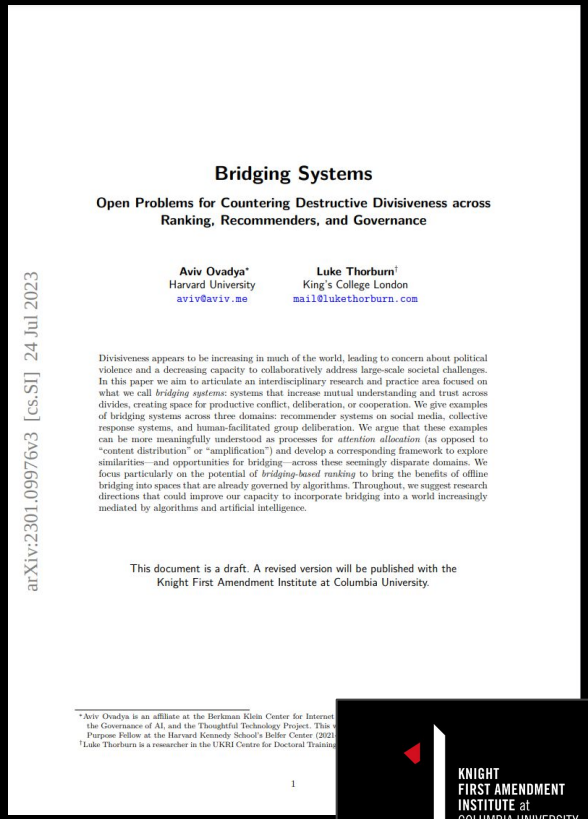
→ bridging.systems



Bridging Systems

Open Problems for Countering Destructive Divisiveness across Ranking, Recommenders, and Governance

Aviv Ovadya[*]
Harvard University
aviv@aviv.me

Luke Thorburn[†]
King's College London
mail@lukethorburn.com

Divisiveness appears to be increasing in much of the world, leading to concern about political violence and a decreasing capacity to collaboratively address large-scale societal challenges. In this paper we aim to articulate an interdisciplinary research and practice area focused on what we call *bridging systems*: systems that increase mutual understanding and trust across divides, creating space for productive conflict, deliberation, or cooperation. We give examples of bridging systems across three domains: recommender systems on social media, collective response systems, and human-facilitated group deliberation. We argue that these examples can be more meaningfully understood as processes for *attention allocation* (as opposed to "content distribution" or "amplification") and develop a corresponding framework to explore similarities—and opportunities for bridging—across these seemingly disparate domains. We focus particularly on the potential of *bridging-based ranking* to bring the benefits of offline bridging into spaces that are already governed by algorithms. Throughout, we suggest research directions that could improve our capacity to incorporate bridging into a world increasingly mediated by algorithms and artificial intelligence.

This document is a draft. A revised version will be published with the Knight First Amendment Institute at Columbia University.

[*]Aviv Ovadya is an affiliate at the Berkman Klein Center for Internet
the Governance of AI, and the Thoughtful Technology Project. This
Purpose Fellow at the Harvard Kennedy School's Belfer Center (2021
[†]Luke Thorburn is a researcher in the UKRI Centre for Doctoral Training

1

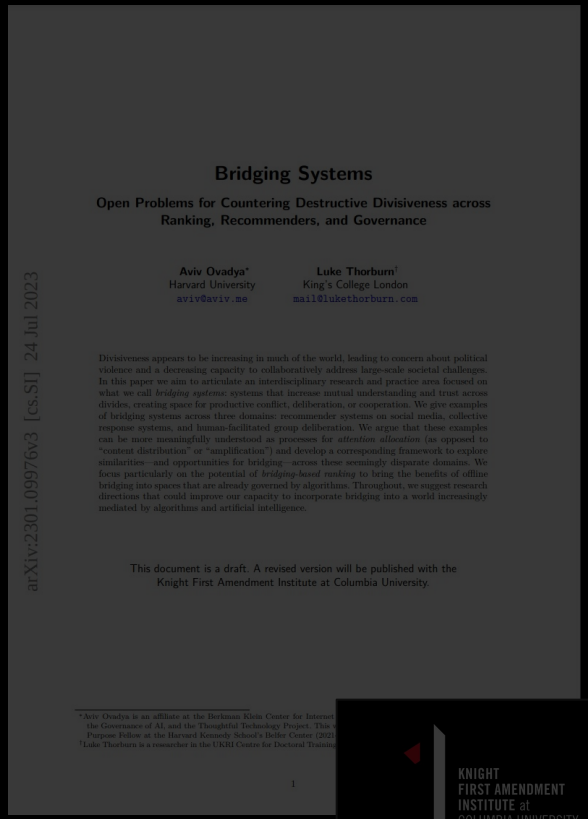
KNIGHT FIRST AMENDMENT INSTITUTE at COLUMBIA UNIVERSITY

Probabilistic Foundations of Partisan (Un)Sorting

Luke Thorburn

October 2023

→ bridging.systems



Bridging Systems

Open Problems for Countering Destructive Divisiveness across Ranking, Recommenders, and Governance

Aviv Ovadya*
Harvard University
aviv@aviv.me

Luke Thorburn†
King's College London
mail@lukethorburn.com

Divisiveness appears to be increasing in much of the world, leading to concern about political violence and a decreasing capacity to collaboratively address large-scale societal challenges. In this paper we aim to articulate an interdisciplinary research and practice area focused on what we call *bridging systems*: systems that increase mutual understanding and trust across divides, creating space for productive conflict, deliberation, or cooperation. We give examples of bridging systems across three domains: recommender systems on social media, collective response systems, and human-facilitated group deliberation. We argue that these examples can be more meaningfully understood as processes for *attention allocation* (as opposed to "content distribution" or "amplification") and develop a corresponding framework to explore similarities—and opportunities for bridging—across these seemingly disparate domains. We focus particularly on the potential of *bridging-based ranking* to bring the benefits of offline bridging into spaces that are already governed by algorithms. Throughout, we suggest research directions that could improve our capacity to incorporate bridging into a world increasingly mediated by algorithms and artificial intelligence.

This document is a draft. A revised version will be published with the Knight First Amendment Institute at Columbia University.

*Aviv Ovadya is an affiliate at the Berkman Klein Center for Internet
the Governance of AI, and the Thoughtful Technology Project. This
Purpose Fellow at the Harvard Kennedy School's Belfer Center (2021
†Luke Thorburn is a researcher in the UKRI Centre for Doctoral Training

1

KNIGHT FIRST AMENDMENT INSTITUTE at COLUMBIA UNIVERSITY

Probabilistic Foundations of Partisan (Un)Sorting

Luke Thorburn

October 2023

Motivation

How to quantify bridging?

Of all the ways to operationalize bridging, why is *diverse approval* so common/successful?

How to ground technical approaches to bridging in political, democratic, and peacebuilding theory?

