation metric has *bounded cost for all p*!

Q2: how to pay for $|N_u^+|^p$ when you do round?
how to pay for (-) edges when you do not?

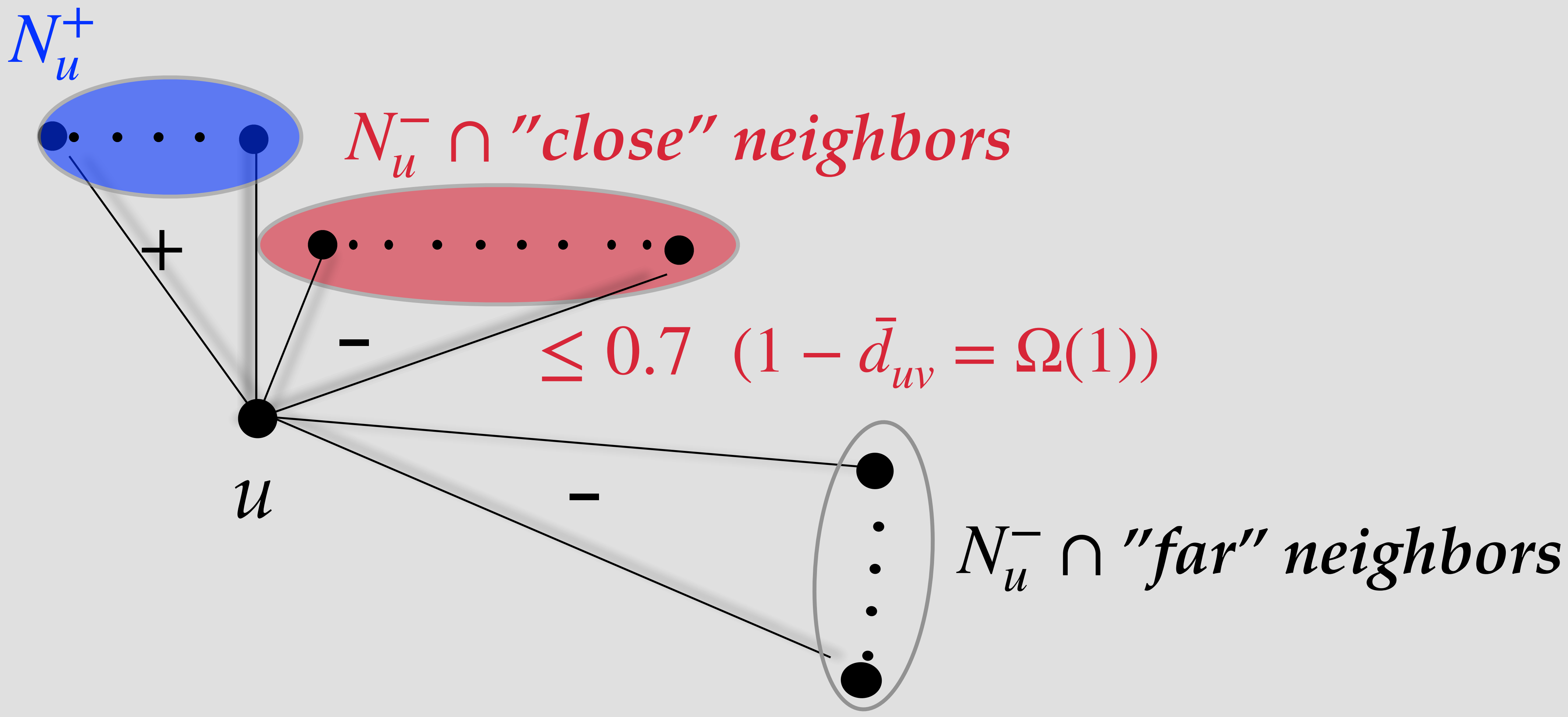


If [red ellipse] >> [blue ellipse] : round all edges out of u to 1 → pay $|N_u^+|^p$ for u

Adjusted correlation metric

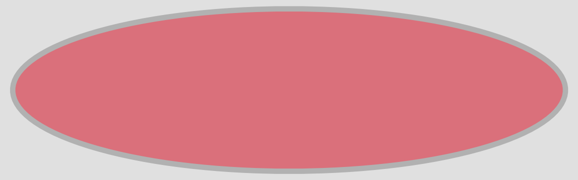
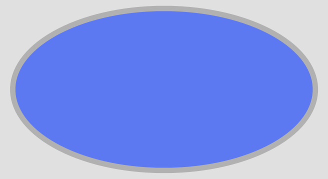
$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

Key Observation 2: for ***positive edges***, correlation metric has *bounded cost for all p*!



