



Timing:  
2:00 - 3:30  
4:00 - 6:00

# Experiments in Computational Social Choice

(Using Maps of Elections)

Niclas Boehmer



Piotr Faliszewski



Stanisław Szufa



Tomasz Wąs





Timing:  
2:00 - 3:30  
4:00 - 6:00

# Using Maps of Elections

Experiments in Computational Social Choice

Niclas Boehmer



Piotr Faliszewski



Stanisław Szufa



Tomasz Wąs



# An Election

$v_1$ : 🐼 > 🐳 > 🐱

$v_2$ : 🐳 > 🐱 > 🐼

$v_3$ : 🐼 > 🐱 > 🐳




$v_4$ : 🐱 > 🐼 > 🐳




$$E = (C, V)$$




$$C = \{ \text{🐼}, \text{🐳}, \text{🐱} \}$$




$$V = (v_1, v_2, v_3, v_4)$$

# An Election

$v_1$ :  >  > 

$v_2$ :  >  > 

$v_3$ :  >  > 

$v_4$ :  >  > 

$$E = (C, V)$$

$$C = \{ \text{panda}, \text{whale}, \text{cat} \}$$

$$V = (v_1, v_2, v_3, v_4)$$

## Also an election

$v_1$ : {  ,  }

$v_2$ : {  ,  }













$v_3$ : {  ,  ,  }

$v_4$ : {  ,  }

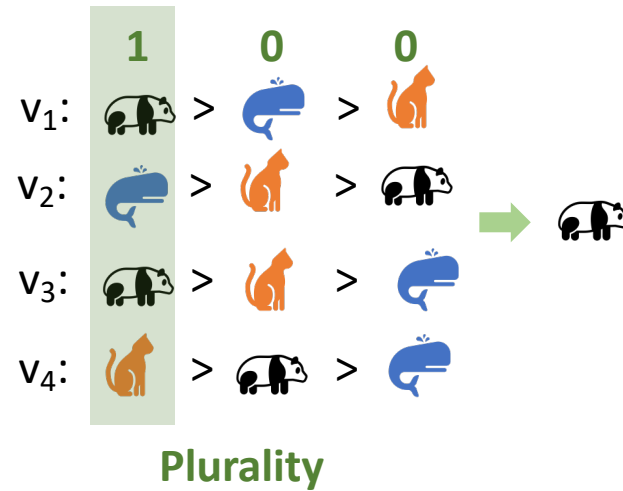
We mostly focus on the ordinal setting, but approvals will come!















## An Election

- $v_1$ :  >  > 
- $v_2$ :  >  > 
- $v_3$ :  >  > 
- $v_4$ :  >  > 

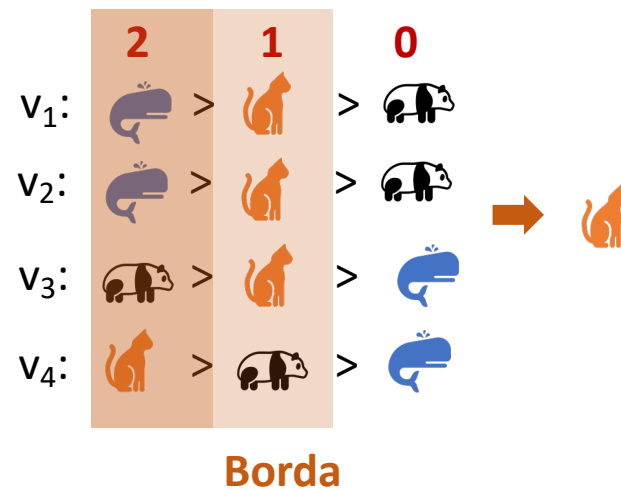
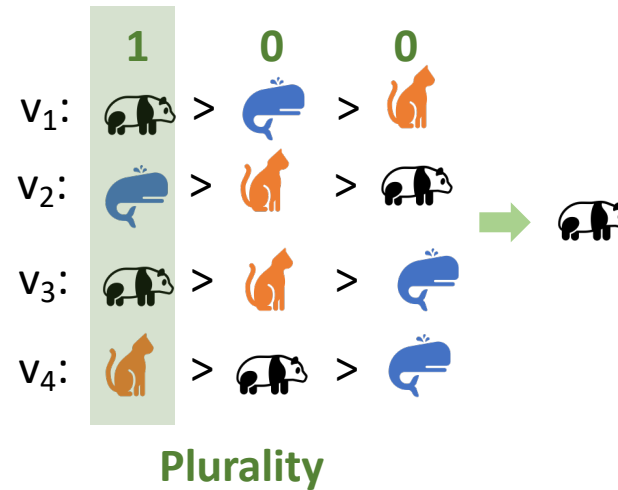
## Winner Determination















## An Election

- $v_1$ :  >  >   
 $v_2$ :  >  >   
 $v_3$ :  >  >   
 $v_4$ :  >  > 

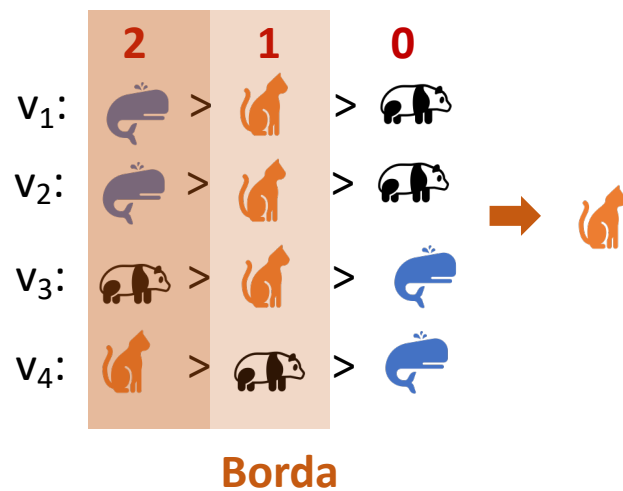
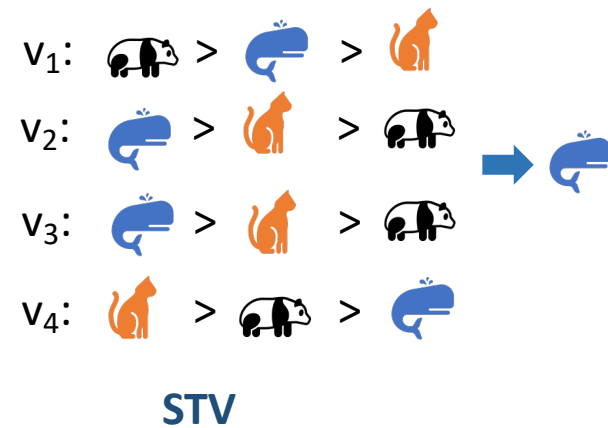
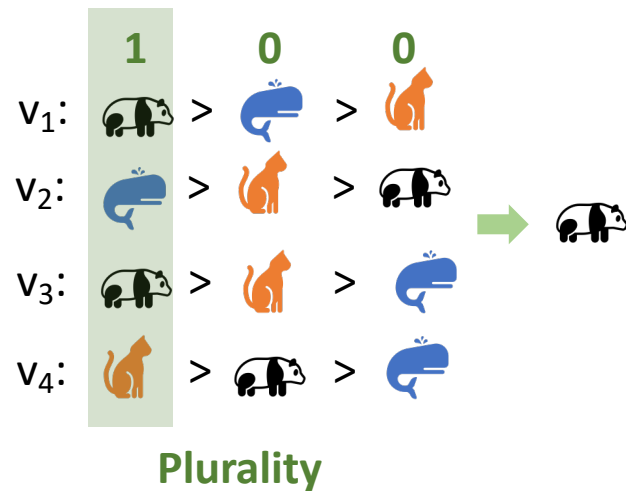
## Winner Determination















## An Election

- $v_1$ :  >  >   
 $v_2$ :  >  >   
 $v_3$ :  >  >   
 $v_4$ :  >  > 

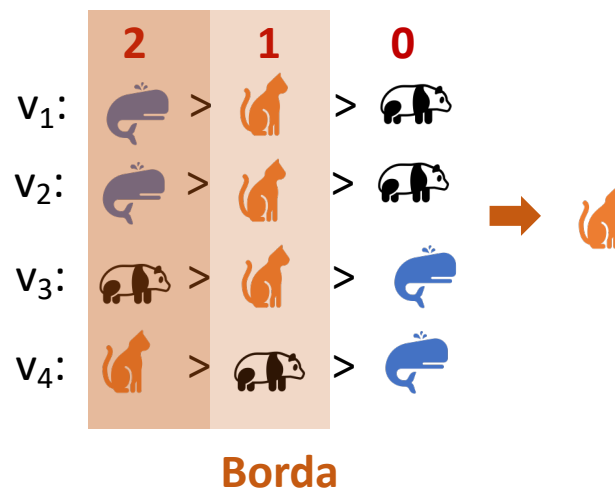
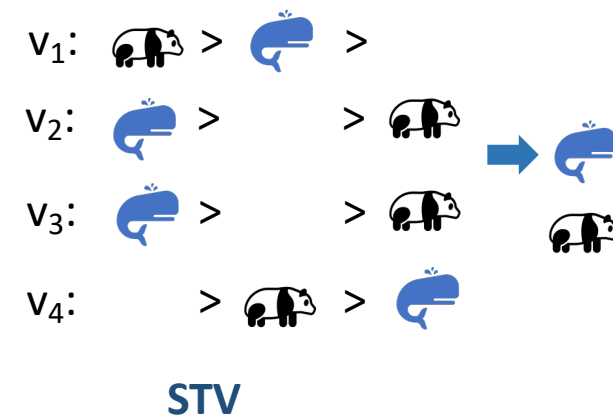
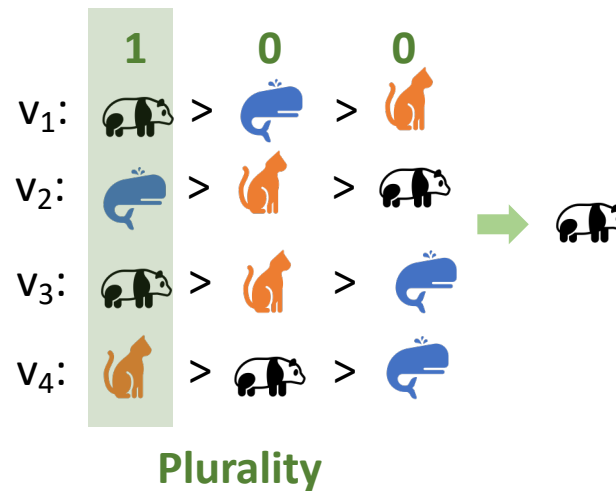
## Winner Determination















## An Election

- $v_1$ :  >  >   
 $v_2$ :  >  >   
 $v_3$ :  >  >   
 $v_4$ :  >  > 

## Winner Determination















## An Election


- $V_1$ :  >  >   
 $V_2$ :  >  >   
 $V_3$ :  >  >   
 $V_4$ :  >  > 

## Winner Determination
















**Plurality**

	<b>1</b>	<b>0</b>	<b>0</b>		
$V_1$ :		>		>	
$V_2$ :		>		>	
$V_3$ :		>		>	
$V_4$ :		>		>	













→ 


**STV**

$V_1$ :		>		>	
$V_2$ :		>		>	
$V_3$ :		>		>	
$V_4$ :		>		>	













→ 


**Borda**

	<b>2</b>	<b>1</b>	<b>0</b>		
$V_1$ :		>		>	
$V_2$ :		>		>	
$V_3$ :		>		>	
$V_4$ :		>		>	

→ 

**Dodgson**

$V_1$ :		>		>	
$V_2$ :		>		>	
$V_3$ :		>		>	
$V_4$ :		>		>	

→ 





Winner Determination

Result Modification/Analysis

An Election

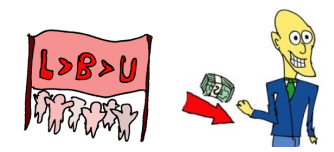
- V<sub>1</sub>: > >
- V<sub>2</sub>: > >
- V<sub>3</sub>: > >
- V<sub>4</sub>: > >

Possible/Necessary Winner (in various shapes)

- V<sub>1</sub>: { , } >
- V<sub>2</sub>: > { , }
- V<sub>3</sub>: > >
- V<sub>4</sub>: { , , }

Manipulating Elections

Bribery/Campaign Management



Electoral Control



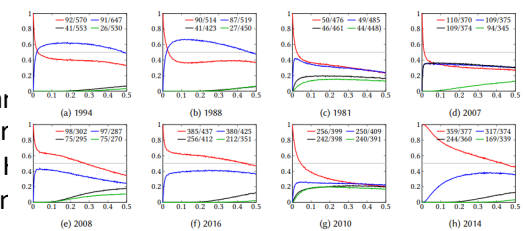
Strategic Voting



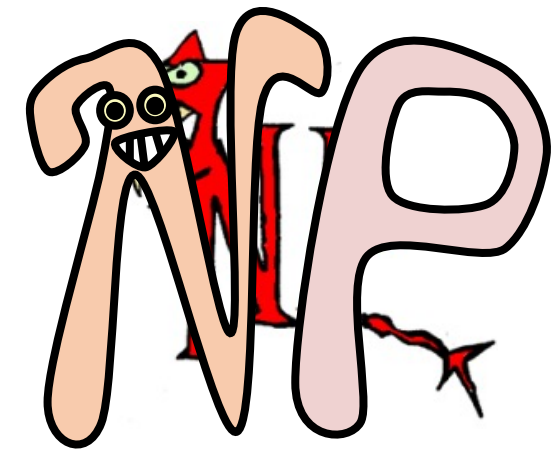
Strategic Candidacy



Robustness / Winner Assessment / Margin of Victory



B. Dutta, M. Jackson, M. Le Breton, Strategic Candidacy at K. Konczak, J. Lang. Voting procedures with incomplete pr V. Conitzer, T. Walsh, Barriers to Manipulation in Voting. I P. Faliszewski, J. Rothe, Control and Bribery in Voting. Har



Winner Determination

Result Modification/Analysis

Manipulating Elections

Bribery/Campaign Management

Electoral Control

Strategic Voting

Strategic Candidacy

Robustness / Winner

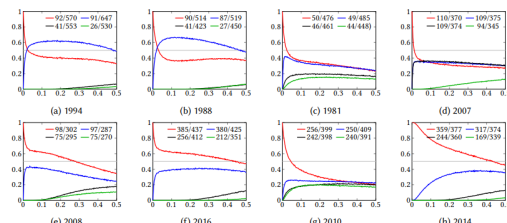
Assessment / Margin of Victory

### An Election

- $v_1$ : > >
- $v_2$ : > >
- $v_3$ : > >
- $v_4$ : > >

Possible/Necessary Winner (in various shapes)

- $v_1$ : { , } >
- $v_2$ : > { , }
- $v_3$ : > >
- $v_4$ : { , , }

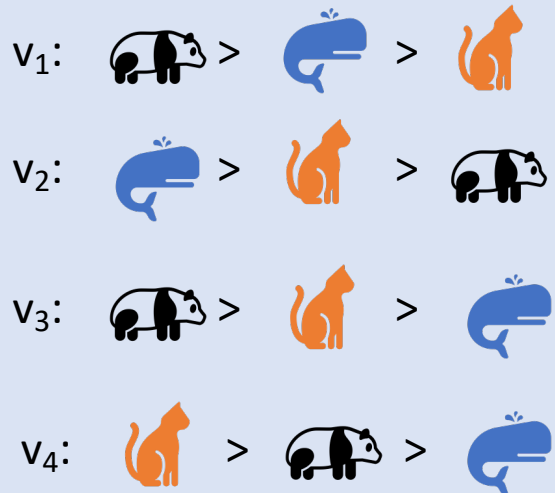


B. Dutta, M. Jackson, M. Le Breton, Strategic Candidacy at K. Konczak, J. Lang. Voting procedures with incomplete pr V. Conitzer, T. Walsh, Barriers to Manipulation in Voting. I P. Faliszewski, J. Rothe, Control and Bribery in Voting. Har





# An Election



	symmetry	consistency	weak efficiency	efficiency	continuity	indep. of irr. alt.	monotonicity	D'Hondt prop.	disjoint equality	disjoint diversity
ABC counting rules	+	+	+	+	+					
Thiele Methods	+	+	+	+	+	+				
Dissatisfaction counting rules	+	+	+	+	+		+			
Multi-winner Approval Voting (AV)	+	+	+	+	+	+	+		+	
Proportional Approval Voting (PAV)	+	+	+	+	+	+		+		
Approval Chamberlin-Courant (CC)	+	+	+	+	+	+				+
Constant Threshold Methods	+	+	+	+	+	+				
Satisfaction Approval Voting	+	+	+	+	+					
Sequential Thiele Methods	+	+	+	+	+		+			
Reverse-sequential Thiele Methods	+	+	+	+	+					
Sequential PAV	+	+	+	+	+	+		+		
Reverse-Sequential PAV	+	+	+	+	+					+

## Winner Determination

## Result Modification/Analysis

## Normative Properties

- Monotonicity
- Homogeneity
- Consistency
- Condorcet Consistency
- (Something) Justified Representation
- Core
- Priceability

Rule	Complexity	JR/PJR/EJR	PR	SMWPI	SMWOP1	Com. Mon.
AV	P <sup>a</sup>	No <sup>d</sup>	No <sup>k</sup>	Str. Thm. 3.4	Str. Thm. 3.6	Yes
SAV	P <sup>a</sup>	No <sup>d</sup>	No <sup>k</sup>	Str. Thm. 3.4	Str. Thm. 3.6	Yes
CC	NP-comp. <sup>b</sup>	JR <sup>d</sup>	Yes <sup>h</sup>	Str. Thm. 3.4	Wk. Thm. 3.3	No <sup>Ex. 4.2</sup>
Monroe	NP-comp. <sup>b</sup>	JR <sup>d,c</sup>	Yes <sup>h</sup>	No <sup>Ex. 3.3</sup>	Wk. Thm. 3.9	No <sup>Ex. 4.2</sup>
PAV	NP-comp. <sup>a</sup>	EJR <sup>d</sup>	No <sup>k</sup>	Str. Thm. 3.4	Wk. Thm. 3.3	No <sup>f</sup>
max-Phragmén	NP-comp. <sup>c</sup>	PJR <sup>d,f</sup>	Yes <sup>g</sup>	Wk. Thm. 3.12	Wk. Thm. 3.12	No <sup>e</sup>

<sup>a</sup> Results by Aziz et al. [4] and Skovron et al. [31].  
<sup>b</sup> Results by Panagiotou et al. [28].  
<sup>c</sup> Results by Brill et al. [7].  
<sup>d</sup> Results by Aziz et al. [2].  
<sup>e</sup> Monroe satisfies PJR if  $k$  divides  $n$  [29].  
<sup>f</sup> max-Phragmén satisfies PJR when combined with certain tie-breaking rule [7].  
<sup>g</sup> CC satisfies PR if ties are broken always in favour of the candidates subsets that provide PR.  
<sup>h</sup> Results by Sánchez-Fernández et al. [29].  
<sup>i</sup> Results by Jaume [15], Mira and Oliver [31], and Phragmén [36].  
<sup>j</sup> Results by Thiele [33].  
<sup>k</sup> Results by Sánchez-Fernández and Fritsch [39].