I. Sortedness

II. Why unsort?

III. How to unsort?

IV. Can we unsort?

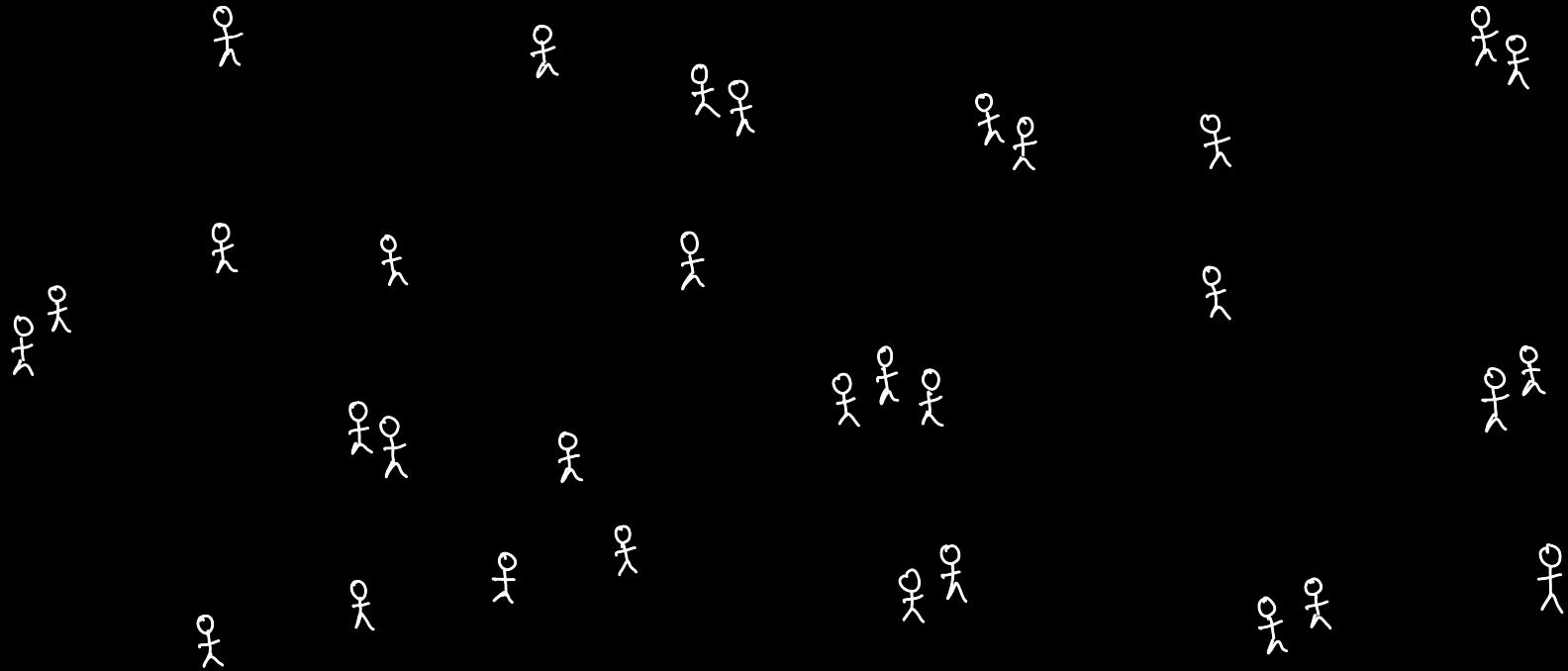
I. Sortedness

II. Why unsort?

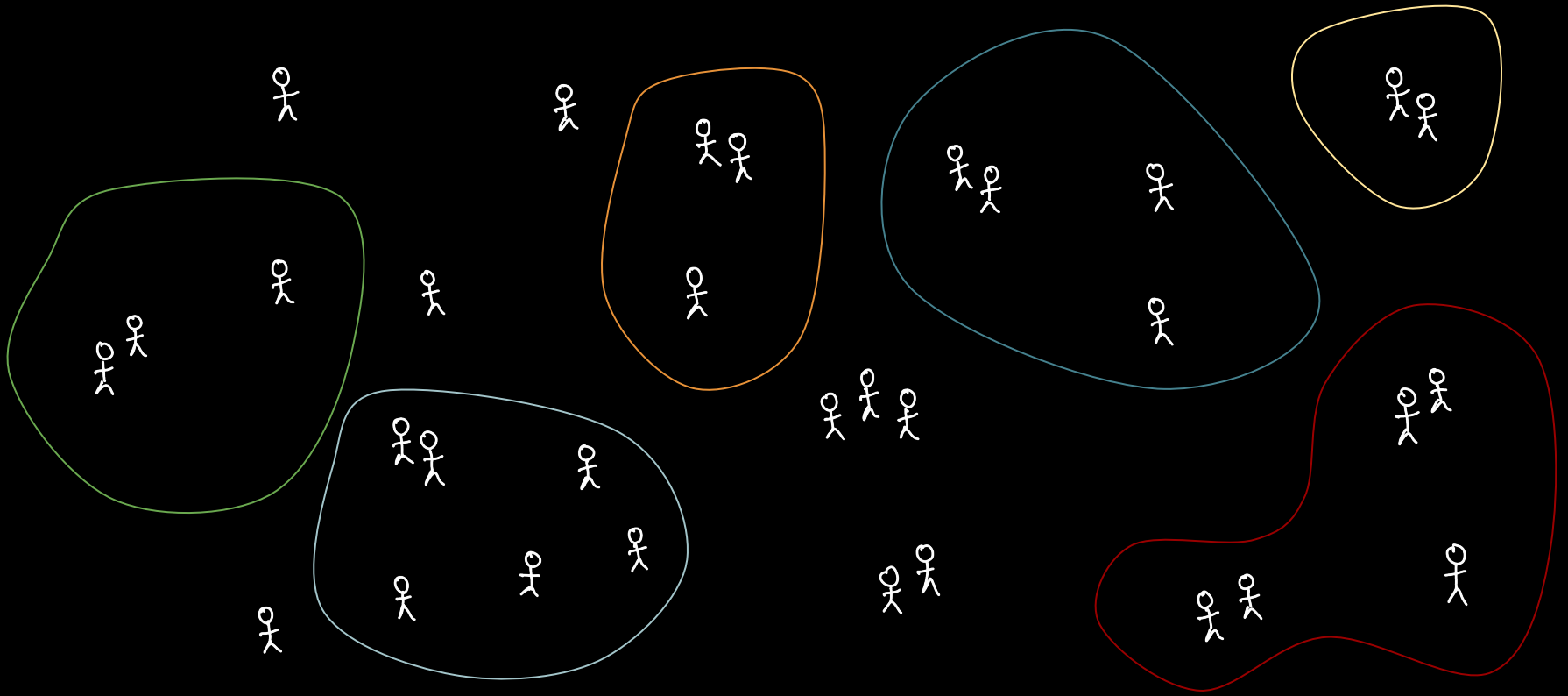
III. How to unsort?

IV. Can we unsort?

Societies are hypergraphs (intersecting groups).

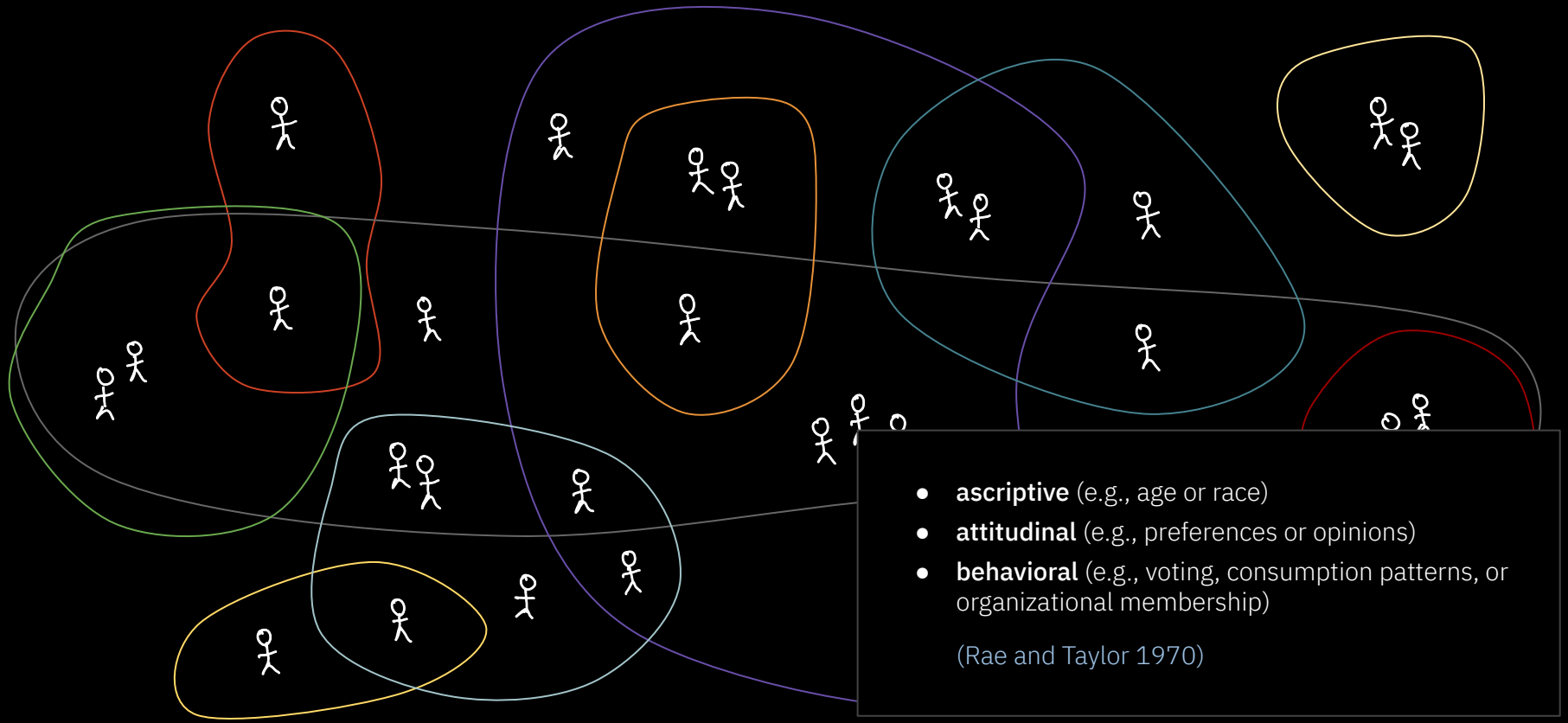


Societies are hypergraphs (intersecting groups).



Societies are hypergraphs (intersecting groups).

Societies are hypergraphs (intersecting groups).



- **ascriptive** (e.g., age or race)
- **attitudinal** (e.g., preferences or opinions)
- **behavioral** (e.g., voting, consumption patterns, or organizational membership)

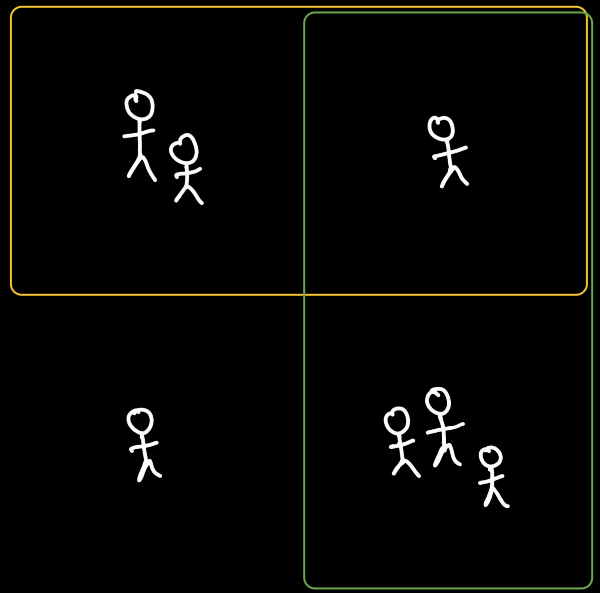
(Rae and Taylor 1970)

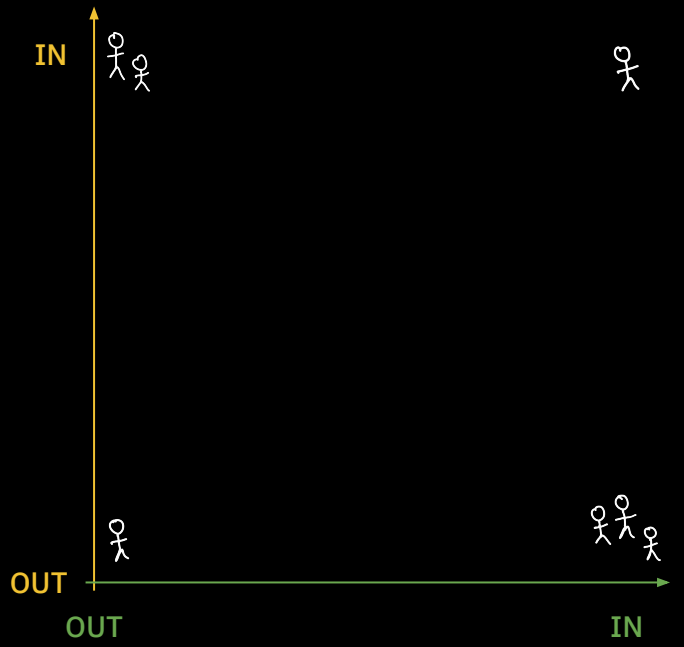
Societies are hypergraphs (intersecting groups).