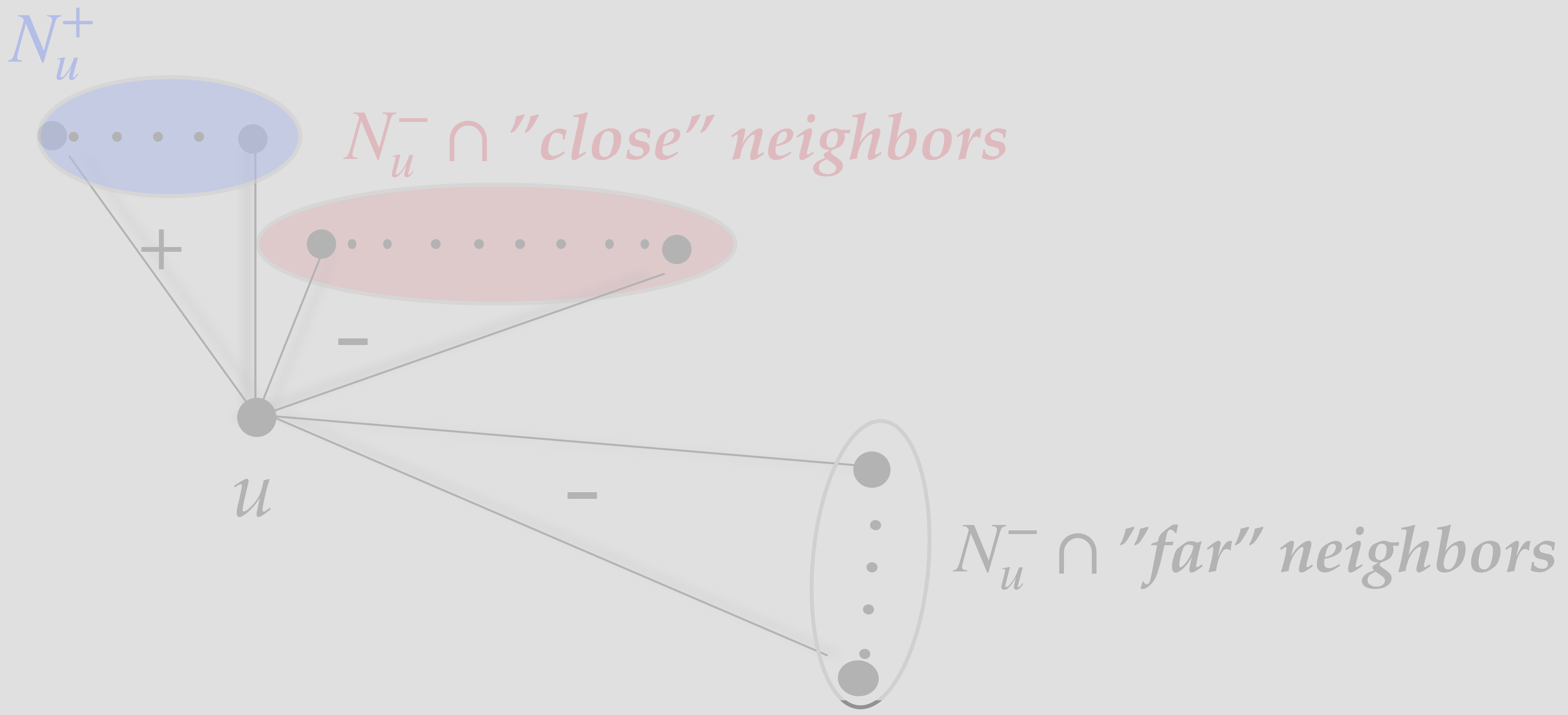
Q2: how to pay for $|N_u^+|^p$ when you do round?
how to pay for (-) edges when you do not?