Table 1: Properties of approval-based multi-winner voting rules

System	Monotonic	Condorcet winner	Majority	Condorcet loser	Majority loser	Mutual majority	Smith	ISDA	LSA	Independence of clones	Reversal	Participation, consistency	Later-no-harm	Later-no-help	Polynomial time	Resolvability
Schulze	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	No	No	Yes	Yes
Ranked pairs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes
Tideman's Alternative	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No	No	No	Yes	Yes
Kemeny-Young	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No	No	No	Yes
Copeland	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	Yes
Nanson	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	No	No	No	Yes	Yes
Black	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	Yes	No	No	No	Yes	Yes
Instant runoff voting	No	No	Yes	Yes	Yes	Yes	No	No	No	Yes	No	No	Yes	Yes	Yes	Yes
Smith/STV	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No	No	No	Yes	Yes
Borda	Yes	No	No	Yes	Yes	No	No	No	No	No	Yes	Yes	No	Yes	Yes	Yes
Baldwin	No	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	No	Yes	Yes
Bucklin	Yes	No	Yes	No	Yes	Yes	No	No	No	No	No	No	No	Yes	Yes	Yes
Plurality	Yes	No	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Condorcet voting	No	No	Yes	Yes	Yes	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Coombs <sup>[22]</sup>	No	No	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	No	No	Yes
Maximal deficit	Yes	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes
Anti-plurality <sup>[22]</sup>	Yes	No	No	No	Yes	No	No	No	No	No	No	Yes	No	No	Yes	Yes
Sri Lankan contingent voting	No	No	Yes	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Supplementary voting	No	No	Yes	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Dodgson <sup>[24]</sup>	No	Yes	Yes	No	No	No	No	No	No	No	No	No	No	No	No	Yes

	M	PO	IAWP	CC	IALP	ICLP
Lexicographic rule	✓	✓	✓			
Condorcet's practical method	✓	✓ ( $m=3$ )	✓		✓ ( $m \leq 4$ )	
Fallback Bargaining	✓	✓				
Majoritarian Compromise	✓	✓	✓		✓	
Obata and Ishii's method	✓	✓	✓			
Contreras, Hinojosa and Mármol's method	✓	✓	✓			
Geometric rule	✓	✓				

M: Monotonicity, PO: Pareto-optimality, IAWP: Immunity to the absolute winner paradox, CC: Condorcet consistency, IALP: Immunity to the absolute loser paradox, ICLP: Immunity to the Condorcet loser paradox.

	Pareto efficiency	committee monoton.	support with add. voters	monot. without add. voters	consist.	inclusion-strategypr.	comput. complexity
AV	strong	✓	✓	/	✓	✓	P
CC	weak	×	✓	/	cand	✓	NP-hard
PAV	strong	×	✓	/	cand	✓	NP-hard
seq-PAV	×	✓	≥cand	/	cand	×	P
seq-CC	×	✓	?	/	?	×	P
rev-seq-PAV	×	✓	≥cand	/	cand	×	P
Monroe	×	×	×	/	cand	×	NP-hard
Greedy Monroe	×	×	×	/	?	×	P
seq-Phragmén	×	✓	cand	/	cand	×	P
lexmin-Phragmén	×	×	cand	/	cand	×	NP-hard
Rule X	?	×	×	/	?	×	P
MAV	weak	×	?	/	?	×	NP-hard
SAV	strong	✓	✓	/	✓	✓	P

# An Election

- V<sub>1</sub>: 🐼 > 🐳 > 🐱
- V<sub>2</sub>: 🐳 > 🐱 > 🐼
- V<sub>3</sub>: 🐼 > 🐱 > 🐳
- V<sub>4</sub>: 🐱 > 🐼 > 🐳

Winner Determination

Result Modification/Analysis

Normative Properties

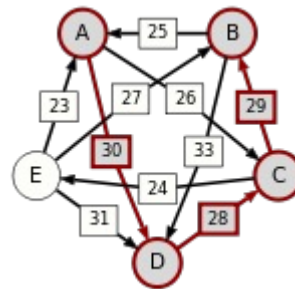
New Rules, New Settings



Portioning



Sortition



Schulze

The **Method of Equal Shares** is a fairer voting rule for participatory budgeting.

It provides proportional representation and allows every voter to decide about an equal part of the budget.

**— Key Benefits —**

- The method is simple to understand and to explain.
- Can be used in any participatory budgeting process, no matter the scale.
- Theoretical guarantees that all interest groups will be represented in the outcome.
- Better reflects voter preferences across project categories.
- The voting experience is unchanged: the Method of Equal Shares works with all standard ballot types (approvals, knapsack voting, rankings, distributing points, etc.)
- Increased transparency: voters can see how their vote influenced the election.
- Straightforward to implement in any software system.

DEVELOPED AND STUDIED BY RESEARCHERS AT UNIVERSITIES AROUND THE WORLD

UNIVERSITY OF MONTREAL | CNRS | UNIVERSITY OF TORONTO | AGH | ETH zürich | TU BERGAKADEMIE FREIBERG

Participatory Budgeting

Winner Determination

Result Modification/Analysis

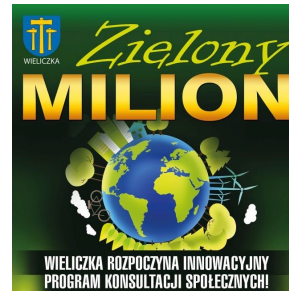
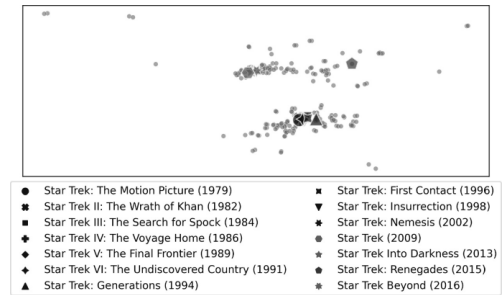
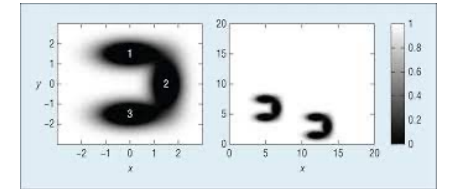
Normative Properties

New Rules, New Settings

Applications

### An Election

- V<sub>1</sub>: 🐼 > 🐳 > 🐱
- V<sub>2</sub>: 🐳 > 🐱 > 🐼
- V<sub>3</sub>: 🐼 > 🐱 > 🐳
- V<sub>4</sub>: 🐱 > 🐼 > 🐳



## An Election

- $v_1$ : 🐼 > 🐳 > 🐱
- $v_2$ : 🐳 > 🐱 > 🐼
- $v_3$ : 🐼 > 🐱 > 🐳
- $v_4$ : 🐱 > 🐼 > 🐳

Largely studied  
theoretically

We want more  
experiments!

## Benefits of Experiments

- More complex settings
- More precise results
  - Exact running time vs asymptotic running time
- Observe actual phenomena instead of merely predicting their possibility
  - Condorcet winners often exist
  - No-show paradox is/is-not a problem
  - Voting rules do/do-not give very different results

## Problems with Experiments

- They don't generalize
- May be misleading
- Some insights are impossible to get experimentally
- You never really know...

**Data!**

Real-Life Data

Preflib

Pabulib

Others...

Synthetic Data

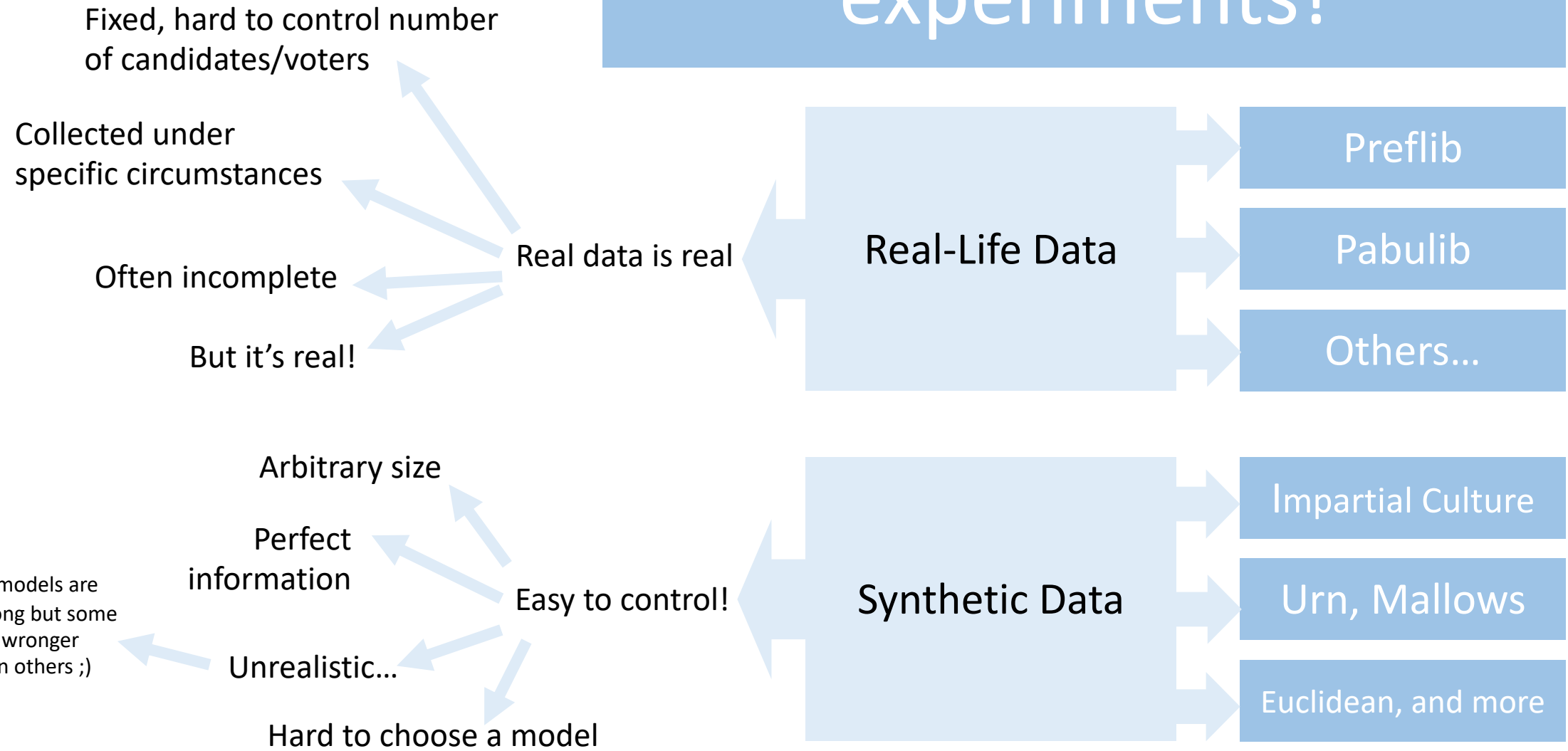
Impartial Culture

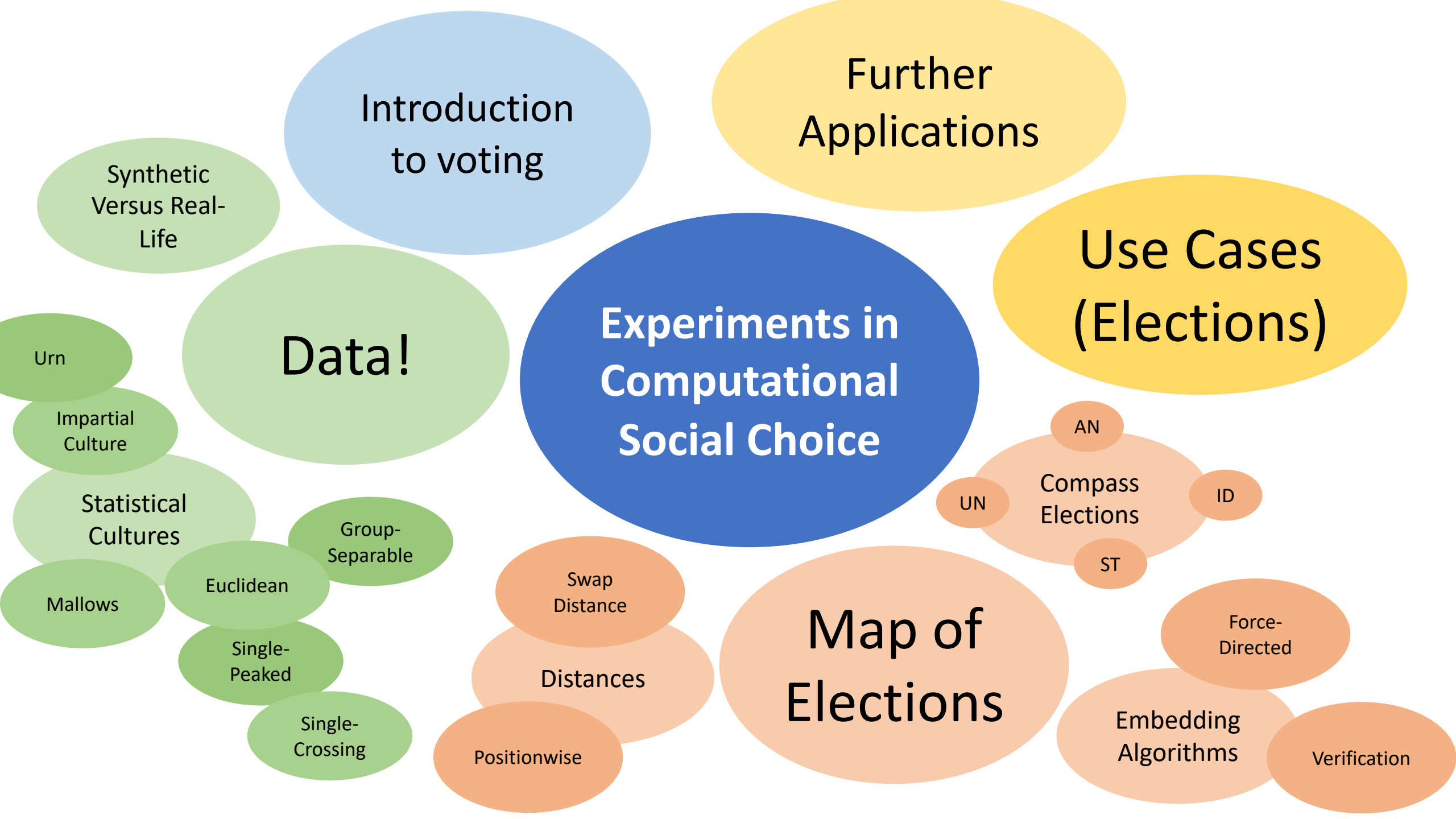
Urn, Mallows

Euclidean, and more

# We want more experiments!

# We want more experiments!





Further Applications

Use Cases (Elections)

Data!