How to quantify the overall amount of crosscutting?

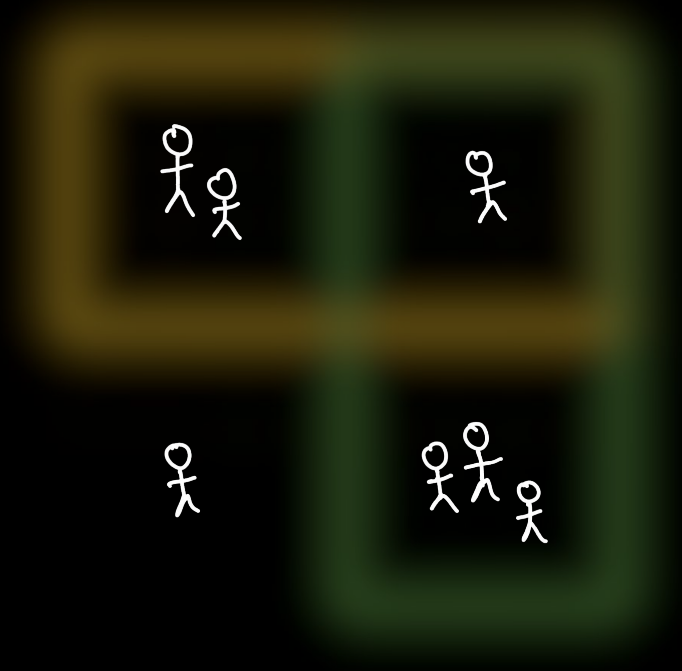
How to quantify the overall amount of crosscutting?

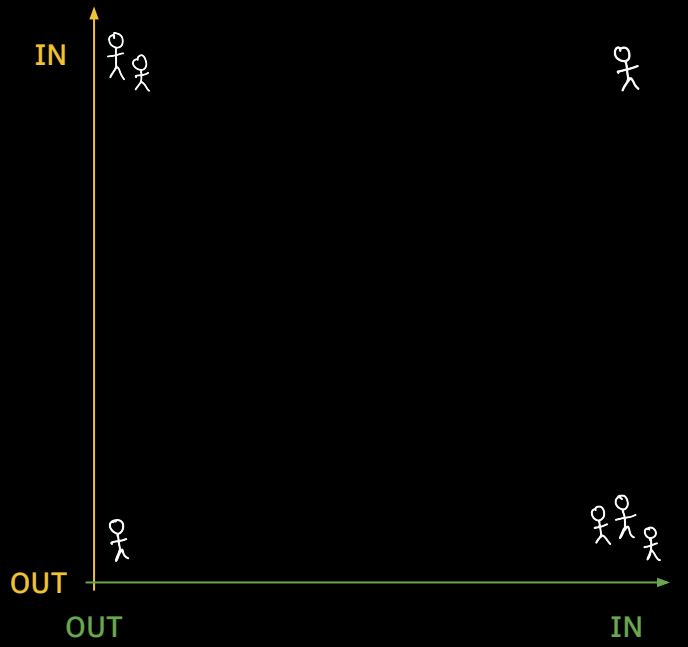
There is a correspondence between hypergraphs and distributions.



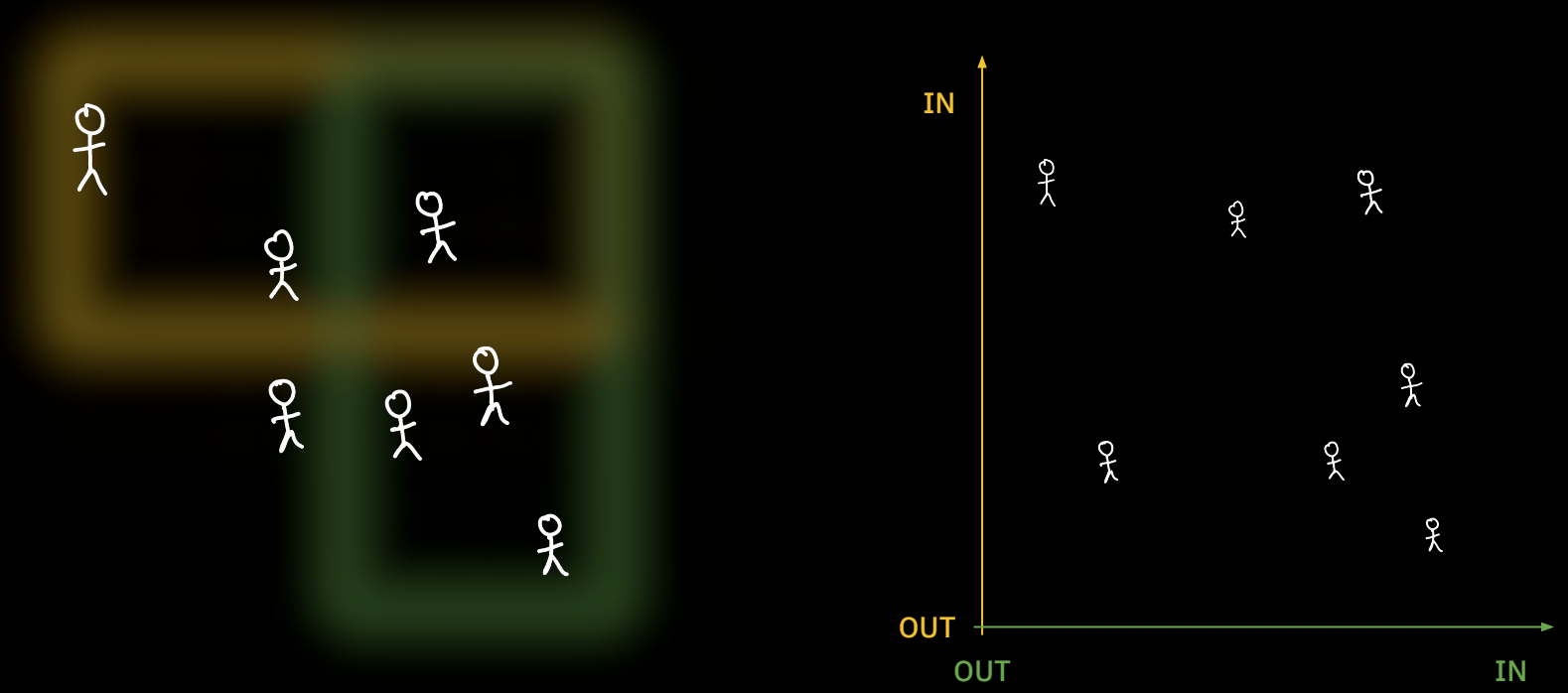


There is a correspondence between hypergraphs and distributions.

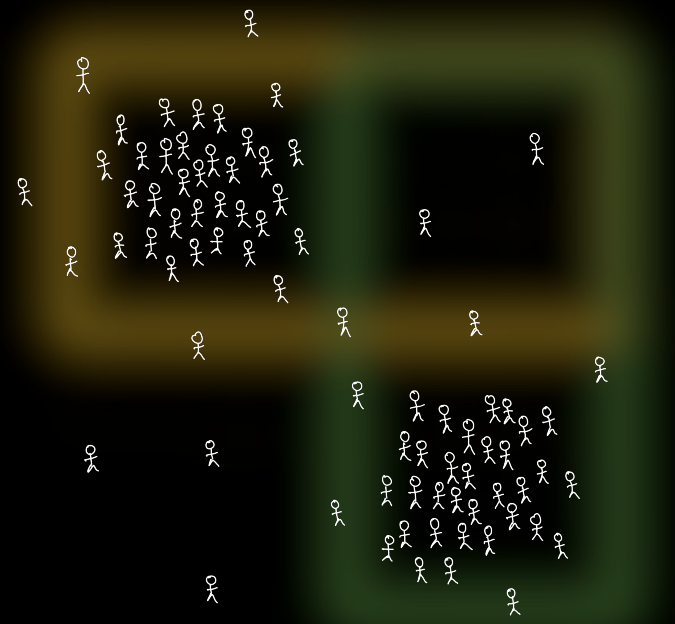


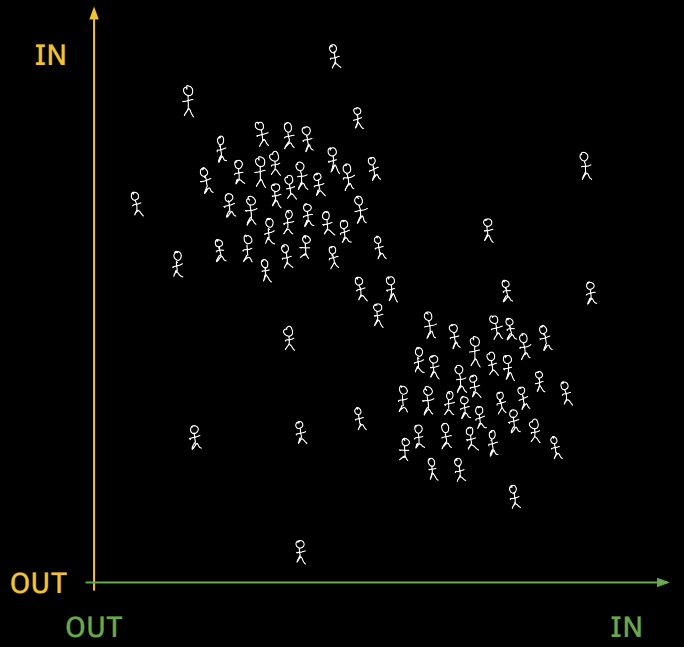


There is a correspondence between hypergraphs and distributions.