A2: non-local charging arguments

If  >>  : round all edges out of u to 1 ➜ pay $|N_u^+|^p$ for u

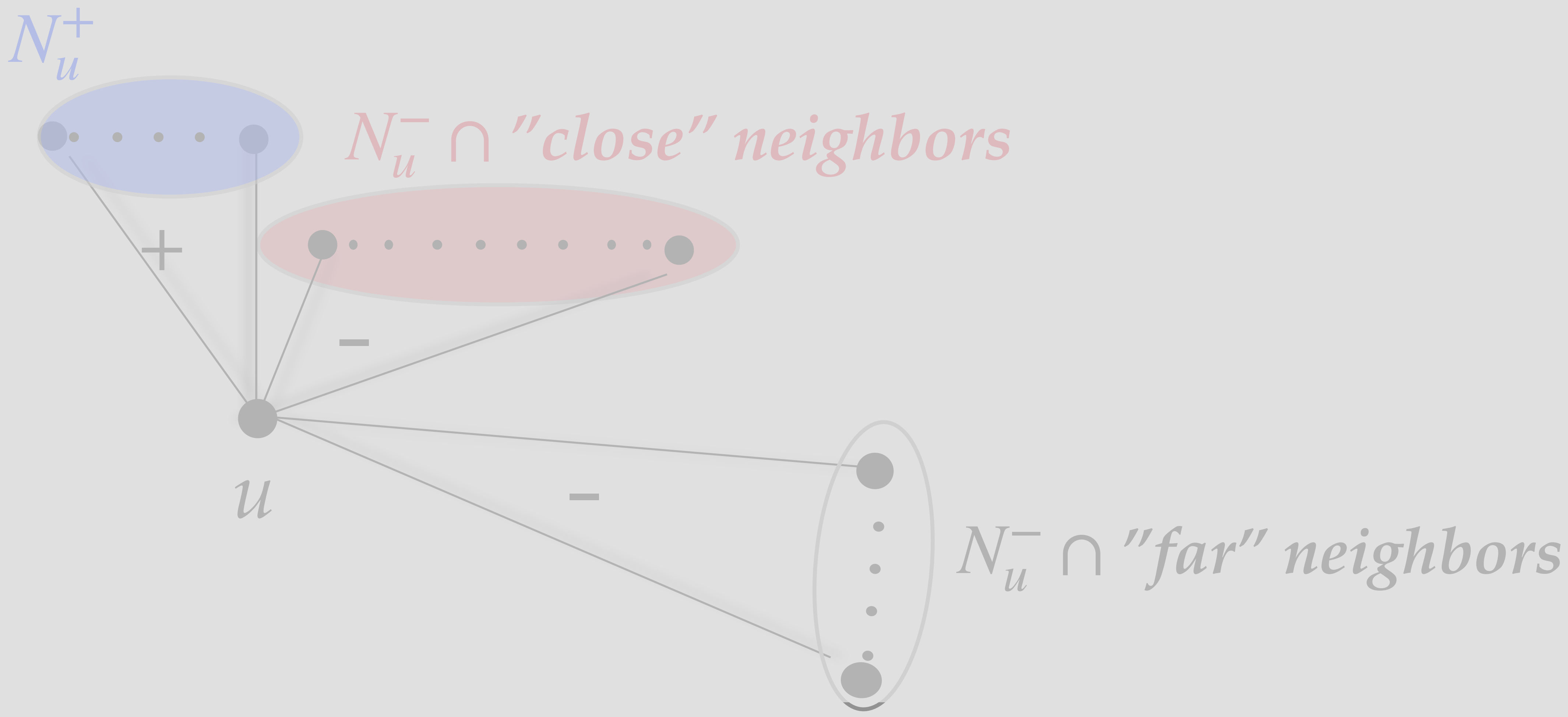
Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$



Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$



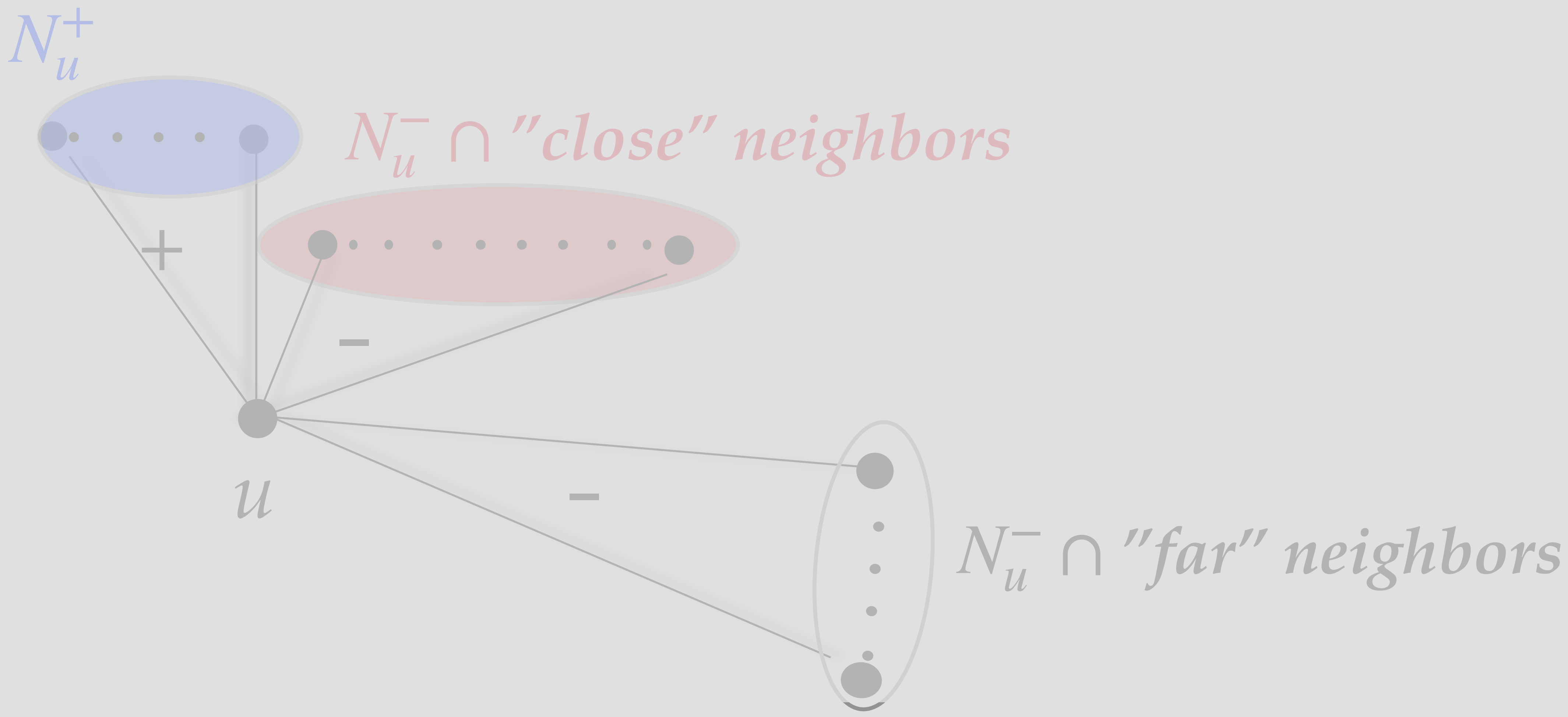
Adjusted correlation metric → Rounding algorithm by Kalhan, Makarychev, Zhou