Experiments in Computational Social Choice

Map of Elections

Introduction to voting

Synthetic Versus Real-Life

Urn

Impartial Culture

Statistical Cultures

Group-Separable

Mallows

Euclidean

Single-Peaked

Single-Crossing

UN

Compass Elections

ID

AN

ST

Swap Distance

Distances

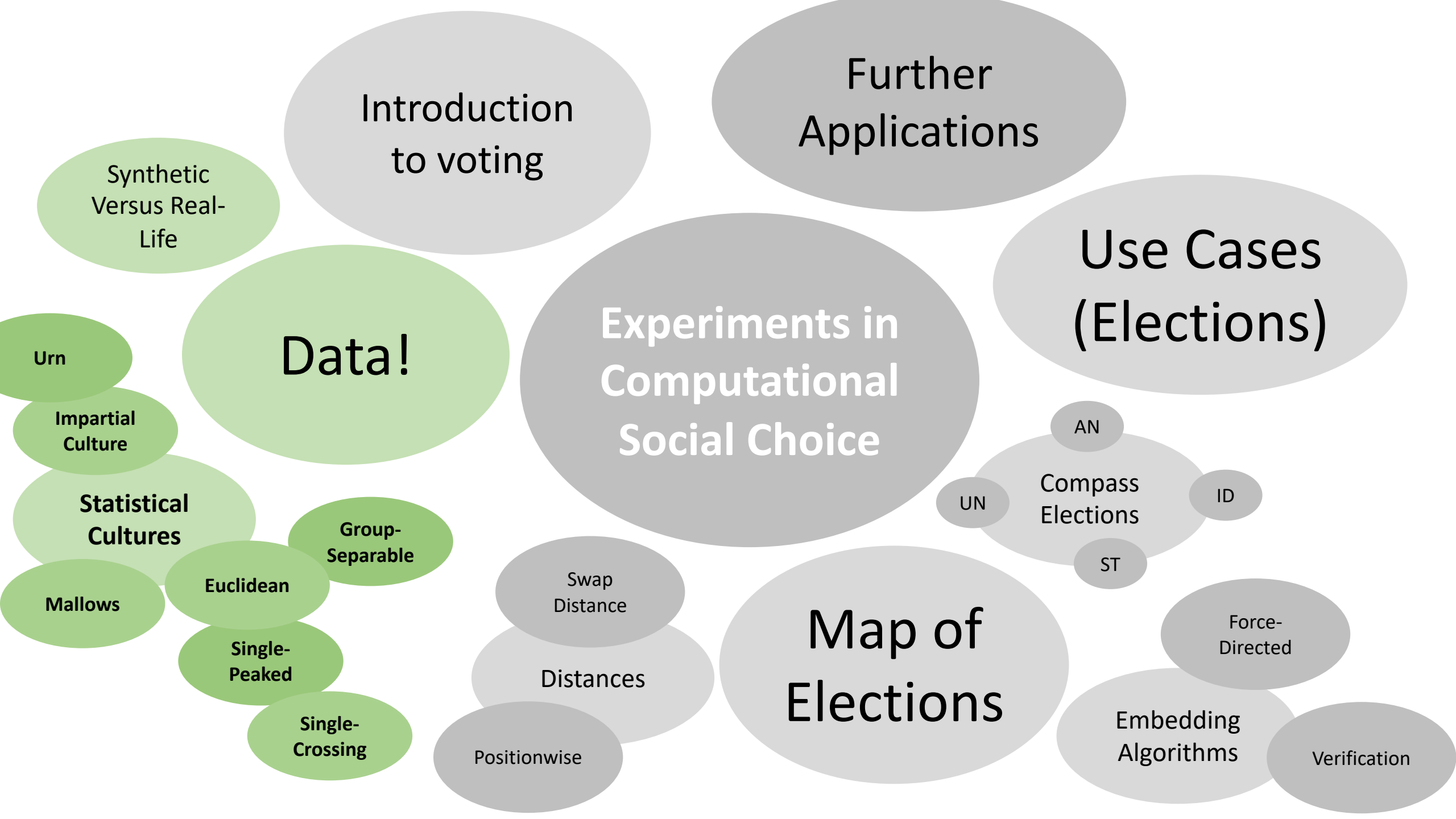
Positionwise

Force-Directed

Embedding Algorithms

Verification







5 minutes

# Basic Statistical Cultures

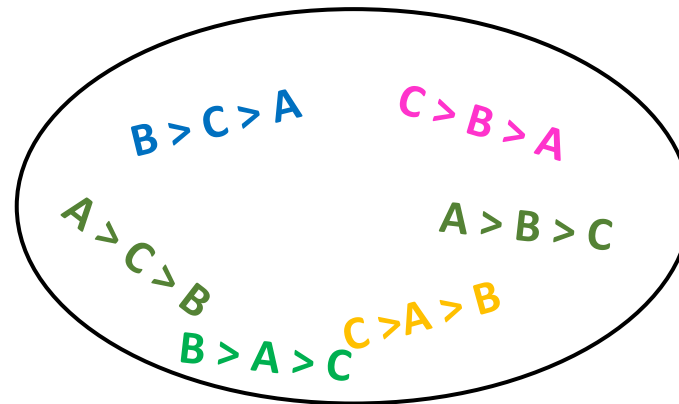
# Statistical Cultures

**Impartial Culture (IC):** Every preference order comes with the same probability (a.k.a. **uniform distribution**)

**Polya-Eggenberger Urn Model:**

Form an urn of **all possible  $m!$  votes**. To generate a vote:

- 1) Choose a vote from the urn and add it to your election
- 2) **Return** the vote to the urn, together with  **$\alpha \cdot m!$  copies**.



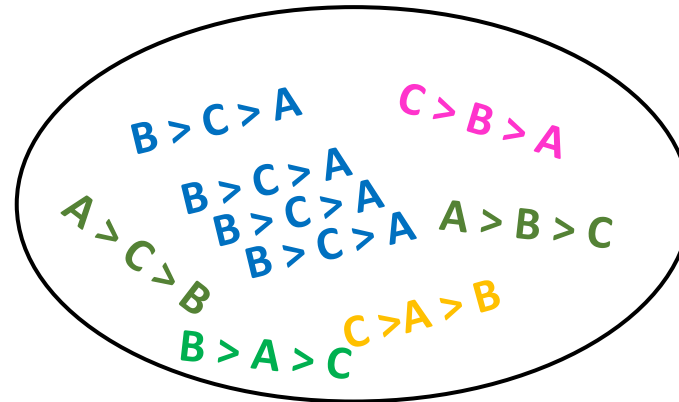
# Statistical Cultures

**Impartial Culture (IC):** Every preference order comes with the same probability (a.k.a. **uniform distribution**)

**Polya-Eggenberger Urn Model:**

Form an urn of **all possible  $m!$  votes**. To generate a vote:

- 1) Choose a vote from the urn and add it to your election
- 2) **Return** the vote to the urn, together with  **$\alpha \cdot m!$  copies**.



# Statistical Cultures

**Impartial Culture (IC):** Every preference order comes with the same probability (a.k.a. **uniform distribution**)

**Polya-Eggenberger Urn Model:**

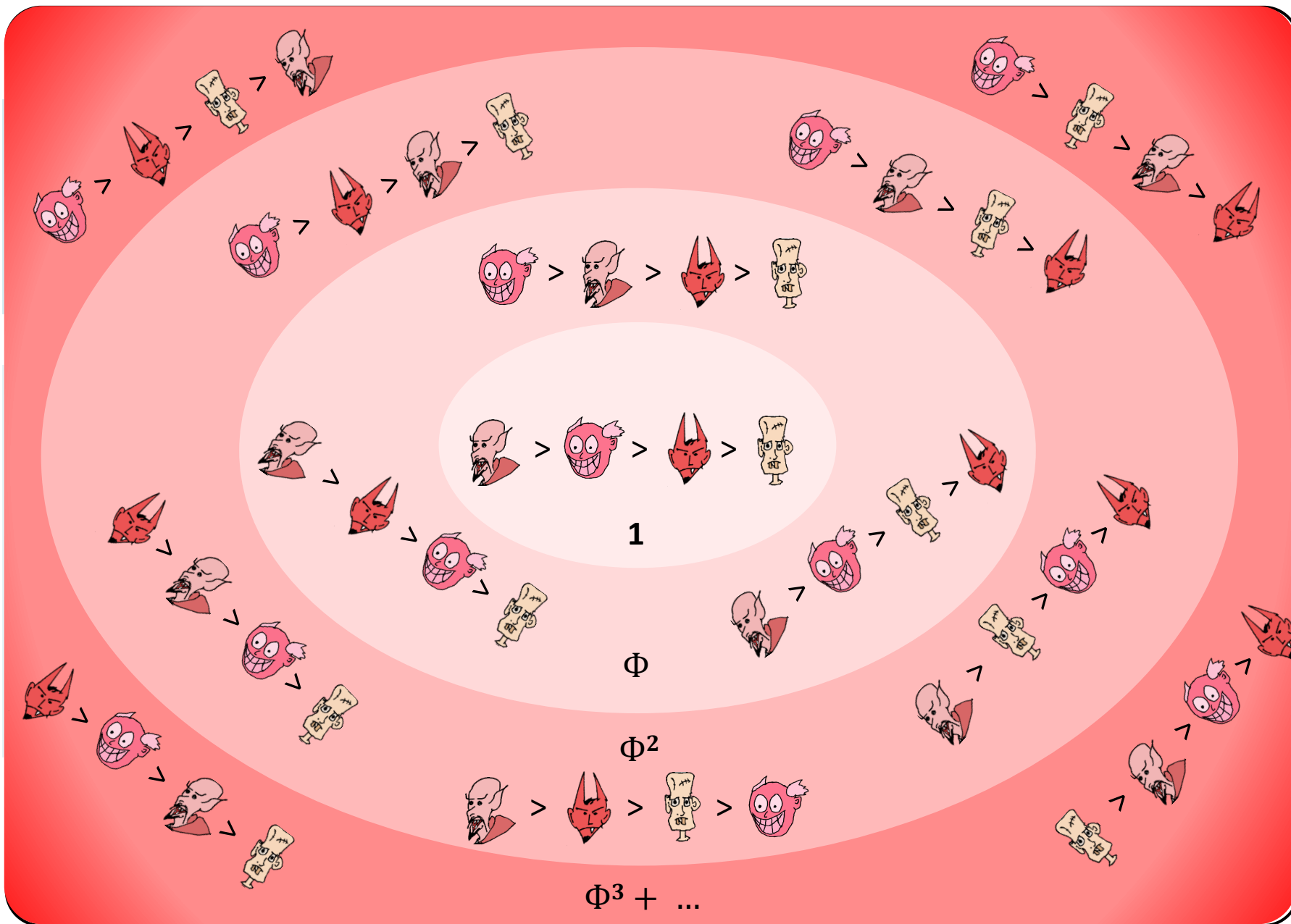
Form an urn of **all possible  $m!$  votes**. To generate a vote:

- 1) Choose a vote from the urn and add it to your election
- 2) **Return** the vote to the urn, together with  **$\alpha \cdot m!$  copies**.

**Mallows Model:** Choose a center vote  $u$ . The probability of generating vote  $v$  is:

$$\frac{1}{Z} \Phi^{swap(u,v)}$$

(There are some algorithms that generate votes from this distribution... effectively.)



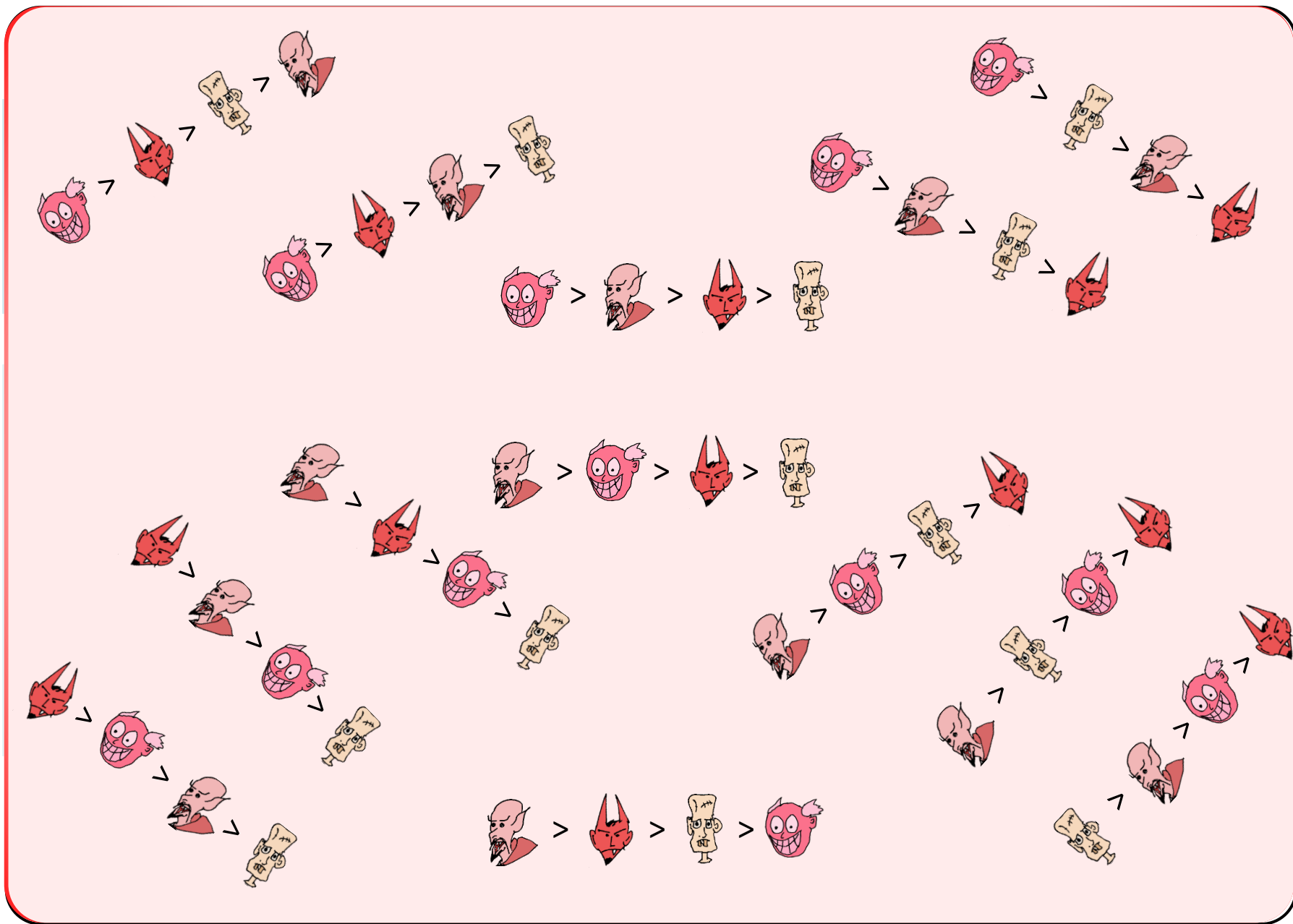
**Impartial Culture (IC):**  
 Every preference order  
 comes with the same  
 probability (a.k.a.  
**uniform distribution**)

**Mallows Model:** Choose a  
 center vote  $u$ . The  
 probability of generating  
 vote  $v$  is:

$$\frac{1}{Z} \Phi^{\text{swap}(u,v)}$$

There are fast sampling  
 algorithms.

$$\Phi = 1$$



**Impartial Culture (IC):**  
 Every preference order  
 comes with the same  
 probability (a.k.a.  
**uniform distribution**)

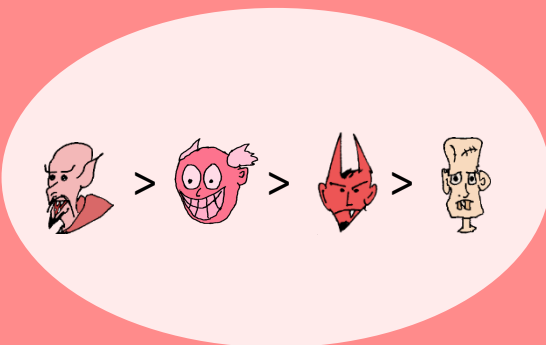
**Mallows Model:** Choose a  
 center vote  $u$ . The  
 probability of generating  
 vote  $v$  is:

$$\frac{1}{Z} \Phi^{swap(u,v)}$$

There are fast sampling  
 algorithms.



$$\Phi = 0$$



**Impartial Culture (IC):**  
Every preference order  
comes with the same  
probability (a.k.a.  
**uniform distribution**)

**Mallows Model:** Choose a  
center vote  $u$ . The  
probability of generating  
vote  $v$  is:

$$\frac{1}{Z} \Phi^{swap(u,v)}$$

There are fast sampling  
algorithms.