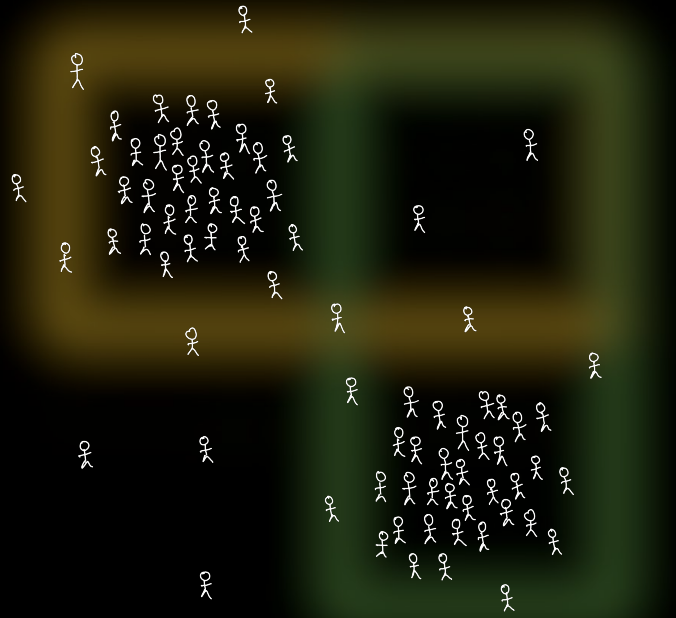


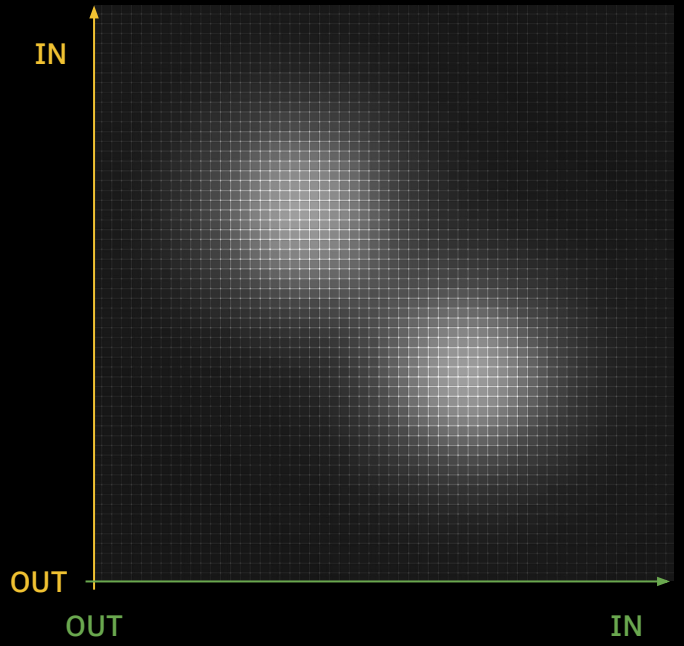
There is a correspondence between hypergraphs and distributions.



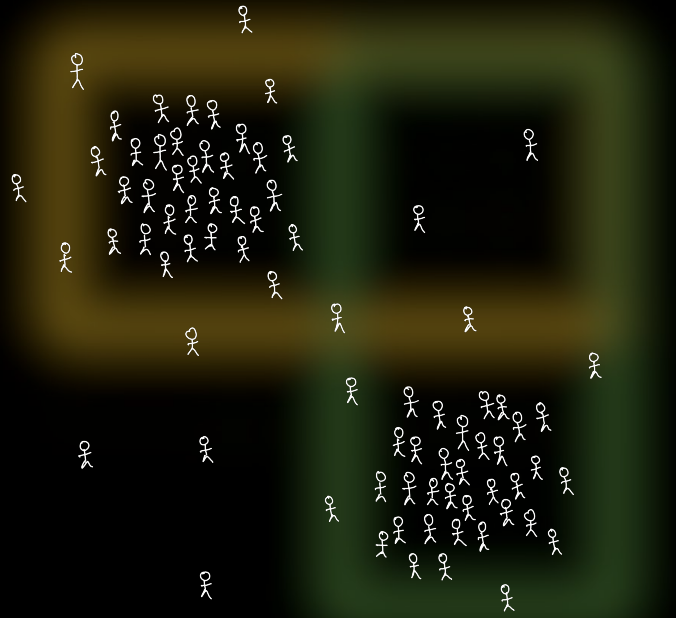


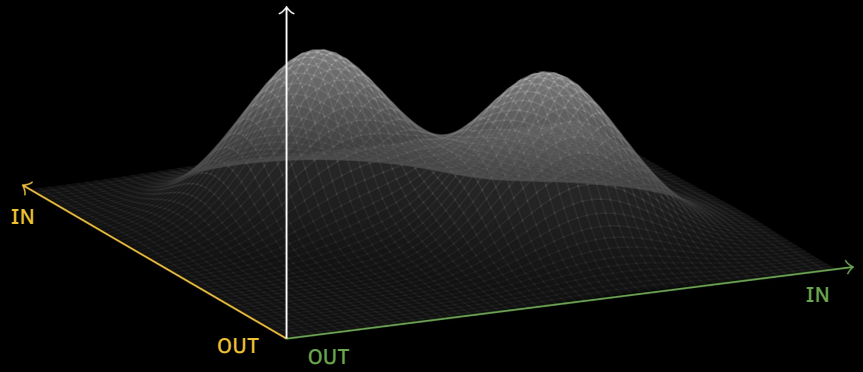
There is a correspondence between hypergraphs and distributions.





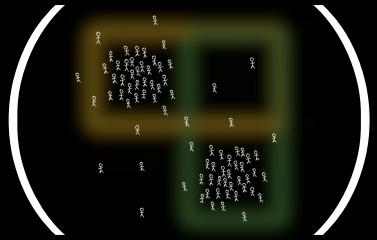
There is a correspondence between hypergraphs and distributions.

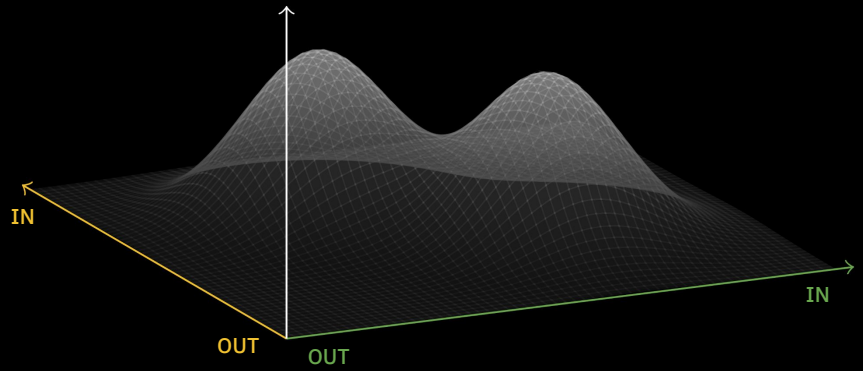




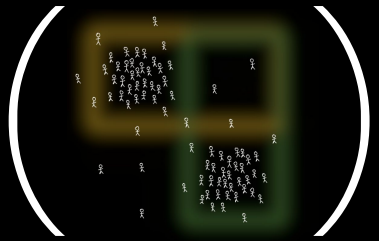
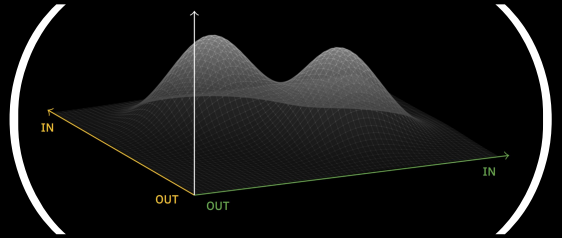
Crosscutting ∝ Dependence

Crosscuttingness

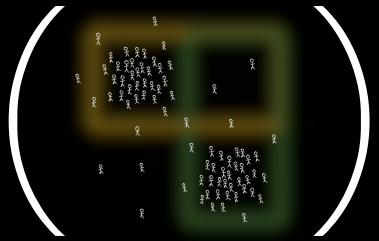
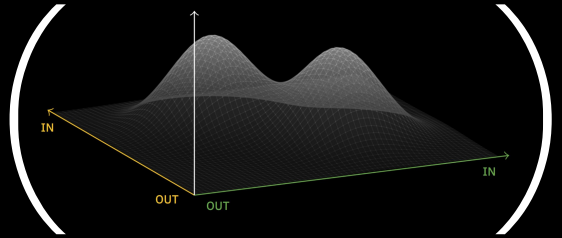


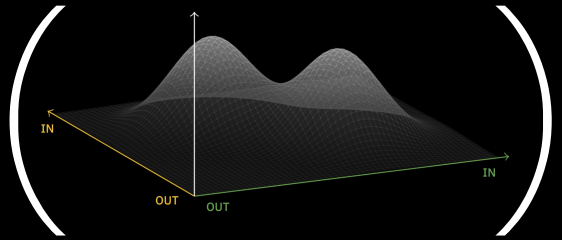


Crosscutting ∝ Dependence

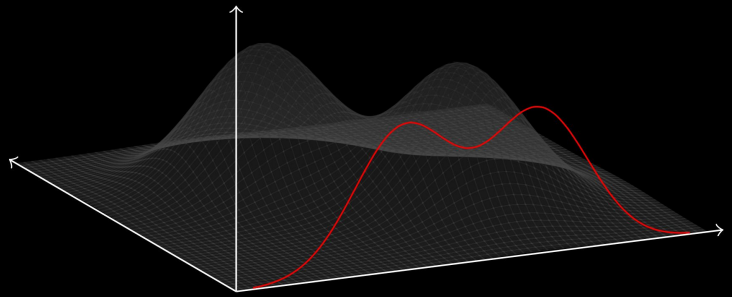
Crosscuttingness  ∝ Dependence 

Crosscutting ∝ Dependence = Sorting

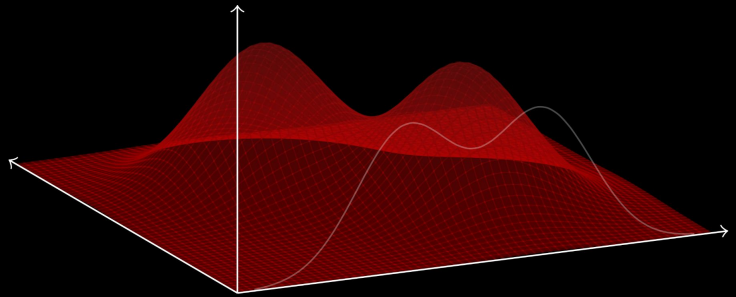
Crosscuttingness  ∝ Dependence 

= Sorting 

What is sorting?