If >> round all d_{uv} to 1 → u now in own cluster

Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$

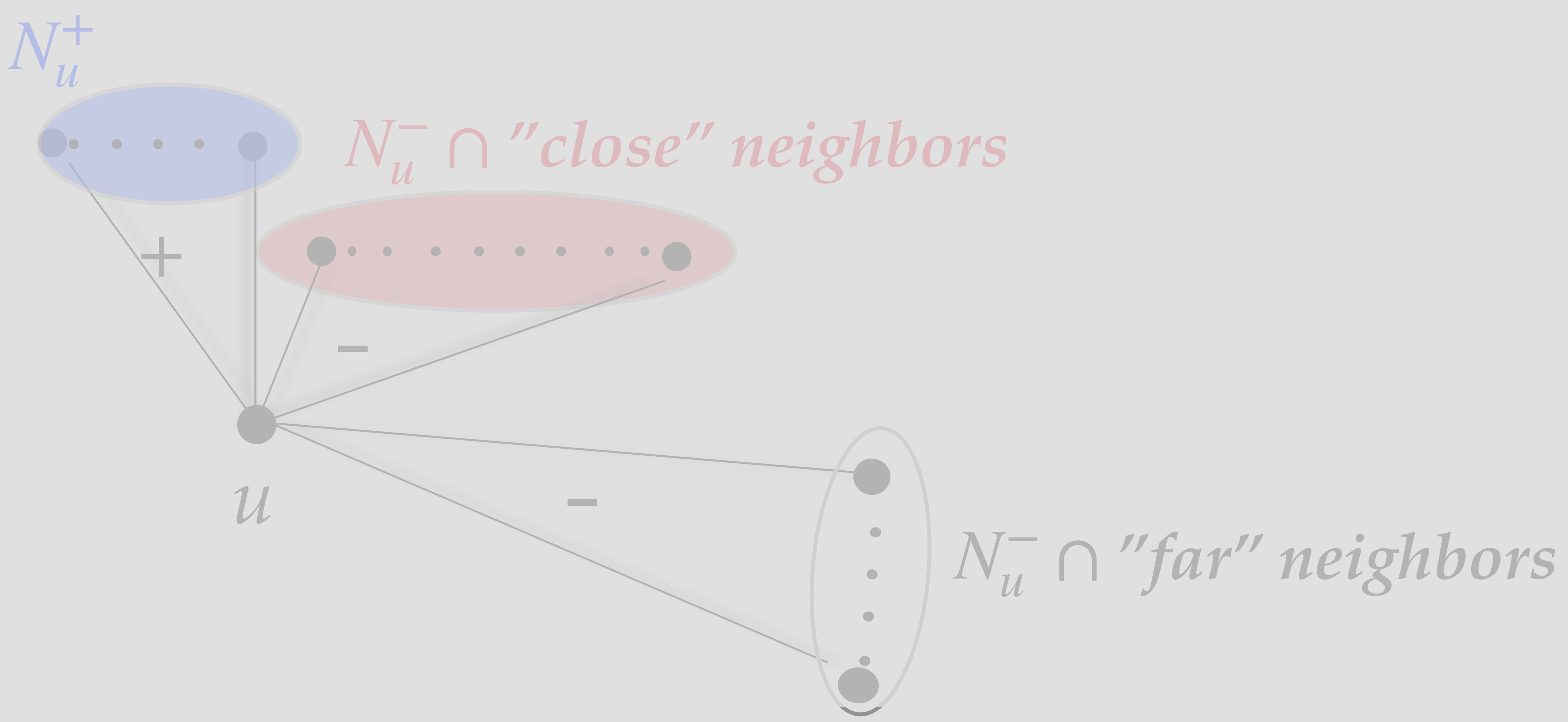
All steps independent of p

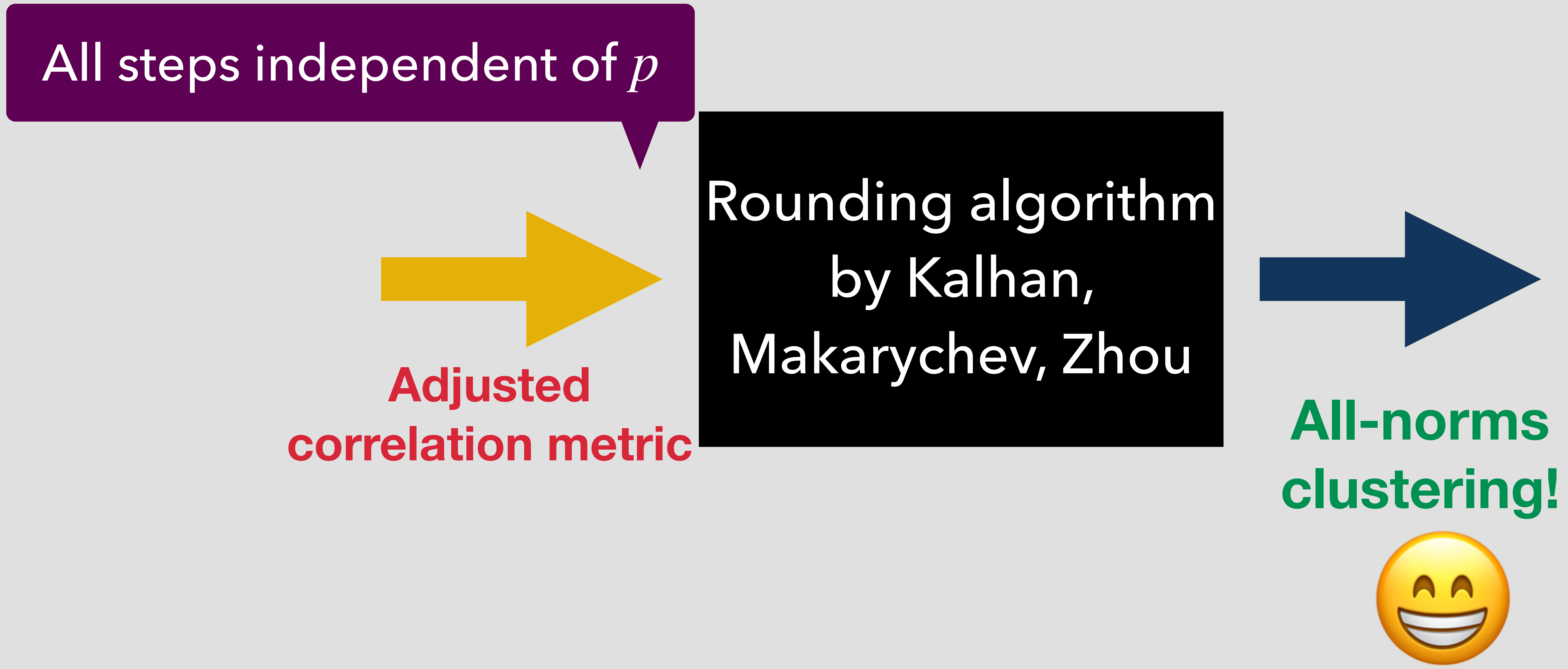


If >> round all d_{uv} to 1 → u now in own cluster

Adjusted correlation metric

$$d_{uv} = \frac{|N_u^+ \cap N_v^-| + |N_u^- \cap N_v^+|}{|N_u^+ \cup N_v^+|}$$





Today

- ✦ ~~The **correlation metric** (constructing a "guess" for the optimal solution to convex relaxation)~~
- ✦ ~~Tweaking correlation metric for **all** ℓ_p norms~~
- ✦ Summary and open questions

Summary

Summary

ℓ_p-norm correlation clustering algs solve a convex program

- Solving *metric constrained* LPs on large networks is slow!
- Not very amenable to online, streaming
- Solution specific to *one fixed* ℓ_p-norm

Summary

ℓ_p-norm correlation clustering algs solve a convex program

Solving *metric constrained* LPs on large networks is slow!

Not very amenable to online, streaming

Solution specific to *one fixed* ℓ_p-norm

Combinatorial techniques can resolve these issues

Summary

Summary

Result 1: $O(1)$-apx alg with run-time $O(min\{\, n{\cdot}\Delta^2{\cdot}log\, n\, ,\, n^{\omega}\})$. Near-linear for sparse graphs.

Summary

First combinatorial alg for $p > 1$

Result 1: $O(1)$-apx alg with run-time $O(min\{ n \cdot \Delta^2 \cdot log\ n\ ,\ n^\omega \})$. Near-linear for sparse graphs.