# Statistical Cultures

**Impartial Culture (IC):** Every preference order comes with the same probability (a.k.a. **uniform distribution**)

**Polya-Eggenberger Urn Model:**

Form an urn of **all possible  $m!$  votes**. To generate a vote:

- 1) Choose a vote from the urn and add it to your election
- 2) **Return** the vote to the urn, together with  **$\alpha \cdot m!$  copies**.

**Mallows Model:** Choose a center vote  $u$ . The probability of generating vote  $v$  is:

$$\frac{1}{Z} \Phi^{swap(u,v)}$$

(There are some algorithms that generate votes from this distribution... effectively.)

So... what do these models actually do?



15 minutes

# Microscope View of Statistical Cultures

# Swap Distance

$$d_{\text{swap}}(\text{panda} > \text{whale} > \text{cat}, \text{whale} > \text{cat} > \text{panda}) = 2$$

Number of swaps of adjacent candidates needed to transform one preference order into the other

# Swap Distance

$$d_{\text{swap}}(\text{panda} > \text{whale} > \text{cat}, \text{whale} > \text{cat} > \text{panda}) = 2$$

Number of swaps of adjacent candidates needed to transform one preference order into the other

# Swap Distance

$$d_{\text{swap}}(\text{whale} > \text{panda} > \text{cat}, \text{whale} > \text{cat} > \text{panda}) = 0$$

Number of swaps of adjacent candidates needed to transform one preference order into the other

# Swap Distance

$$d_{\text{swap}}(\text{whale} > \text{cat} > \text{panda}, \text{whale} > \text{cat} > \text{panda}) = 0$$

Number of swaps of adjacent candidates needed to transform one preference order into the other

Election microscope:

1. Generate an election from a statistical culture
2. Compute swap distances between all pairs of votes
3. Represent each vote as a dot in 2D space, so that Euclidean distances are similar to the swap distances → **map!**



# The Map Idea

We have some objects:

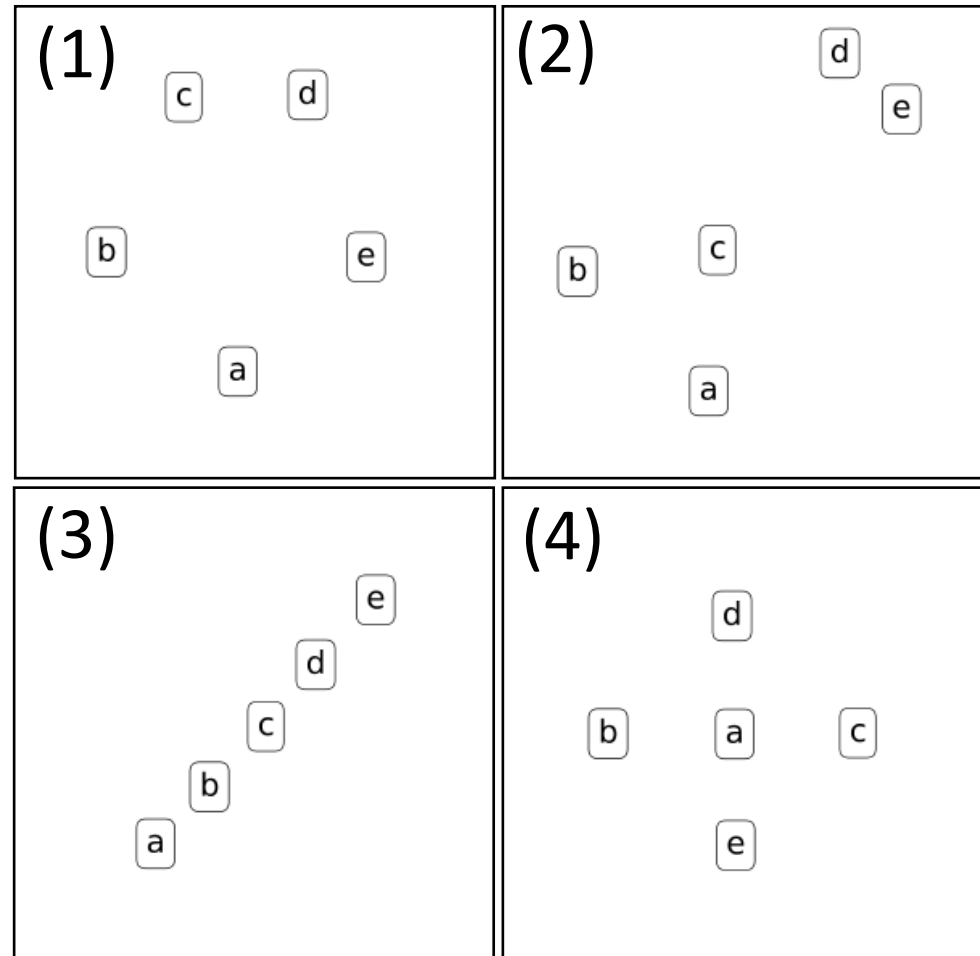
$a, b, c, d, e$

We (somehow) know the distances between each pair

—	$a$	$b$	$c$	$d$	$e$
$a$	—	2	2	4	4
$b$	2	—	2	4	4
$c$	2	2	—	3	3
$d$	4	4	3	—	1
$e$	4	4	3	1	—

(a) Distance Matrix

Can we arrange them in 2D space?



# The Map Idea

We have some objects:

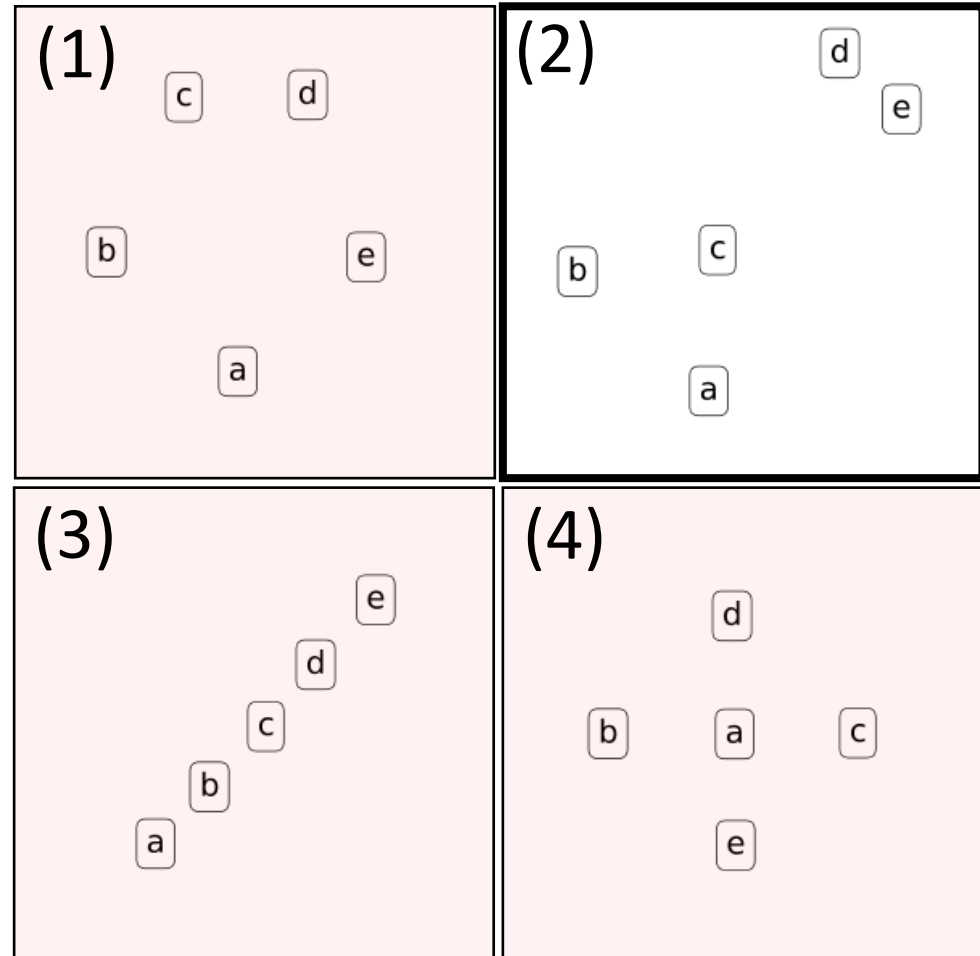
$a, b, c, d, e$

We (somehow) know the distances between each pair

—	$a$	$b$	$c$	$d$	$e$
$a$	—	2	2	4	4
$b$	2	—	2	4	4
$c$	2	2	—	3	3
$d$	4	4	3	—	1
$e$	4	4	3	1	—

(a) Distance Matrix

Can we arrange them in 2D space?



# The Map Idea: Sometimes You Fail

Consider objects:

$$z_1, z_2, z_3, \dots, z_{100}$$

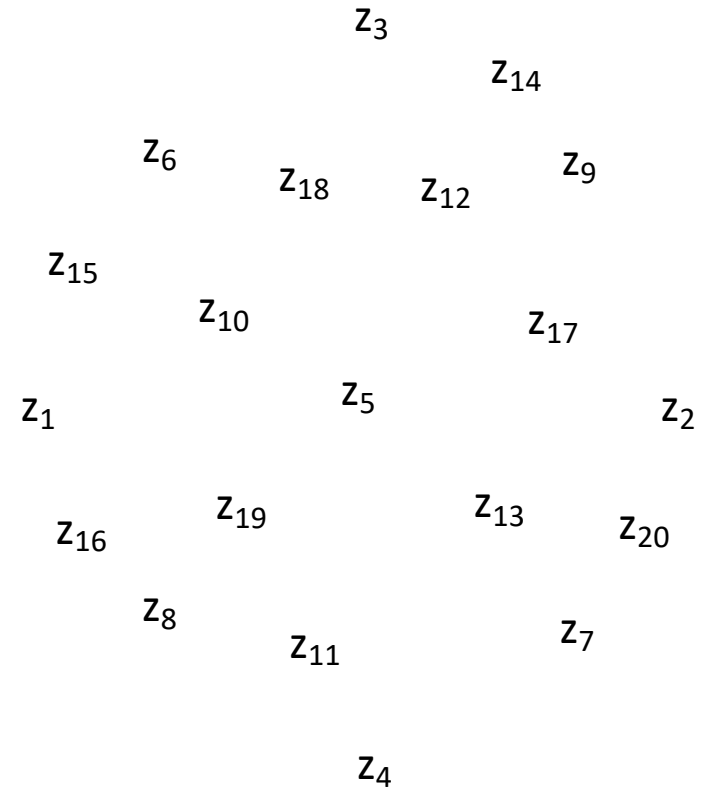
For each  $i, j \in [100]$ , we have:

$$d(z_i, z_j) = 1$$

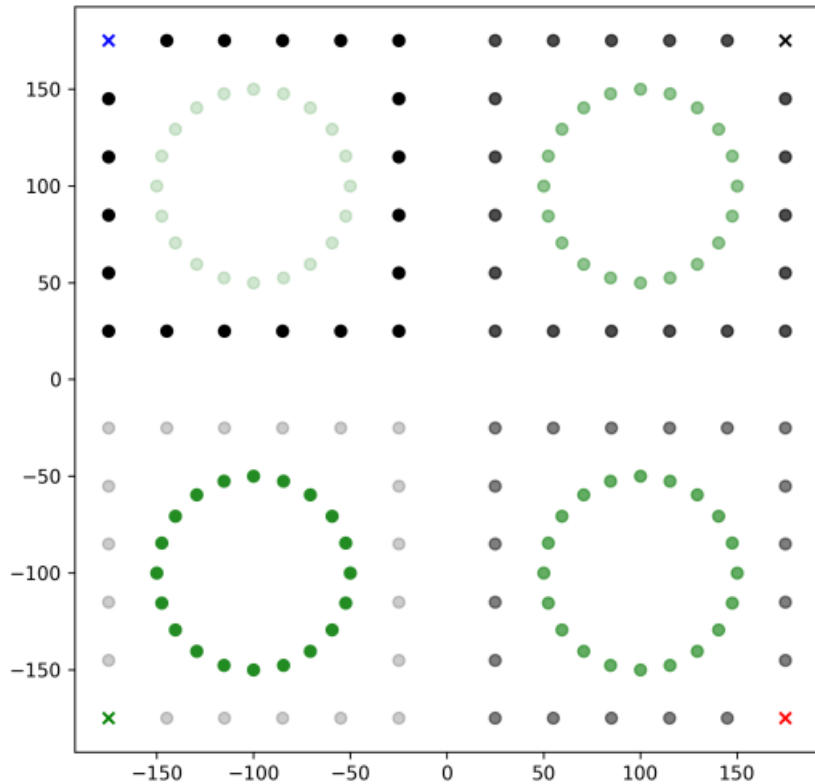
How to arrange these in the  
2D space?

Not much you can do without errors...

But we still do it



# The Map Idea: Computing The Embedding

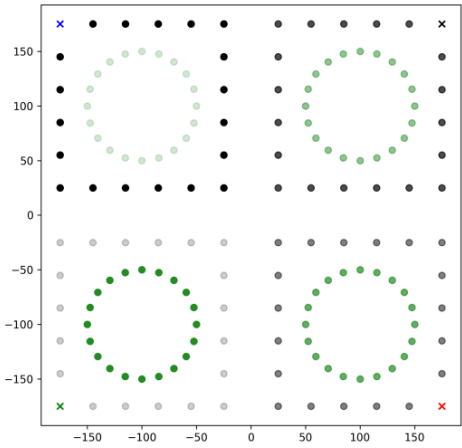


$a_{00}$	$a_{01}$	$a_{02}$	$a_{03}$	$a_{04}$	$a_{05}$	$a_{06}$	$a_{07}$
$a_{10}$	$a_{11}$	$a_{12}$	$a_{13}$	$a_{14}$	$a_{15}$	$a_{16}$	$a_{17}$
$a_{20}$	$a_{21}$	$a_{22}$	$a_{23}$	$a_{24}$	$a_{25}$	$a_{26}$	$a_{27}$
$a_{30}$	$a_{31}$	$a_{32}$	$a_{33}$	$a_{34}$	$a_{35}$	$a_{36}$	$a_{37}$
$a_{40}$	$a_{41}$	$a_{42}$	$a_{43}$	$a_{44}$	$a_{45}$	$a_{46}$	$a_{47}$
$a_{50}$	$a_{51}$	$a_{52}$	$a_{53}$	$a_{54}$	$a_{55}$	$a_{56}$	$a_{57}$
$a_{60}$	$a_{61}$	$a_{62}$	$a_{63}$	$a_{64}$	$a_{65}$	$a_{66}$	$a_{67}$
$a_{70}$	$a_{71}$	$a_{72}$	$a_{73}$	$a_{74}$	$a_{75}$	$a_{76}$	$a_{77}$
$a_{80}$	$a_{81}$	$a_{82}$	$a_{83}$	$a_{84}$	$a_{85}$	$a_{86}$	$a_{87}$
$a_{90}$	$a_{91}$	$a_{92}$	$a_{93}$	$a_{94}$	$a_{95}$	$a_{96}$	$a_{97}$

simple geometric dataset

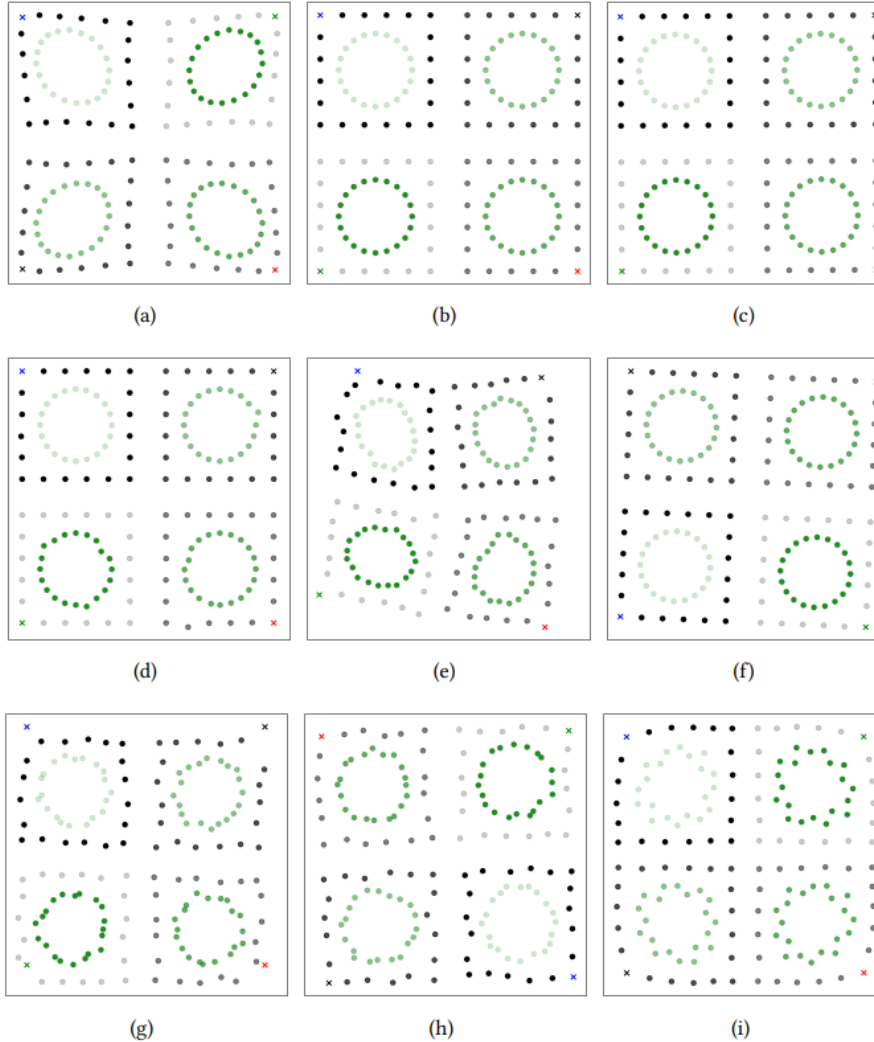
(embedding algorithms only have Euclidean distances of points as inputs)