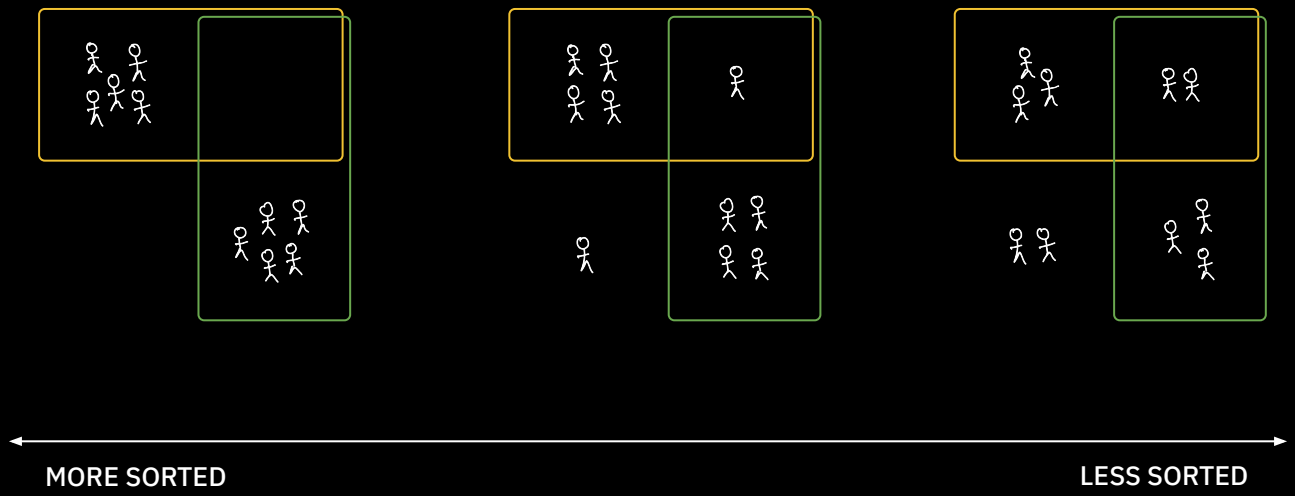


polarization

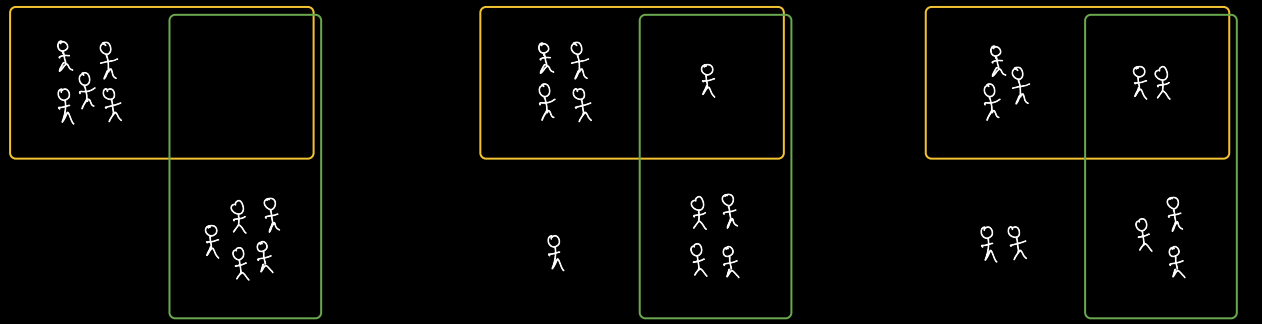


sorting

What is sorting?



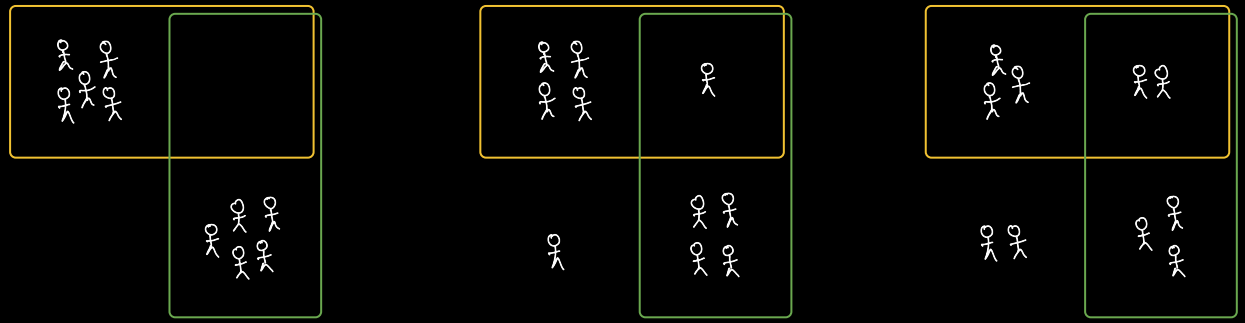
What is sorting?



MORE SORTED — LESS SORTED

PERFECT DEPENDENCE — ~INDEPENDENCE

What is sorting?



MORE SORTED — LESS SORTED

PERFECT DEPENDENCE — ~INDEPENDENCE

NO CROSSCUTTING — EVERYBODY CROSSCUTTING

I. Sortedness

II. Why unsort?

III. How to unsort?

IV. Can we unsort?

I. Sortedness

II. Why unsort?

III. How to unsort?

IV. Can we unsort?

Sorting is real.

Sorting is bad.

Unsorting is good.

Sorting is real.

Direct measurement:
(e.g., Abrams et al. 2015, Kaplan et al. 2022)

- ↑ correlations between issue positions
- “no evidence of polarization; the middle has not shrunk”
- ↑ geographic sorting

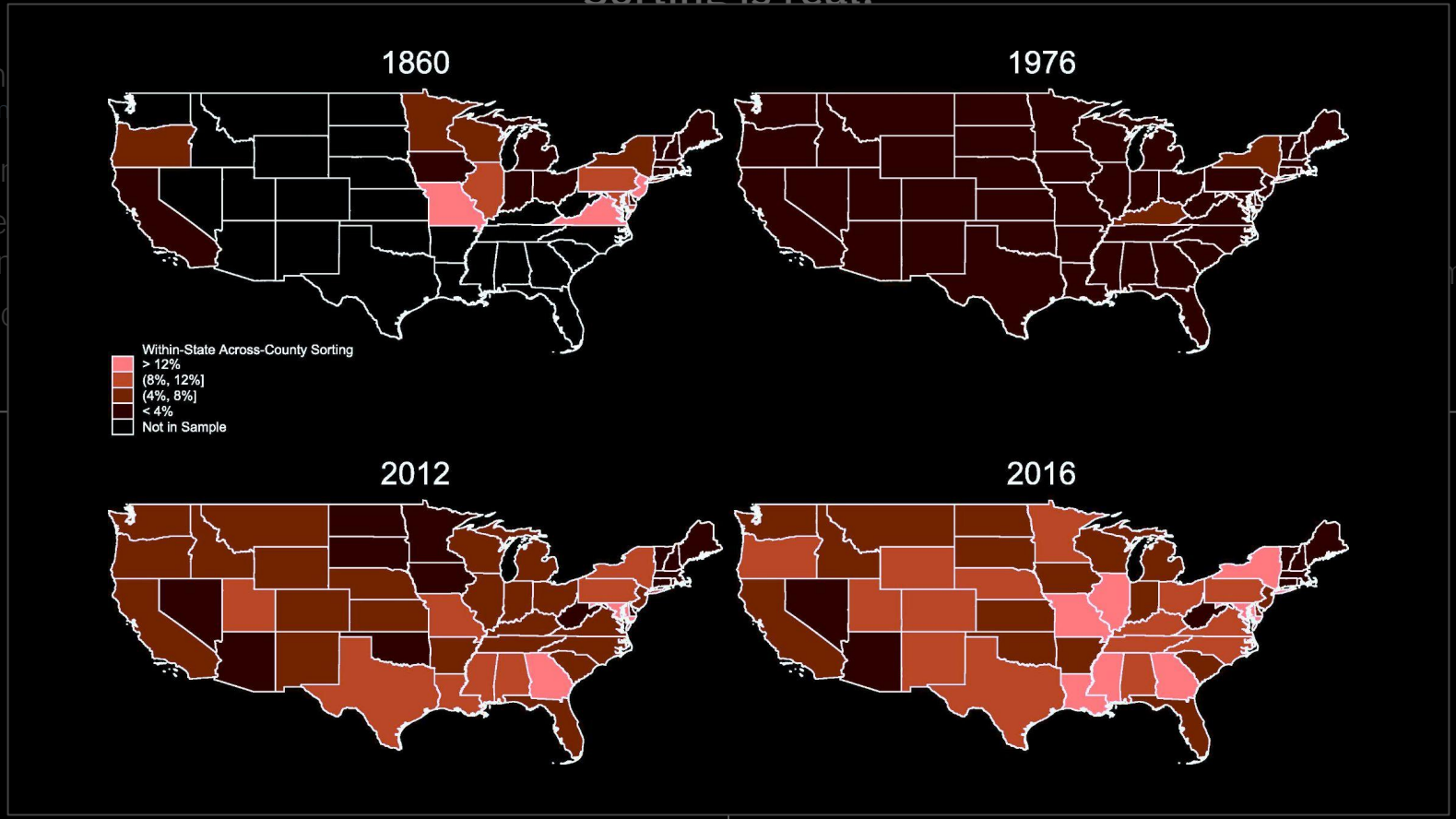
Explains empirical trends:
(Levendusky 2009; Abrams & Fiorina 2015; Fiorina 2017)

- ↑ difference in approval ratings
- ↓ split ticket voting
- campaigning to convince → campaigning to mobilize
- ↑ affective polarization

Sorting is bad.