Summary

First combinatorial alg for $p > 1$

Δ = max (+) degree of any vertex
ω = matrix multiplication exponent

Result 1: $O(1)$-apx alg with run-time $O(\min\{ n \cdot \Delta^2 \cdot \log n , n^{\omega}\})$. Near-linear for sparse graphs.

Summary

First combinatorial alg for $p > 1$

Δ = max (+) degree of any vertex
ω = matrix multiplication exponent

Result 1: $O(1)$-apx alg with run-time $O(\min\{ n \cdot \Delta^2 \cdot \log n , n^{\omega}\})$. Near-linear for sparse graphs.

Result 2: ∃ an alg producing a clustering that is $O(1)$-apx for all ℓ_p-norms, simultaneously.

Summary

First combinatorial alg for $p > 1$

Δ = max (+) degree of any vertex
ω = matrix multiplication exponent

Result 1: $O(1)$-apx alg with run-time $O(\min\{ n \cdot \Delta^2 \cdot \log n , n^\omega\})$. Near-linear for sparse graphs.

Result 2: ∃ an alg producing a clustering that is $O(1)$-apx for all ℓ_p-norms, simultaneously.

Sometimes called *universality* property

Summary

First combinatorial alg for $p > 1$

Δ = max (+) degree of any vertex
ω = matrix multiplication exponent

Result 1: $O(1)$-apx alg with run-time $O(\min\{ n \cdot \Delta^2 \cdot \log n , n^{\omega}\})$. Near-linear for sparse graphs.

Result 2: ∃ an alg producing a clustering that is $O(1)$-apx for all ℓ_p-norms, simultaneously.

Sometimes called *universality* property

Summary

First combinatorial alg for $p > 1$

Δ = max (+) degree of any vertex
ω = matrix multiplication exponent

Result 1: $O(1)$-apx alg with run-time $O(\min\{ n \cdot \Delta^2 \cdot \log n , n^\omega\})$. Near-linear for sparse graphs.

Result 2: ∃ an alg producing a clustering that is $O(1)$-apx for all ℓ_p-norms, simultaneously.

Sometimes called *universality* property

Correlation clustering has interesting combinatorial structure that can be exploited

What's next?

What's next?

What's next?

- 🔥**Hot conjecture:** Exists a combinatorial alg simultaneously 4-approximating all ℓ_p-norms running in $O(n^\omega)$ time

What's next?

- 🔥**Hot conjecture:** Exists a combinatorial alg simultaneously 4-approximating all ℓ_p-norms running in $O(n^\omega)$ time
- 😶‍🌫️ **Broader Qs:**

What's next?

- 🔥**Hot conjecture:** Exists a combinatorial alg simultaneously 4-approximating all ℓ_p-norms running in $O(n^\omega)$ time
- 😶‍🌫️ **Broader Qs:**
    1. Combinatorial algorithms by designing "approximate LP solution"

What's next?

- 🔥**Hot conjecture:** Exists a combinatorial alg simultaneously 4-approximating all ℓ_p-norms running in $O(n^\omega)$ time
- 😶‍🌫️ **Broader Qs:**
  1. Combinatorial algorithms by designing "approximate LP solution"
  2. Non-constructive existence proof?

What's next?

- 🔥**Hot conjecture:** Exists a combinatorial alg simultaneously 4-approximating all ℓ_p-norms running in $O(n^\omega)$ time
- 😶‍🌫️ **Broader Qs:**
  1. Combinatorial algorithms by designing "approximate LP solution"
  2. Non-constructive existence proof?
  3. Further study on the all-norms objective

What's next?

- 🔥**Hot conjecture:** Exists a combinatorial alg simultaneously 4-approximating all ℓ_p-norms running in $O(n^\omega)$ time
- 😶‍🌫️ **Broader Qs:**
  1. Combinatorial algorithms by designing "approximate LP solution"
  2. Non-constructive existence proof?
  3. Further study on the all-norms objective

Thank you!

Thank you!

hanewman@andrew.cmu.edu