# The Map Idea: Computing The Embedding



$a_{03}$	$a_{04}$	$a_{05}$	$a_{06}$	$a_{07}$			
$a_{13}$	$a_{14}$	$a_{15}$	$a_{16}$	$a_{17}$			
$a_{23}$	$a_{24}$	$a_{25}$	$a_{26}$	$a_{27}$			
$a_{30}$	$a_{31}$	$a_{32}$	$a_{33}$	$a_{34}$	$a_{35}$	$a_{36}$	$a_{37}$
$a_{40}$	$a_{41}$	$a_{42}$	$a_{43}$	$a_{44}$	$a_{45}$	$a_{46}$	$a_{47}$
$a_{50}$	$a_{51}$	$a_{52}$	$a_{53}$	$a_{54}$	$a_{55}$	$a_{56}$	$a_{57}$
$a_{60}$	$a_{61}$	$a_{62}$	$a_{63}$	$a_{64}$	$a_{65}$	$a_{66}$	$a_{67}$
$a_{70}$	$a_{71}$	$a_{72}$	$a_{73}$	$a_{74}$	$a_{75}$	$a_{76}$	$a_{77}$
$a_{80}$	$a_{81}$	$a_{82}$	$a_{83}$	$a_{84}$	$a_{85}$	$a_{86}$	$a_{87}$
$a_{90}$	$a_{91}$	$a_{92}$	$a_{93}$	$a_{94}$	$a_{95}$	$a_{96}$	$a_{97}$

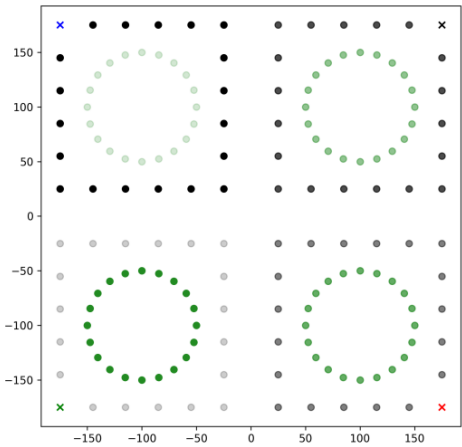
simple geometric dataset  
(embedding algorithms only have Euclidean distances of points as inputs)



Examples of embeddings

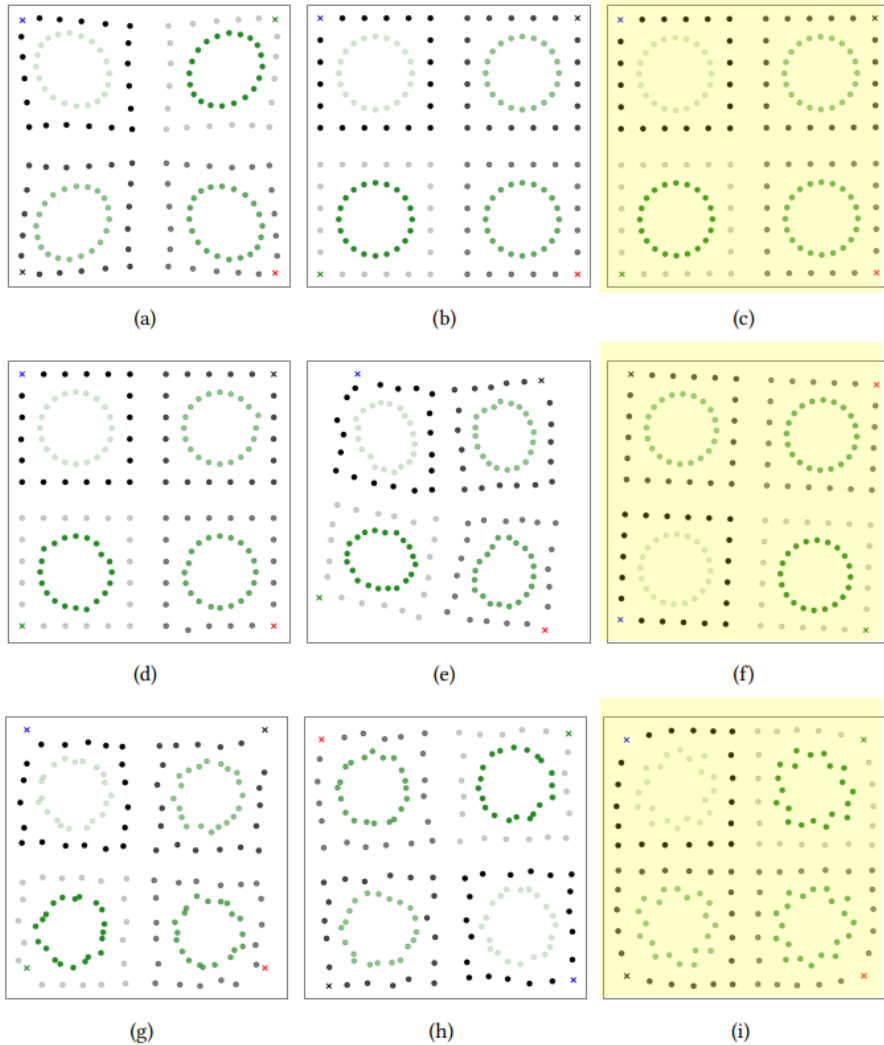
- (a) ISOMAP
- (b) Kamada-Kawai (KK) with positions of corner points fixed
- (c) KK without fixing
- (d) KK with Newton-Raphson + fixing
- (e) KK with Newton-Raphson without fixing
- (f) MDS
- (g) Simulated annealing with fixing
- (h) Simulated annealing without fixing
- (i) Fruchterman-Reingold

# The Map Idea: Computing The Embedding



$a_{03}$	$a_{04}$	$a_{05}$	$a_{06}$	$a_{07}$			
$a_{13}$	$a_{14}$	$a_{15}$	$a_{16}$	$a_{17}$			
$a_{23}$	$a_{24}$	$a_{25}$	$a_{26}$	$a_{27}$			
$a_{30}$	$a_{31}$	$a_{32}$	$a_{33}$	$a_{34}$	$a_{35}$	$a_{36}$	$a_{37}$
$a_{40}$	$a_{41}$	$a_{42}$	$a_{43}$	$a_{44}$	$a_{45}$	$a_{46}$	$a_{47}$
$a_{50}$	$a_{51}$	$a_{52}$	$a_{53}$	$a_{54}$	$a_{55}$	$a_{56}$	$a_{57}$
$a_{60}$	$a_{61}$	$a_{62}$	$a_{63}$	$a_{64}$	$a_{65}$	$a_{66}$	$a_{67}$
$a_{70}$	$a_{71}$	$a_{72}$	$a_{73}$	$a_{74}$	$a_{75}$	$a_{76}$	$a_{77}$
$a_{80}$	$a_{81}$	$a_{82}$	$a_{83}$	$a_{84}$	$a_{85}$	$a_{86}$	$a_{87}$
$a_{90}$	$a_{91}$	$a_{92}$	$a_{93}$	$a_{94}$	$a_{95}$	$a_{96}$	$a_{97}$

simple geometric dataset  
(embedding algorithms only have Euclidean distances of points as inputs)



Examples of embeddings

- (a) ISOMAP
- (b) Kamada-Kawai (KK) with positions of corner points fixed
- (c) KK without fixing**
- (d) KK with Newton-Rhapson + fixing
- (e) KK with Newton-Rhapson without fixing
- (f) MDS**
- (g) Simulated annealing with fixing
- (h) Simulated annealing without fixing
- (i) Fruchterman-Reingold**

# Swap Distance

$$d_{\text{swap}}(\text{whale} > \text{cat} > \text{panda}, \text{whale} > \text{cat} > \text{panda}) = 0$$

Number of swaps of adjacent candidates needed to transform one preference order into the other

Election microscope:

1. Generate an election from a statistical culture
2. Compute swap distances between all pairs of votes
3. Represent each vote as a dot in 2D space, so that Euclidean distances are similar to the swap distances → **map!**

**Impartial Culture (IC):** Every preference order comes with the same probability (a.k.a. **uniform distribution**)

**Polya-Eggenberger Urn Model:** Form an urn of **all possible  $m!$  votes**. To generate a vote:  
 1) Choose a vote from the urn and add it to your election  
 2) **Return** the vote to the urn, together with  **$\alpha \cdot m!$  copies**.

**Mallows Model:** Choose a center vote  $u$ . The probability of generating vote  $v$  is:

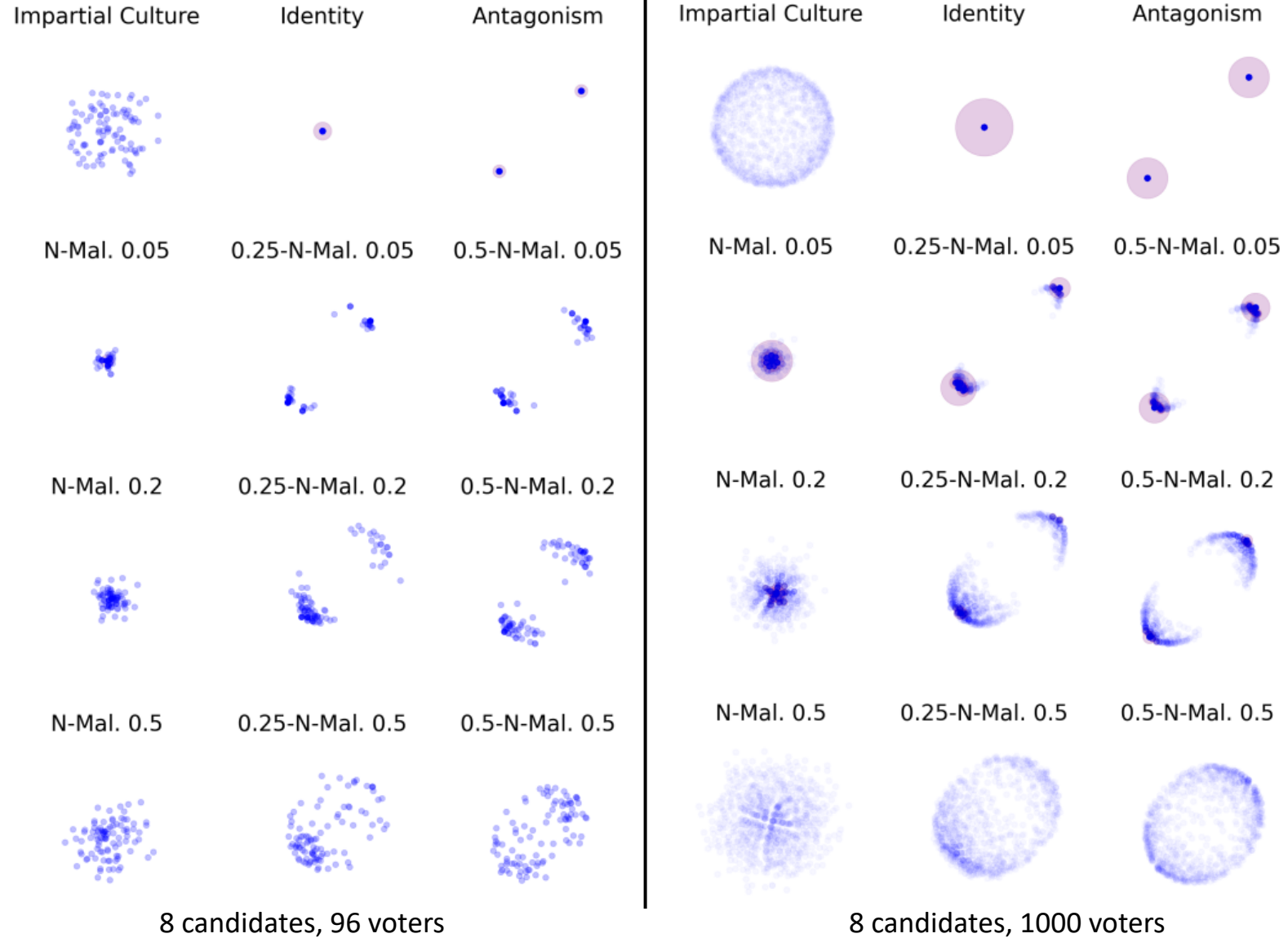
$$\frac{1}{Z} \Phi^{swap(u,v)}$$

(There are some algorithms that generate votes from this distribution... effectively.)

**Election microscope:**

1. Generate an election from a statistical culture
2. Compute swap distances between all pairs of votes
3. Represent each vote as a dot in 2D space, so that Euclidean distances are similar to the swap distances (MDS)

# Microscope





**Impartial Culture (IC):** Every preference order comes with the same probability (a.k.a. **uniform distribution**)

**Polya-Eggenberger Urn Model:** Form an urn of **all possible  $m!$  votes**. To generate a vote:

- 1) Choose a vote from the urn and add it to your election
- 2) **Return** the vote to the urn, together with  $\alpha \cdot m!$  **copies**.