Unsorting is good.

Sorting is real.



Sorting is real.

Direct measurement:
(e.g., Abrams et al. 2015, Kaplan et al. 2022)

- ↑ correlations between issue positions
- “no evidence of polarization; the middle has not shrunk”
- ↑ geographic sorting

Explains empirical trends:
(Levendusky 2009; Abrams & Fiorina 2015; Fiorina 2017)

- ↑ difference in approval ratings
- ↓ split ticket voting
- campaigning to convince → campaigning to mobilize
- ↑ affective polarization

Sorting is bad.

Unsorting is good.

Sorting is real.

Direct measurement:
(e.g., Abrams et al. 2015, Kaplan et al. 2022)

- ↑ correlations between issue positions
- "no evidence of polarization; the middle has not shrunk"
- ↑ geographic sorting

Explains empirical trends:
(Levendusky 2009; Abrams & Fiorina 2015; Fiorina 2017)

- ↑ difference in approval ratings
- ↓ split ticket voting
- campaigning to convince → campaigning to mobilize
- ↑ affective polarization

Sorting is bad.

- ↑ pigeonholing / stereotyping
- ↓ representation of nuanced, crosscutting positions (Fiorina 2016)
- ↑ risk of civil war
  (Selway 2011; Gubler and Selway 2012; Siroky and Hechter 2016)

Unsorting is good.

Sorting is real.

Direct measurement:
(e.g., Abrams et al. 2015, Kaplan et al. 2022)

- ↑ correlations between issue positions
- "no evidence of polarization; the middle has not shrunk"
- ↑ geographic sorting

Explains empirical trends:
(Levendusky 2009; Abrams & Fiorina 2015; Fiorina 2017)

- ↑ difference in approval ratings
- ↓ split ticket voting
- campaigning to convince → campaigning to mobilize
- ↑ affective polarization

Sorting is bad.

- ↑ pigeonholing / stereotyping
- ↓ representation of nuanced, crosscutting positions (Fiorina 2016)
- ↑ risk of civil war
  (Selway 2011; Gubler and Selway 2012; Siroky and Hechter 2016)

Unsorting is good.

- ↑ "surprising validators" (Glaeser and Sunstein 2014)
- ↑ extent to which any majority will have something in common with a minority
- ↑ weak/long ties → ↑ economic outcomes
  (Jahani et al., 2023)

Sortedness by other names:

safeguards against factionalism
James Madison, *The Federalist No. 10* (1787)

“intersection of social circles”
Georg Simmel, *Soziologie* (1908)

“overlapping memberships”
Robert Dahl, *A Preface to Democratic Theory* (1956)

“overlapping consensus”
John Rawls, *Political Liberalism* (1993)

“diverse diversities”
Amartya Sen, *Identity and Violence* (2006)

“connected society” / “polypolitanism”
Danielle Allen, *Justice by Means of Democracy* (2023)

Why not unsort?

- Probably **too neutral a goal**
- Maybe **absurd if taken to the extreme**
- Maybe **the necessary interventions would be unethical**
- Makes **engaging in politics more complicated**

I. Sortedness

II. Why unsort?

III. How to unsort?

IV. Can we unsort?

I. Sortedness

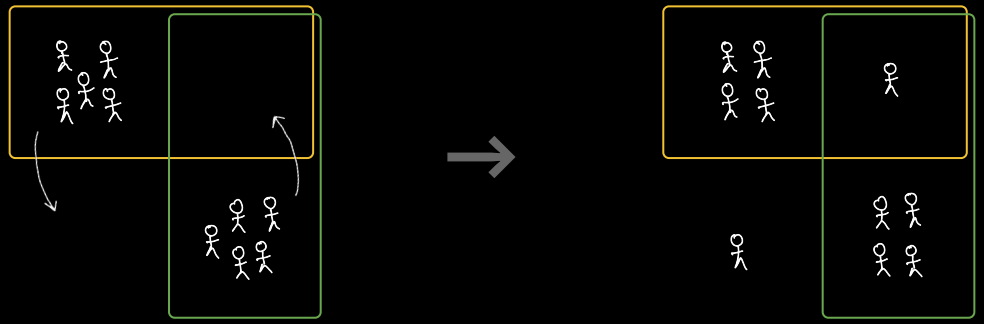
II. Why unsort?

III. How to unsort?

IV. Can we unsort?

Mechanism 1

Conversion



EXAMPLES

common ground

symmachy / common enemies

surprising validation

consilience

weak / long ties

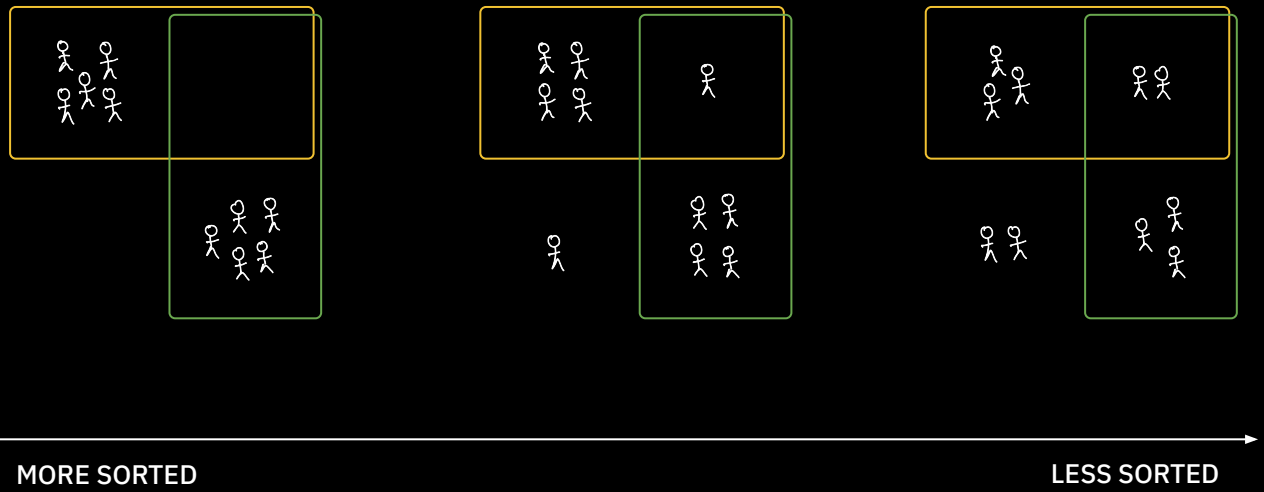
“complicate the narratives”

Most bridging algorithms use “diverse approval”.

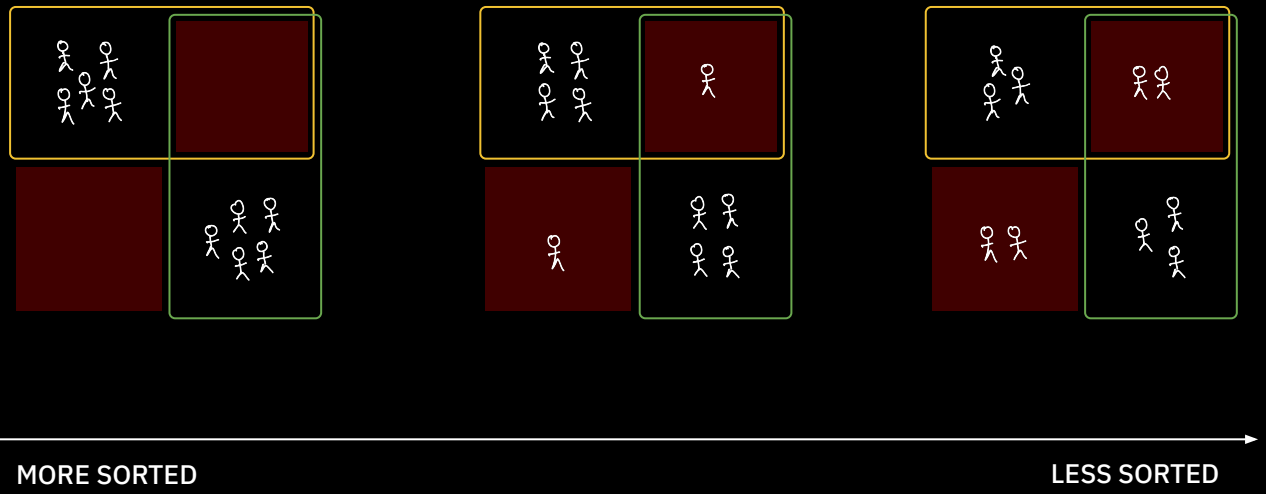
Most bridging algorithms use “diverse approval”.

EXAMPLES

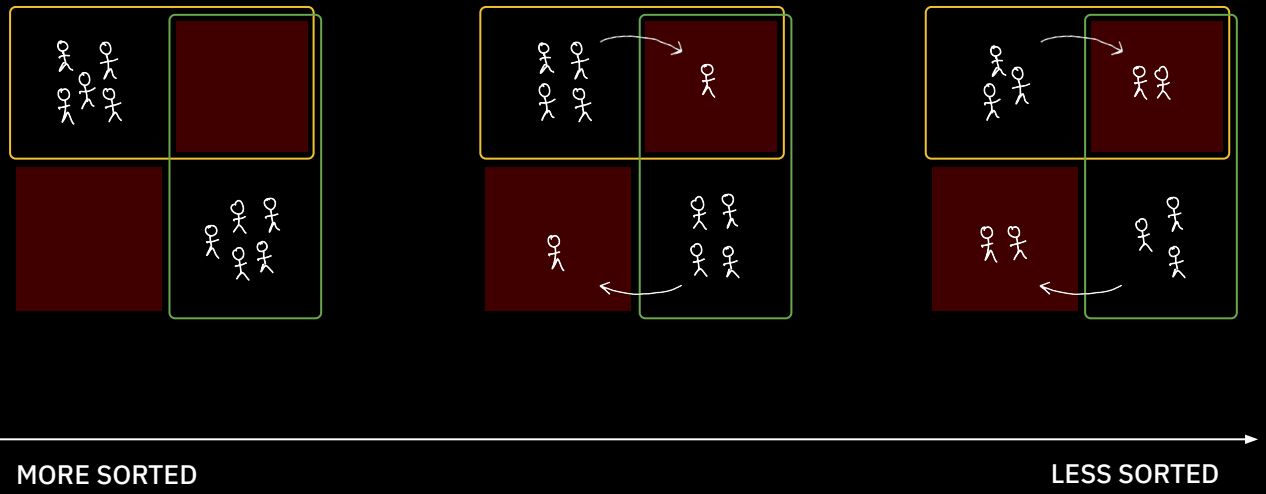
Diverse approval selects for unlikely combinations of attributes.



Diverse approval selects for unlikely combinations of attributes.



Diverse approval selects for unlikely combinations of attributes.



Diverse approval selects for unlikely combinations of attributes.

Cognitive “biases”

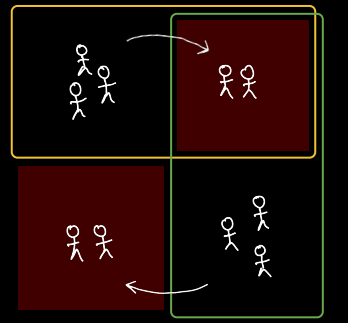
- mere-exposure effect
- illusory truth effect
- experience effects
- anchoring
- pressures to conform

Learning

- learning new facts
- discovering new interests

Algorithmic incentives (Brady et al., 2023)

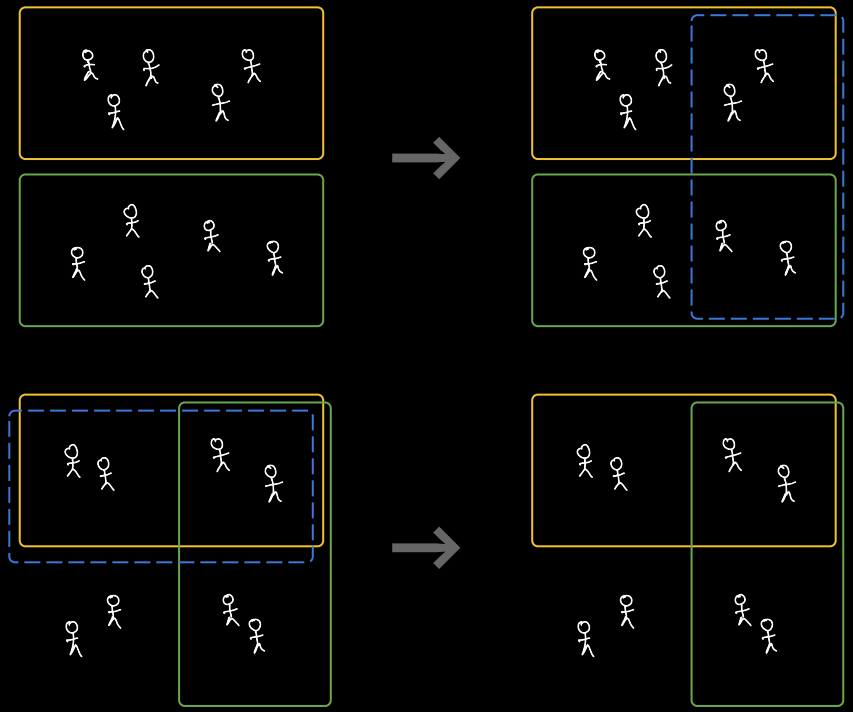
- observational learning
- reinforcement learning
- intentional strategic adaptation



LESS SORTED

Mechanism 2

Adding or removing groups/dimensions



EXAMPLES

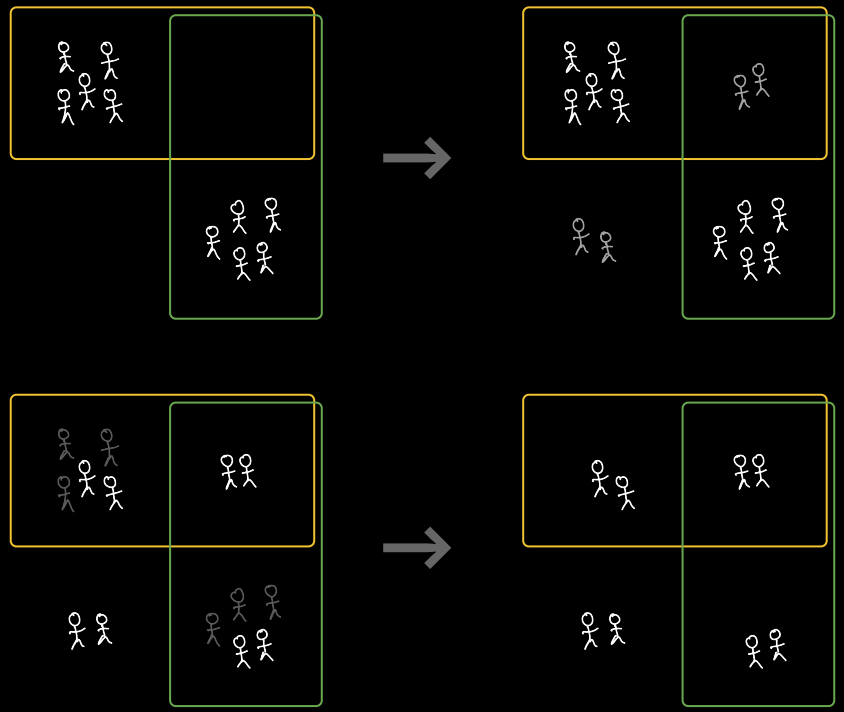
humanization

intergroup contact

depoliticization

Mechanism 3

Entry or exit of people



EXAMPLES

migration

generational change

account creation / deletion

I. Sortedness

II. Why unsort?

III. How to unsort?

IV. Can we unsort?

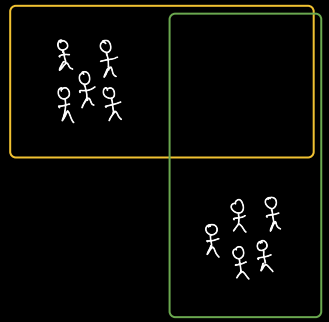
I. Sortedness

II. Why unsort?

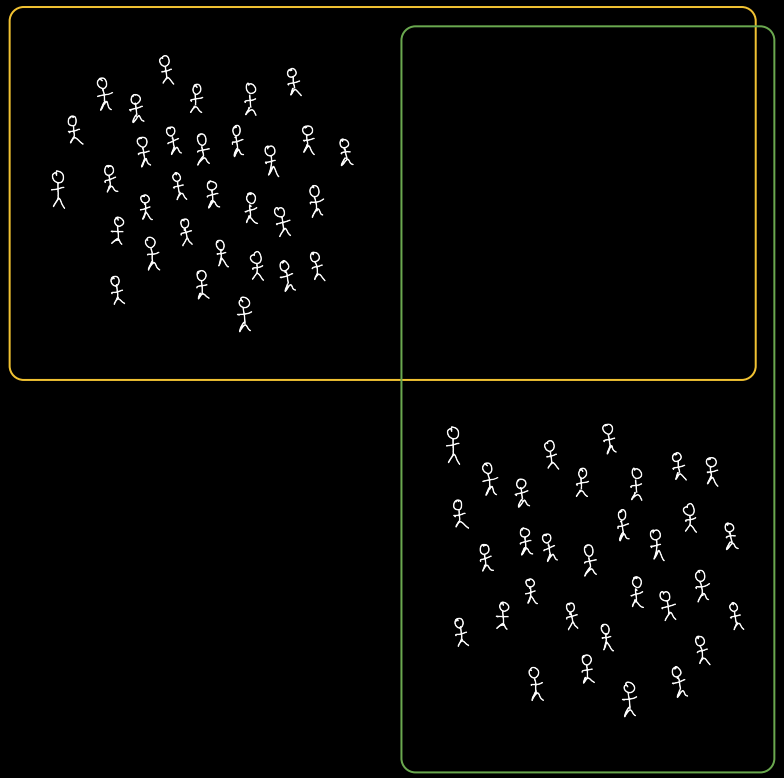
III. How to unsort?

IV. Can we unsort?

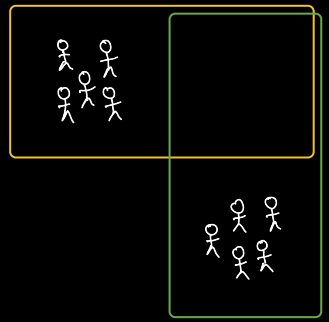
Sortedness can be reduced globally by intervening locally.