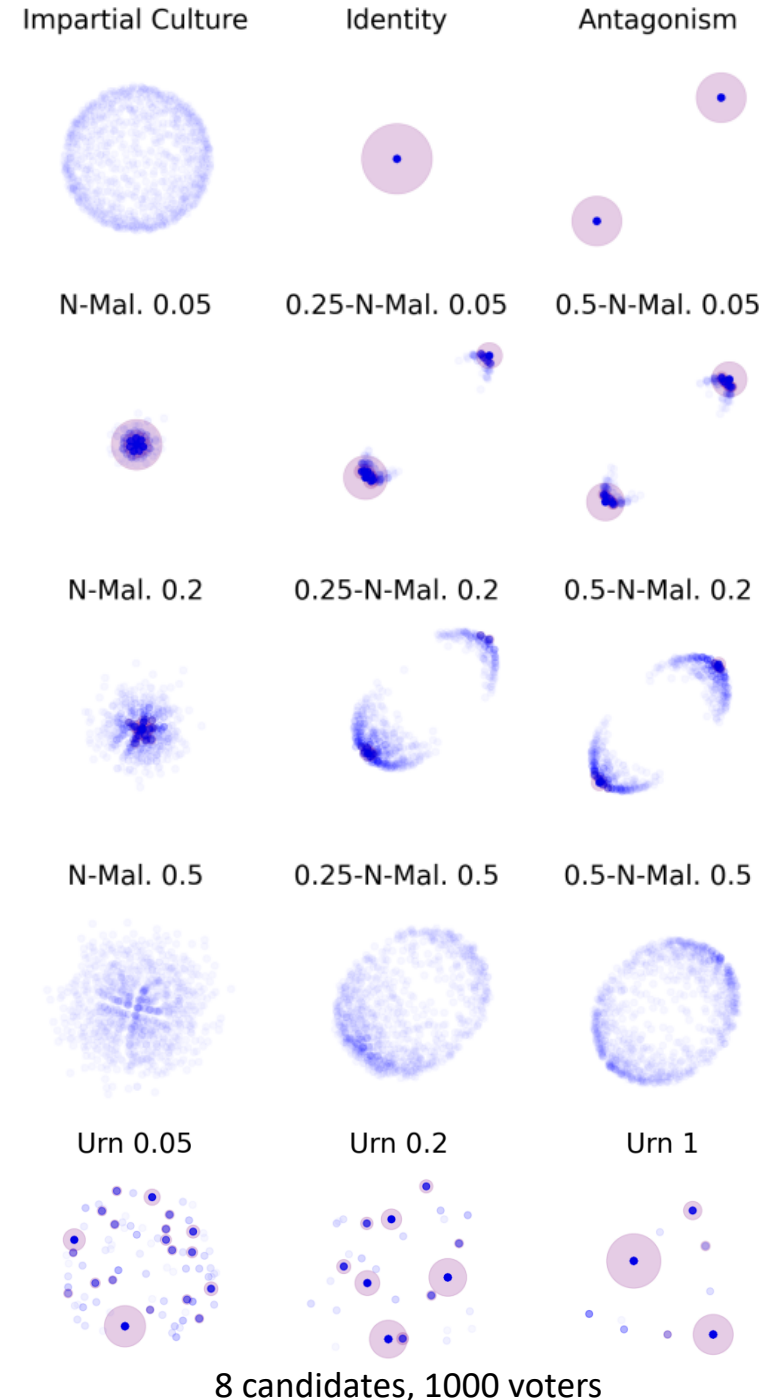
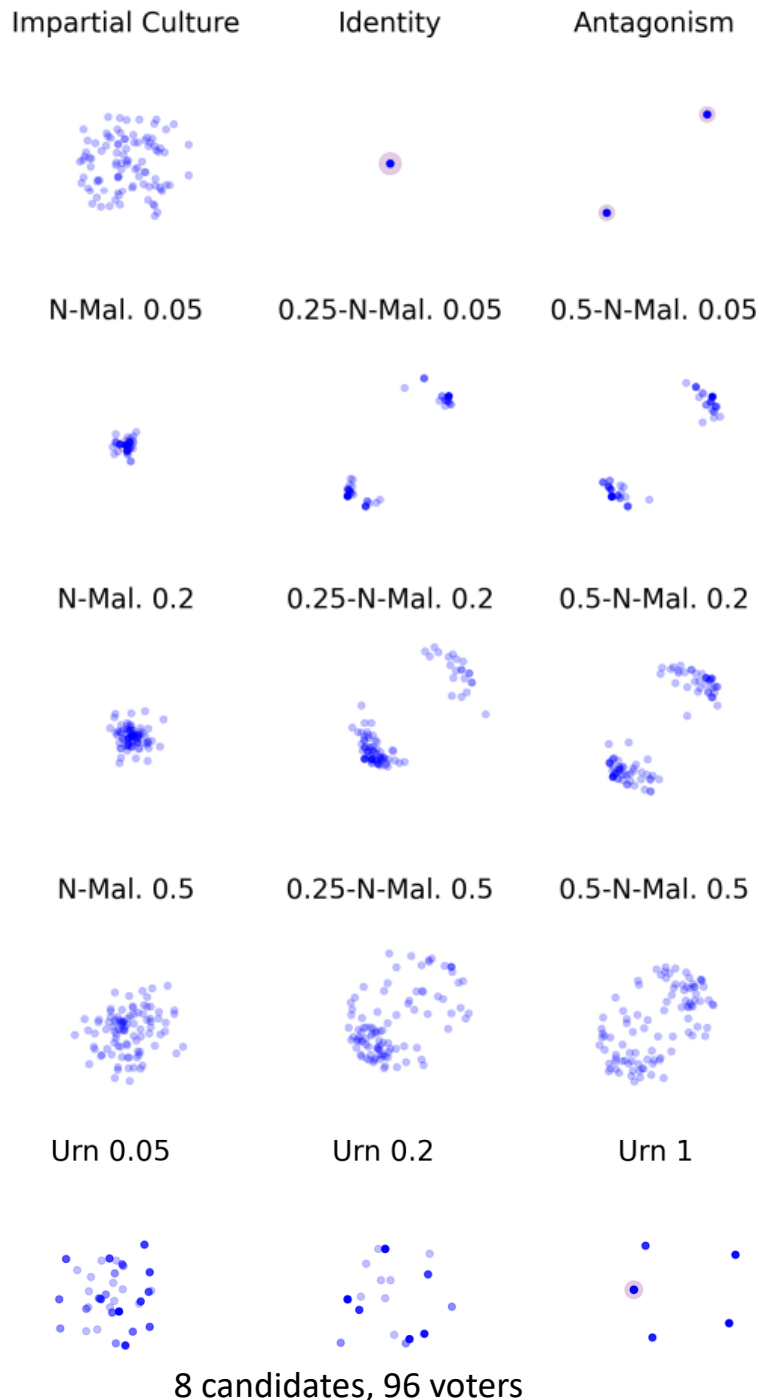
**Mallows Model:** Choose a center vote  $u$ . The probability of generating vote  $v$  is:

$$\frac{1}{Z} \Phi^{swap(u,v)}$$

(There are some algorithms that generate votes from this distribution... effectively.)

**Election microscope:**

1. Generate an election from a statistical culture
2. Compute swap distances between all pairs of votes
3. Represent each vote as a dot in 2D space, so that Euclidean distances are similar to the swap distances (MDS)



**Impartial Culture (IC):** Every preference order comes with the same probability (a.k.a. **uniform distribution**)

**Polya-Eggenberger Urn Model:** Form an urn of **all possible  $m!$  votes**. To generate a vote:

- 1) Choose a vote from the urn and add it to your election
- 2) **Return** the vote to the urn, together with  **$\alpha \cdot m!$  copies**.

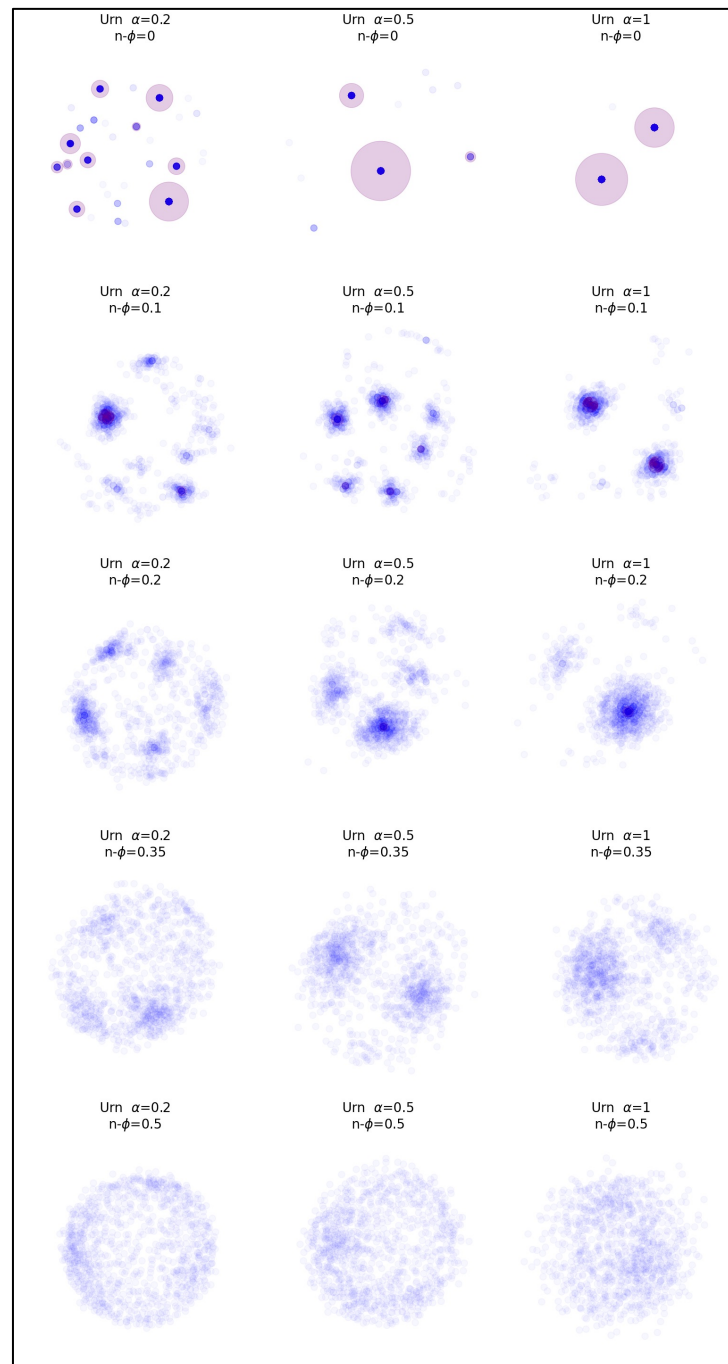
**Mallows Model:** Choose a center vote  $u$ . The probability of generating vote  $v$  is:

$$\frac{1}{Z} \Phi^{swap(u,v)}$$

(There are some algorithms that generate votes from this distribution... effectively.)

### Election microscope:

1. Generate an election from a statistical culture
2. Compute swap distances between all pairs of votes
3. Represent each vote as a dot in 2D space, so that Euclidean distances are similar to the swap distances (MDS)



**Urn-Mallows Model:** First generate an election according to the urn model and then replace each vote  $v$  with one generated using Mallows model, with  $v$  as the center vote.

### Comparison to real-life elections:

Sushi contains preferences about sushi types. Grenoble and Irish are political elections





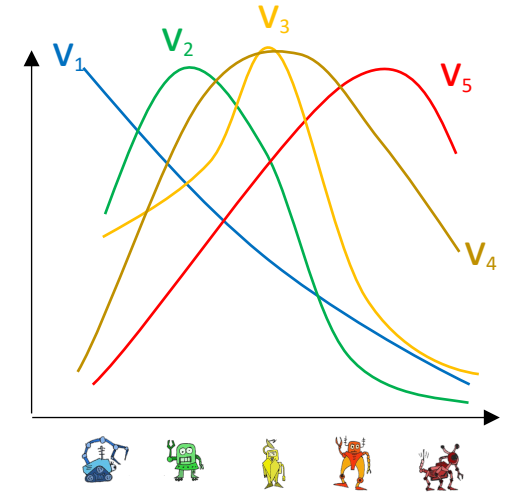
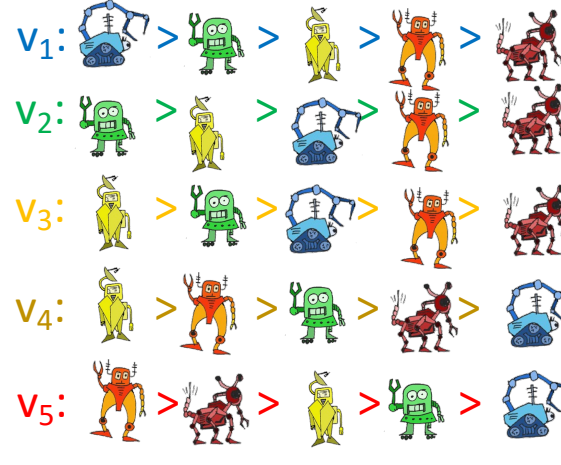
10 minutes

# Restricted Domains

# Restricted Domains

**Single-Peaked (SP):** Fix a societal axis, e.g., the following ordering of the candidates. Every single-peaked vote for this axis satisfies the property that „for each t, the top t candidates form an interval on the axis).

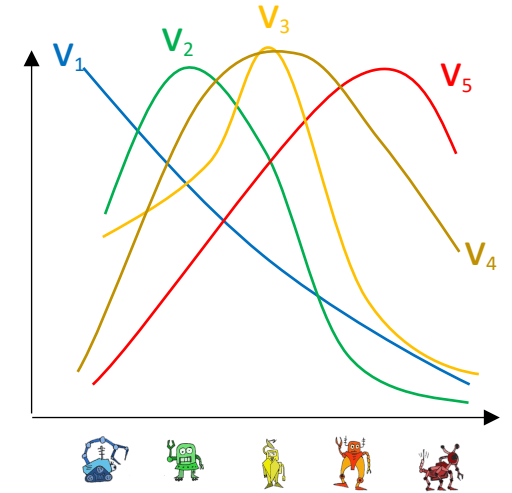
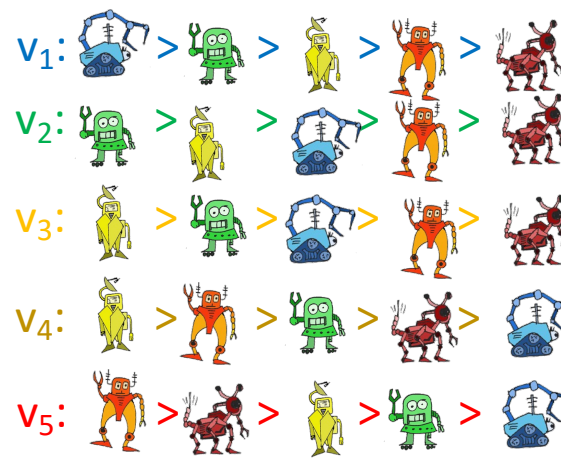
single-peakedness



# Restricted Domains

**Single-Peaked (SP):** Fix a societal axis, e.g., the following ordering of the candidates. Every single-peaked vote for this axis satisfies the property that „for each t, the top t candidates form an interval on the axis).

single-peakedness



## Conitzer model (top-down)

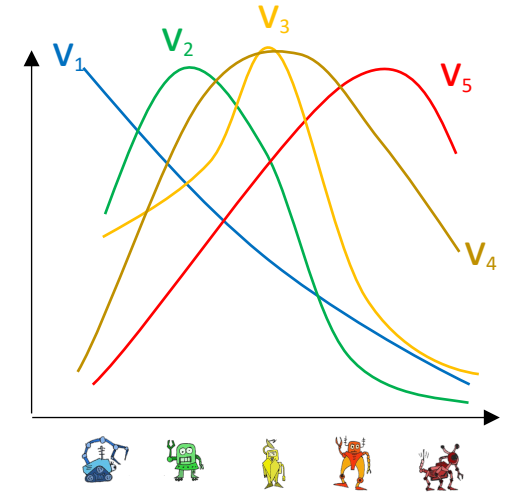
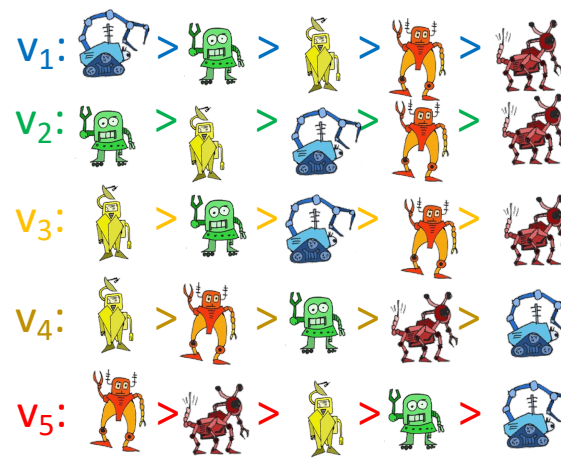
$$1/n * 1/2 * 1/2 * 1 * 1$$



# Restricted Domains

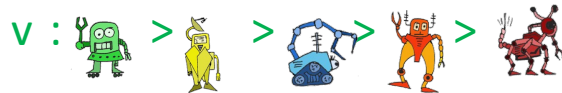
**Single-Peaked (SP):** Fix a societal axis, e.g., the following ordering of the candidates. Every single-peaked vote for this axis satisfies the property that „for each t, the top t candidates form an interval on the axis).

single-peakedness



## Conitzer model (top-down)

$$1/5 * 1/2 * 1/2 * 1 * 1$$



## Walsh model (bottom-up)

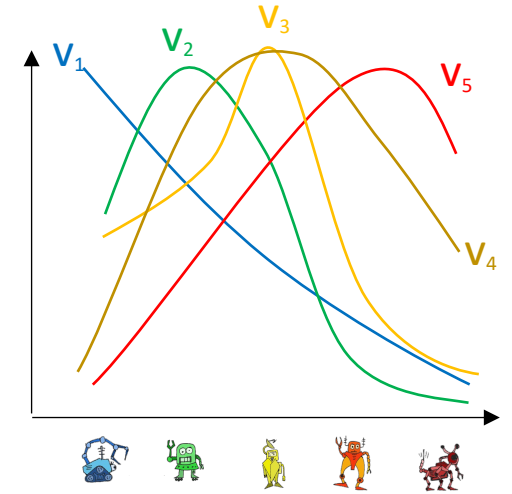
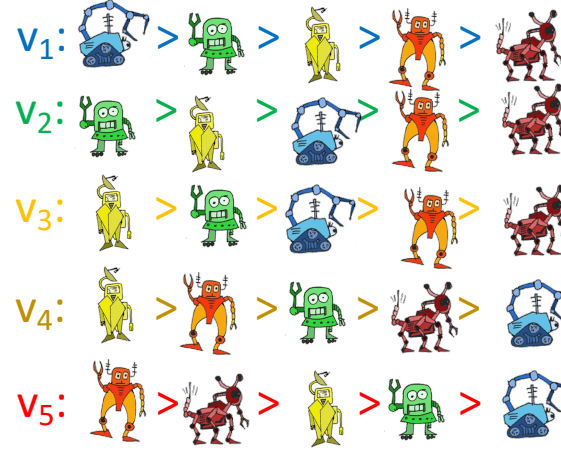
Uniform distribution

# Restricted Domains

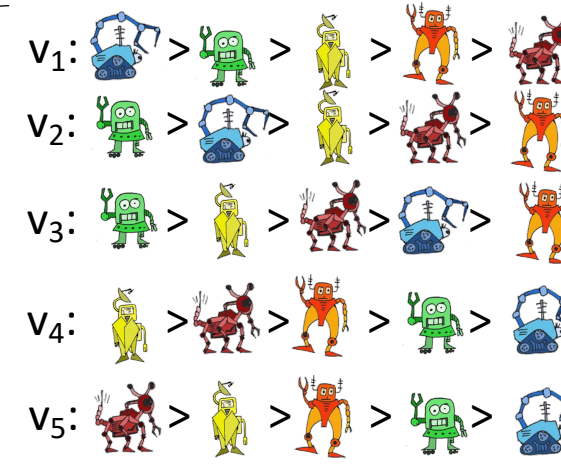
**Single-Peaked (SP):** Fix a societal axis, e.g., the following ordering of the candidates. Every single-peaked vote for this axis satisfies the property that „for each  $t$ , the top  $t$  candidates form an interval on the axis).