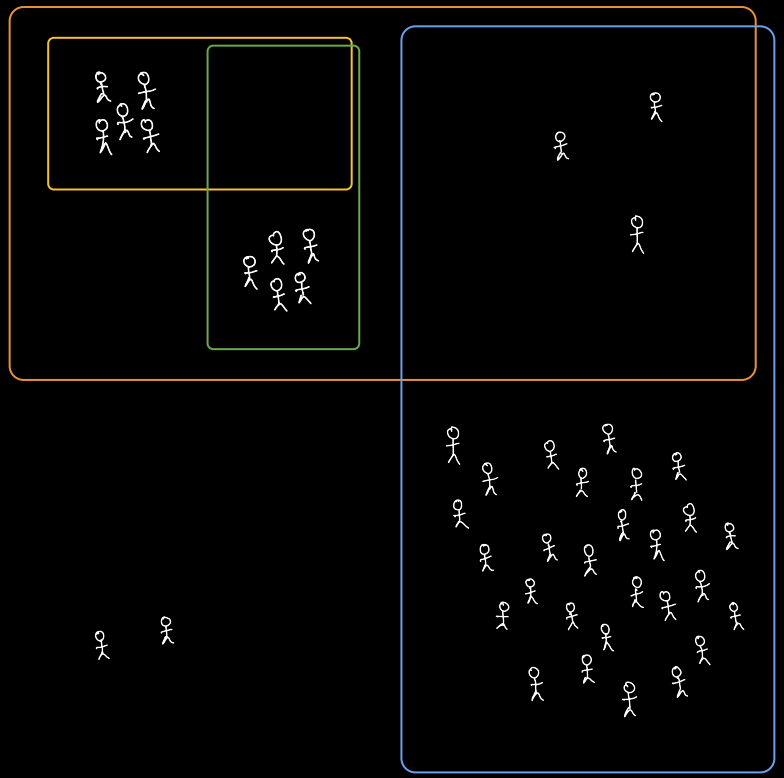
Sortedness can be reduced globally by intervening locally.



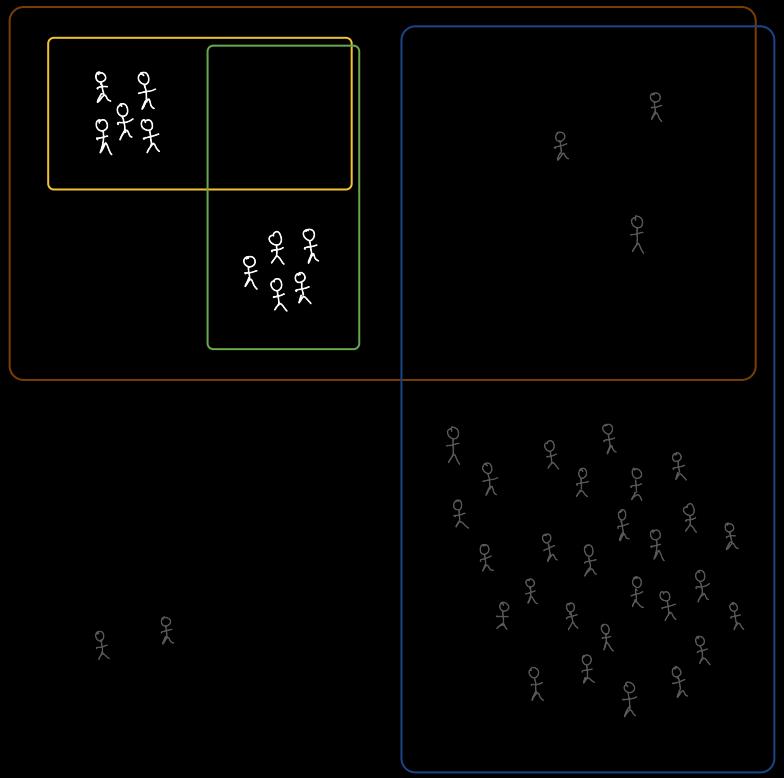
Sortedness can be reduced globally by intervening locally.



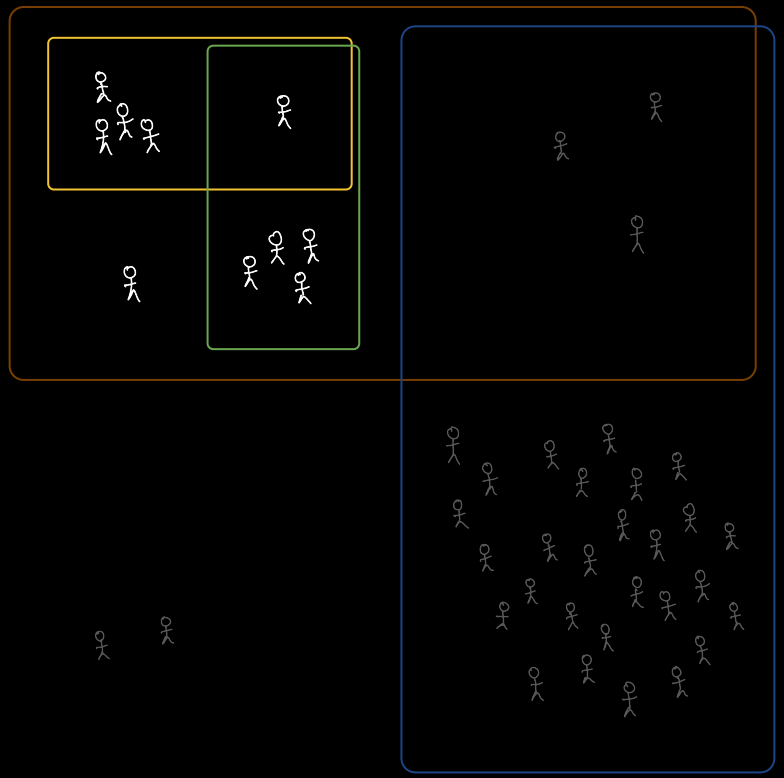
Sortedness can be reduced globally by intervening locally.



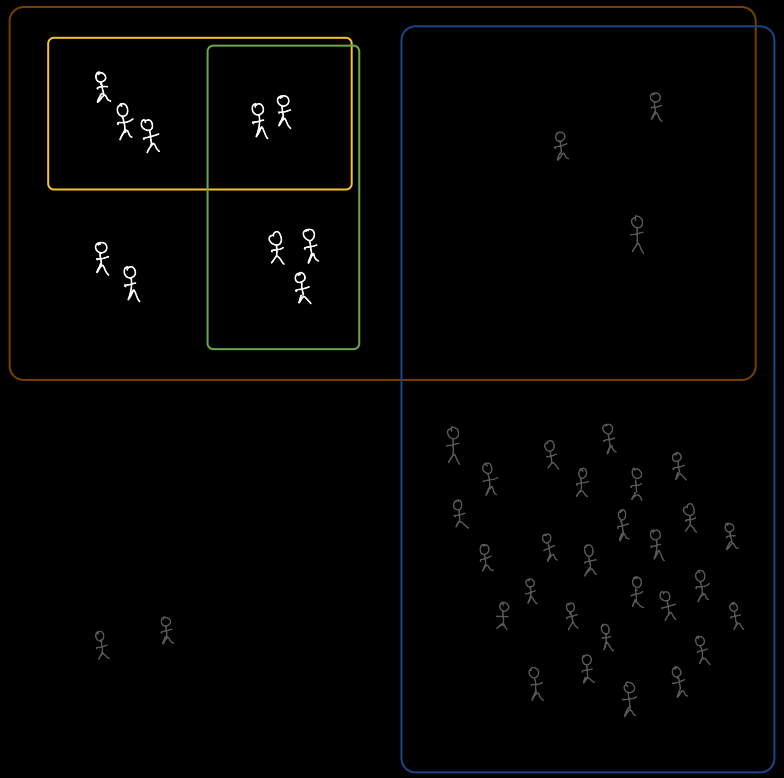
Sortedness can be reduced globally by intervening locally.



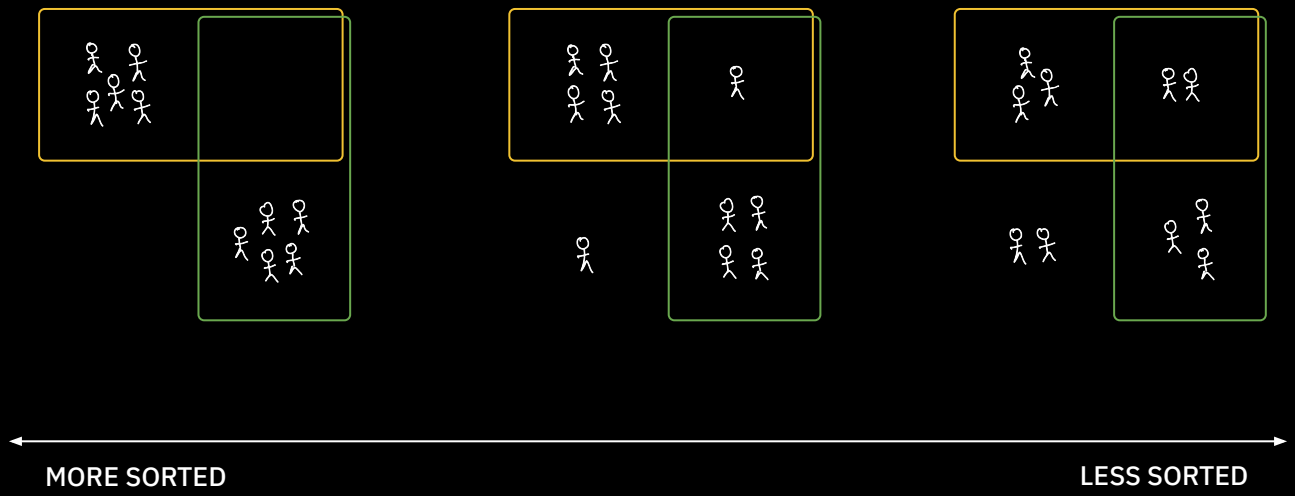
Sortedness can be reduced globally by intervening locally.



Sortedness can be reduced globally by intervening locally.



Sortedness is orthogonal to “public opinion”.



Sortedness is orthogonal to “public opinion”.