**Single-Crossing:** Order voters so going from top to bottom, each pair of candidates crosses at most once.

single-peakedness



single-crossingness



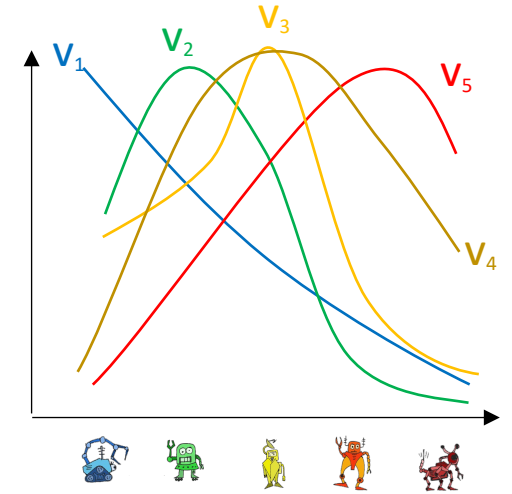
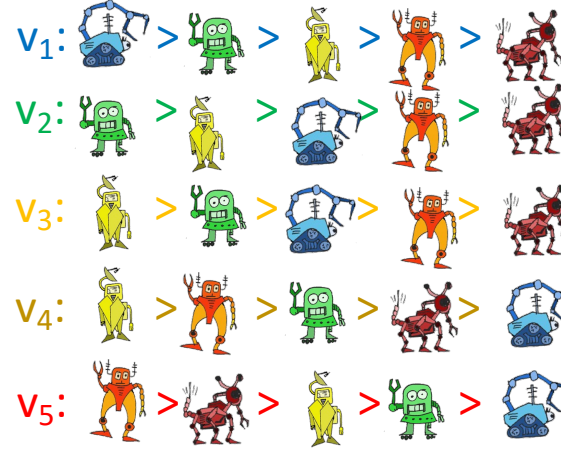


# Restricted Domains

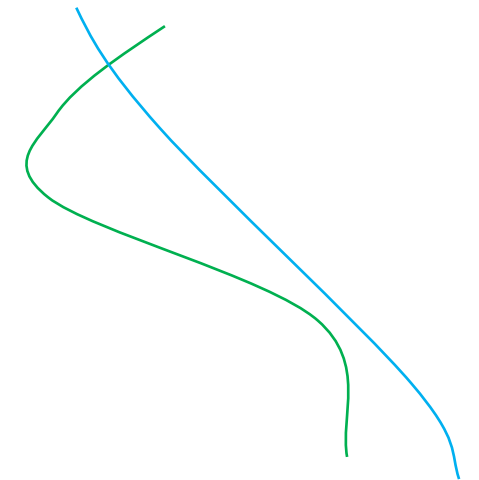
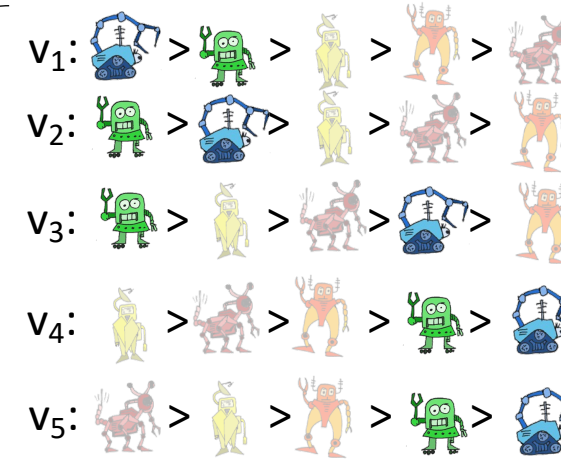
**Single-Peaked (SP):** Fix a societal axis, e.g., the following ordering of the candidates. Every single-peaked vote for this axis satisfies the property that „for each  $t$ , the top  $t$  candidates form an interval on the axis).

**Single-Crossing:** Order voters so going from top to bottom, each pair of candidates crosses at most once.

single-peakedness



single-crossingness



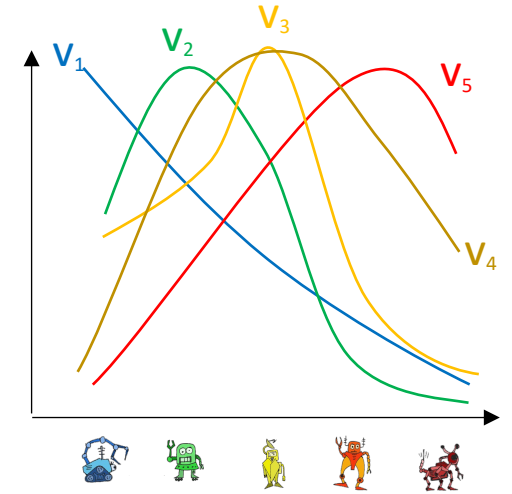
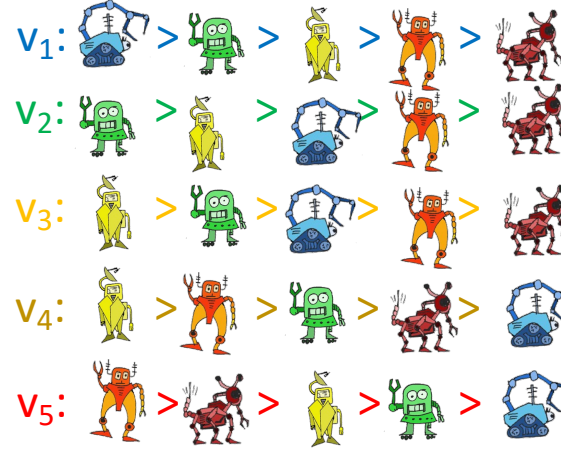


# Restricted Domains

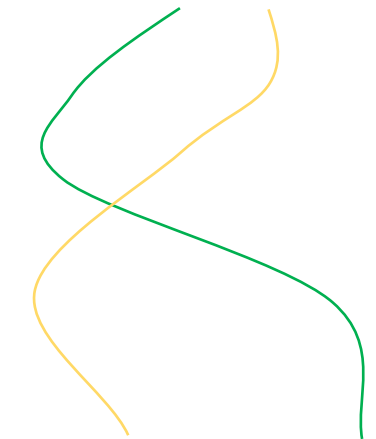
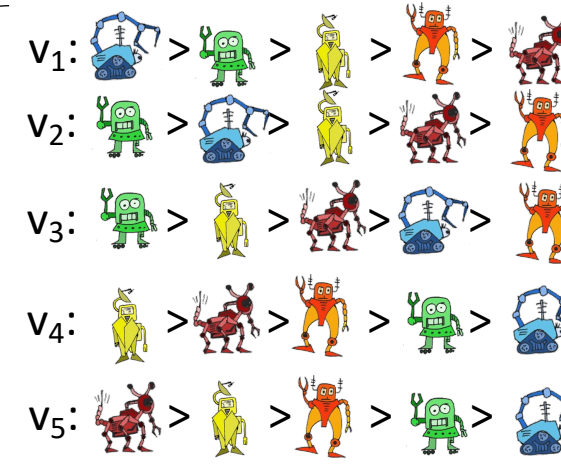
**Single-Peaked (SP):** Fix a societal axis, e.g., the following ordering of the candidates. Every single-peaked vote for this axis satisfies the property that „for each  $t$ , the top  $t$  candidates form an interval on the axis).

**Single-Crossing:** Order voters so going from top to bottom, each pair of candidates crosses at most once.

single-peakedness



single-crossingness

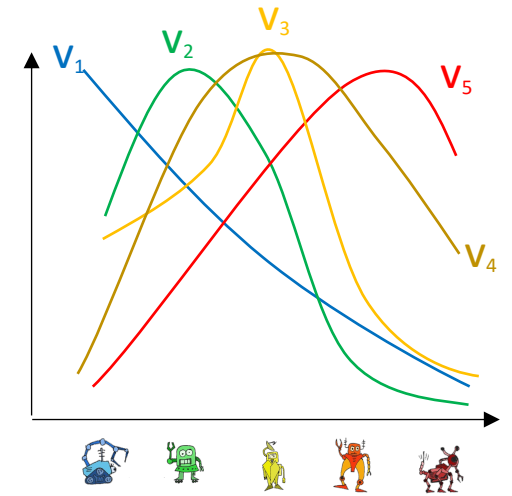
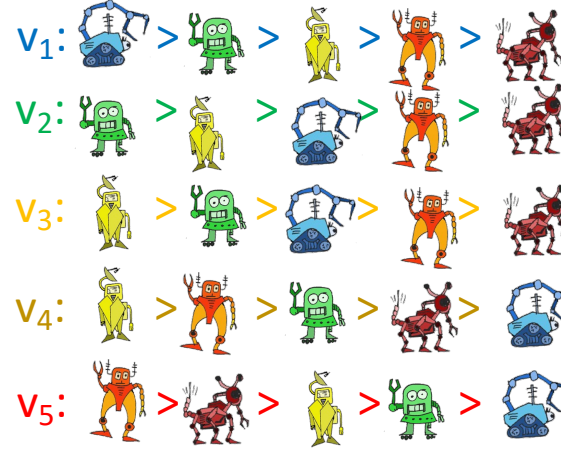


# Restricted Domains

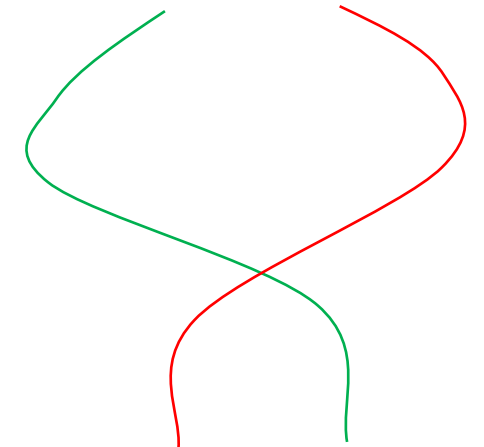
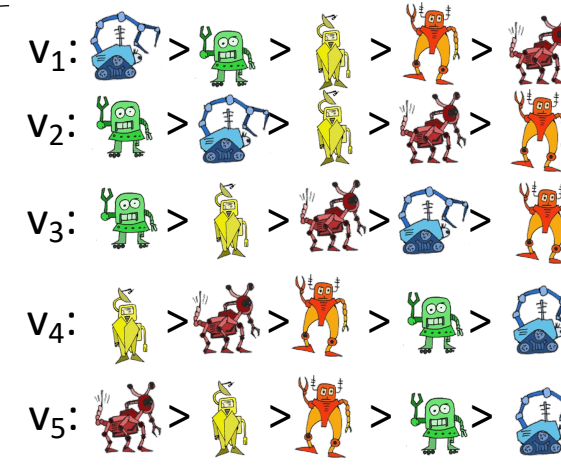
**Single-Peaked (SP):** Fix a societal axis, e.g., the following ordering of the candidates. Every single-peaked vote for this axis satisfies the property that „for each  $t$ , the top  $t$  candidates form an interval on the axis).

**Single-Crossing:** Order voters so going from top to bottom, each pair of candidates crosses at most once.

single-peakedness



single-crossingness

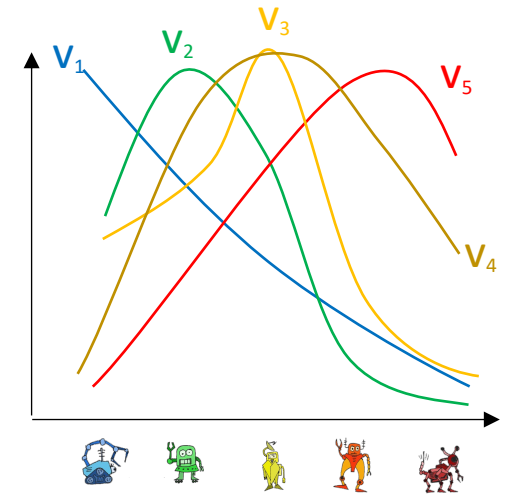
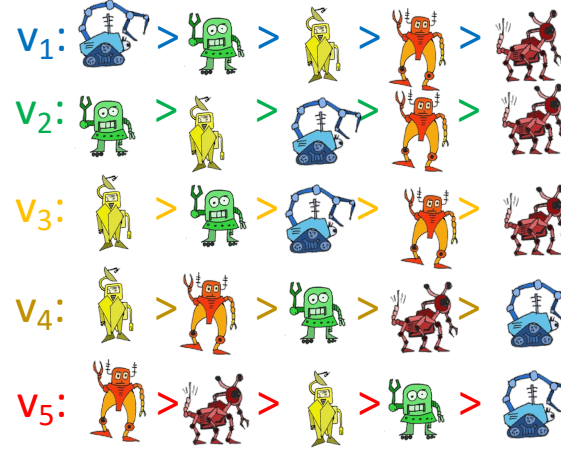


# Restricted Domains

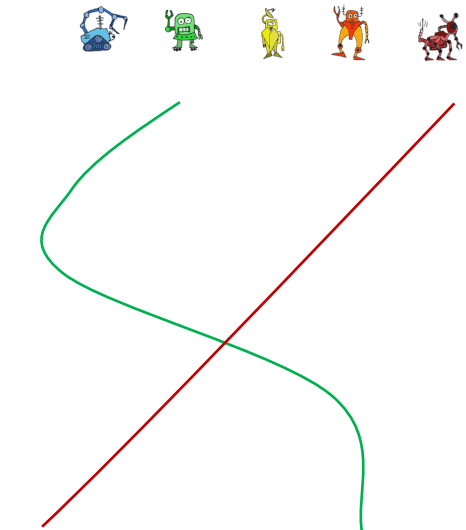
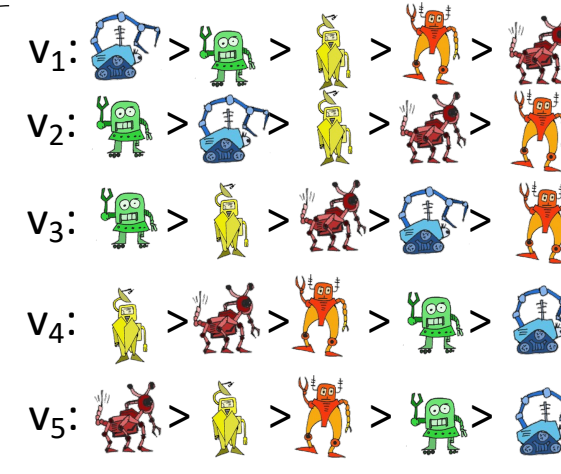
**Single-Peaked (SP):** Fix a societal axis, e.g., the following ordering of the candidates. Every single-peaked vote for this axis satisfies the property that „for each  $t$ , the top  $t$  candidates form an interval on the axis).

**Single-Crossing:** Order voters so going from top to bottom, each pair of candidates crosses at most once.

single-peakedness

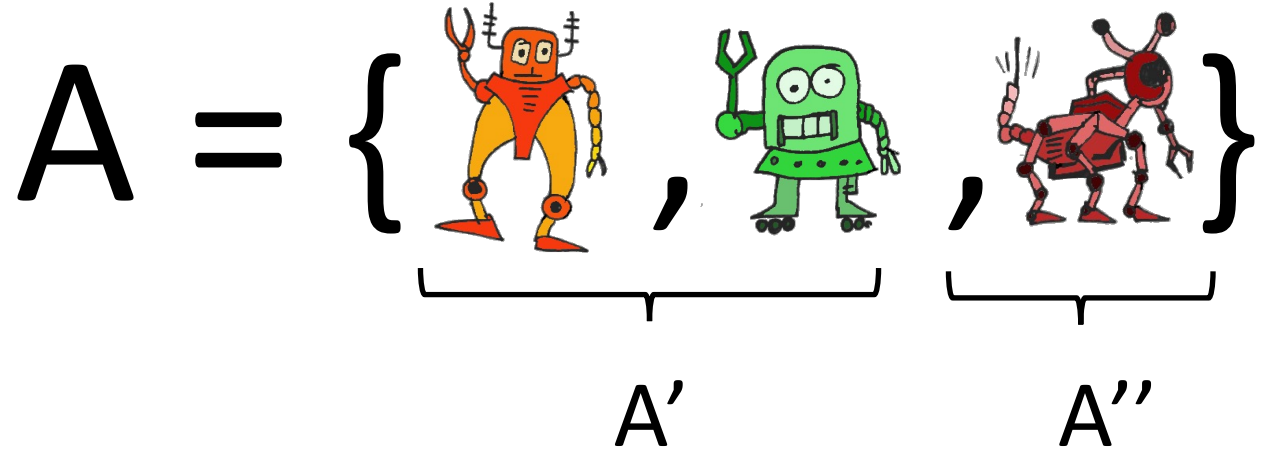


single-crossingness



# Group-Separable Preferences

A profile is group-separable if each subset  $A$ ,  $|A| \geq 2$ , of candidates can be partitioned into  $A'$  and  $A''$  so that each voter prefers all members of one to all members of the other



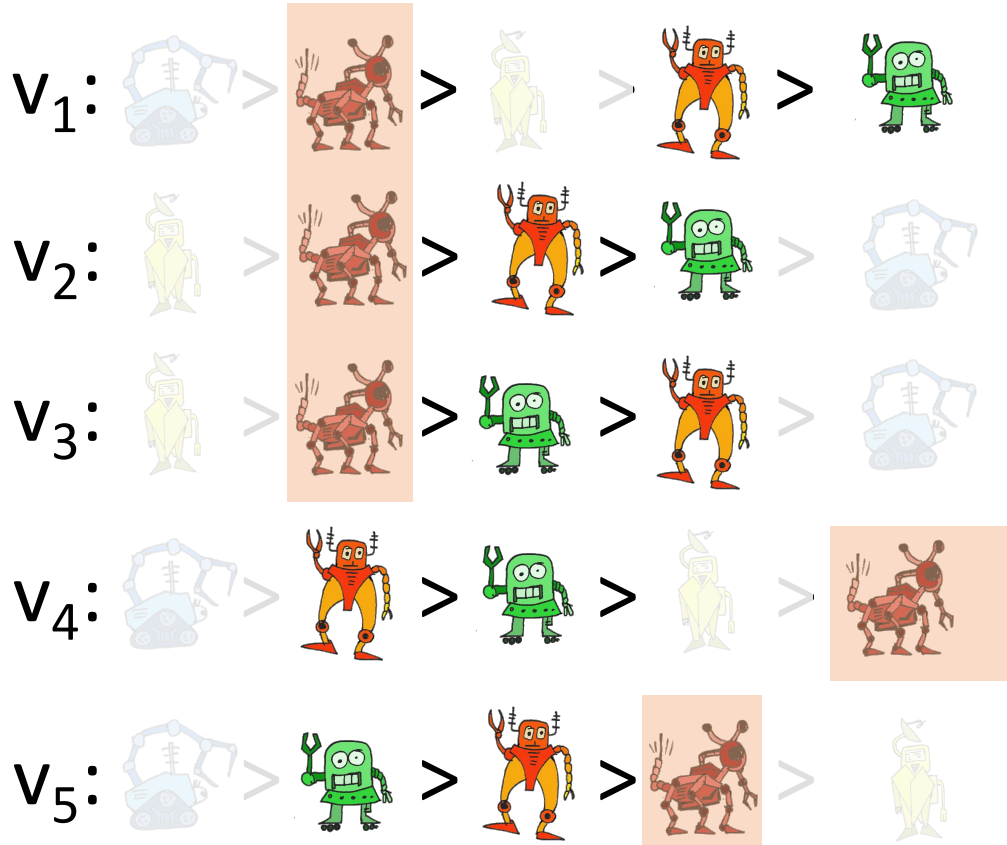
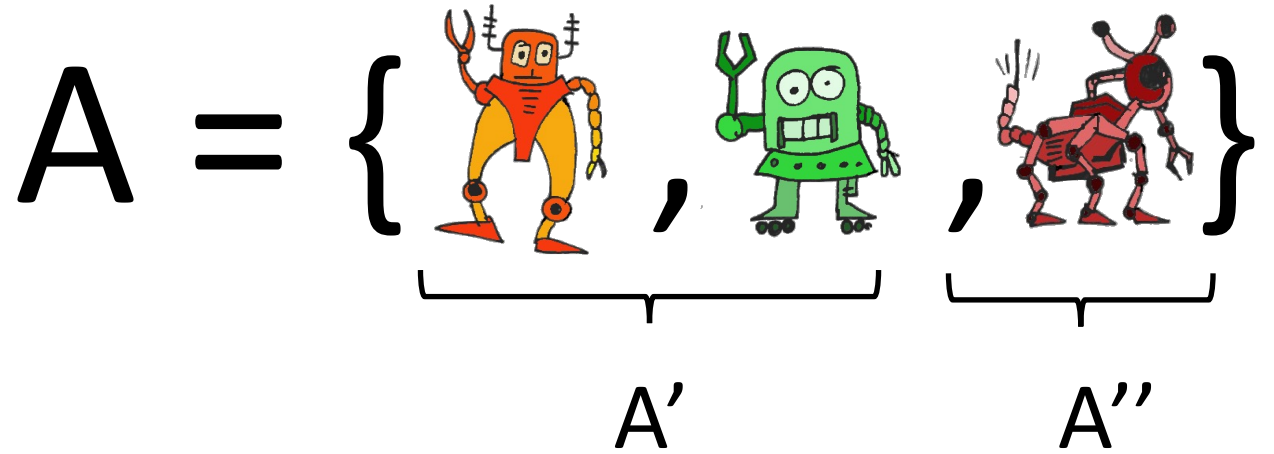
- $V_1$ : > > > >
- $V_2$ : > > > >
- $V_3$ : > > > >
- $V_4$ : > > > >
- $V_5$ : > > > >

K. Inada, A Note on the Simple Majority Decision Rule, *Econometrica*, 1964.

K. Inada, The Simple Majority Decision Rule, *Econometrica*, 1969

# Group-Separable Preferences

A profile is group-separable if each subset  $A$ ,  $|A| \geq 2$ , of candidates can be partitioned into  $A'$  and  $A''$  so that each voter prefers all members of one to all members of the other

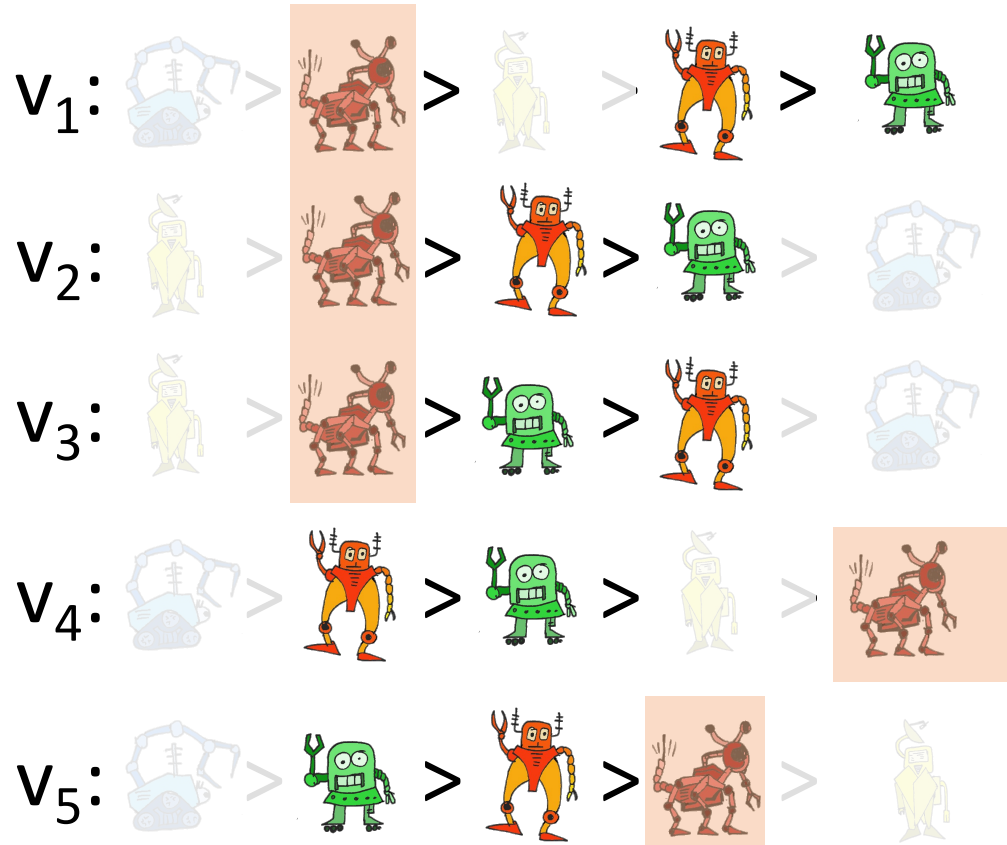


K. Inada, A Note on the Simple Majority Decision Rule, *Econometrica*, 1964.

K. Inada, The Simple Majority Decision Rule, *Econometrica*, 1969

# Group-Separable Preferences

A profile is group-separable if each subset  $A$ ,  $|A| \geq 2$ , of candidates can be partitioned into  $A'$  and  $A''$  so that each voter prefers all members of one to all members of the other



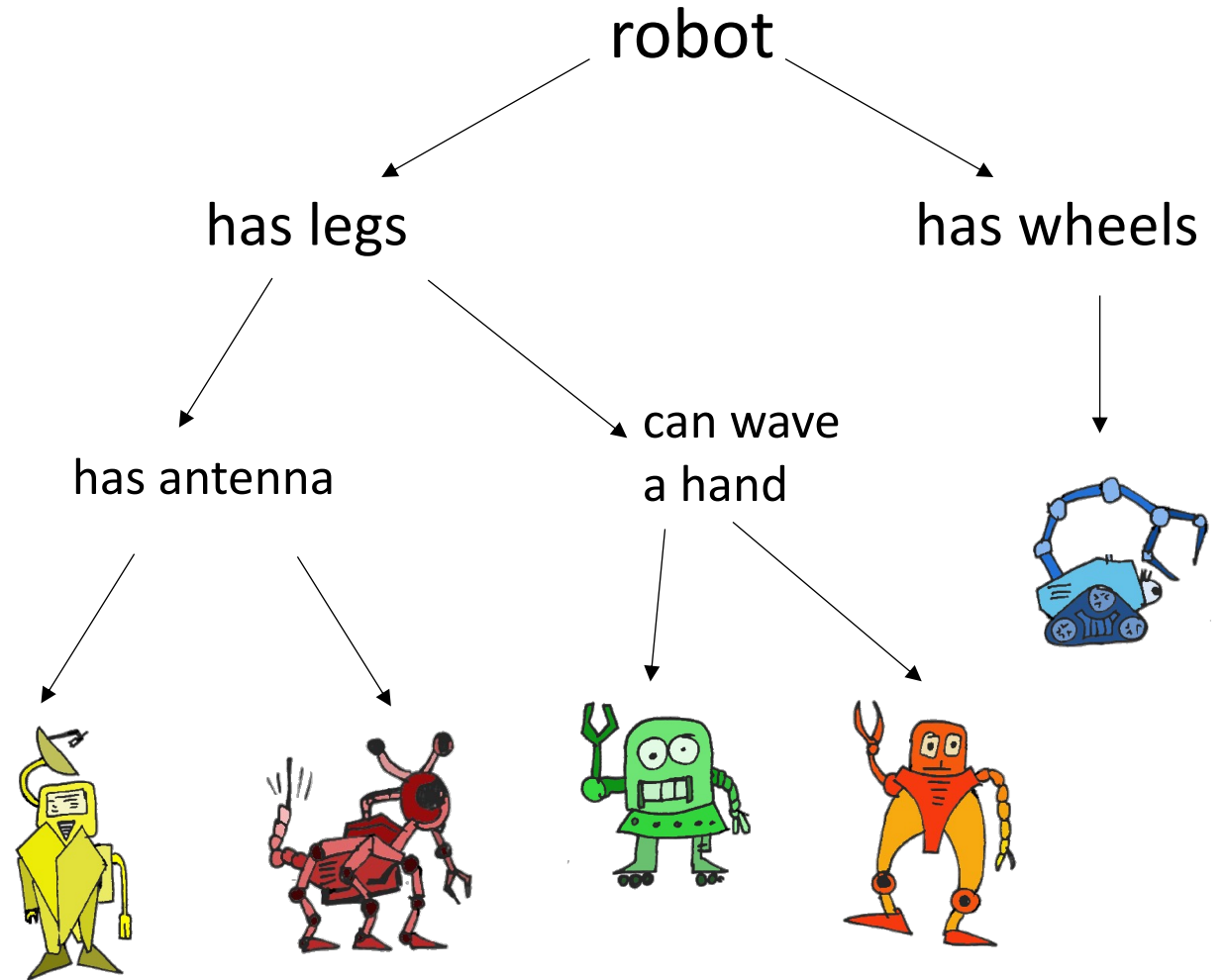
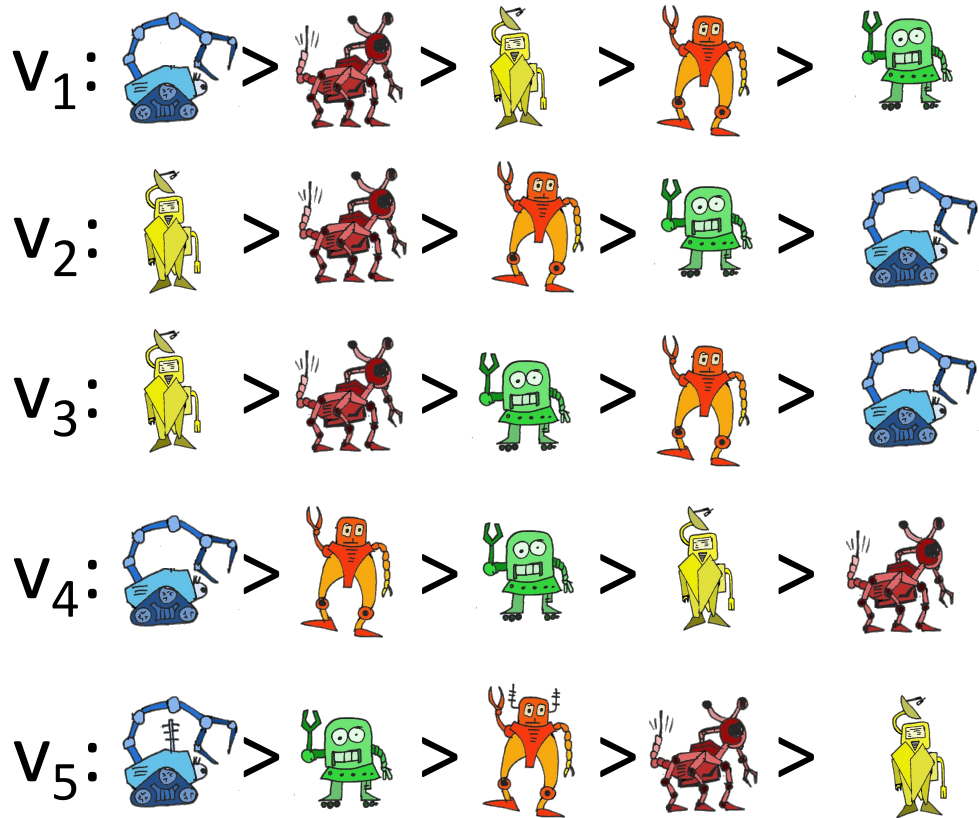
K. Inada, A Note on the Simple Majority Decision Rule, *Econometrica*, 1964.

K. Inada, The Simple Majority Decision Rule, *Econometrica*, 1969



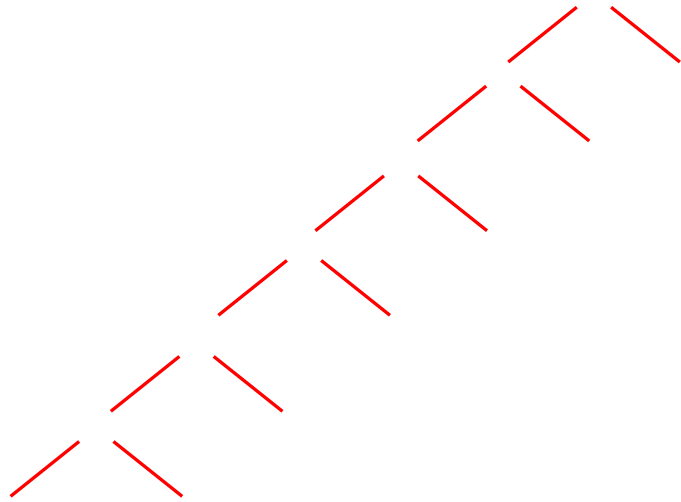
# Group-Separable Preferences

A profile is group-separable if each subset  $A$ ,  $|A| \geq 2$ , of candidates can be partitioned into  $A'$  and  $A''$  so that each voter prefers all members of one to all members of the other

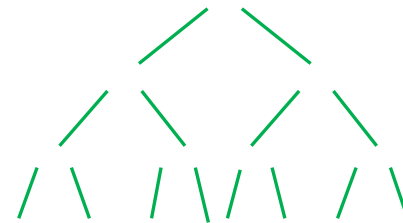


# Group-Separable Preferences

Caterpillar Trees



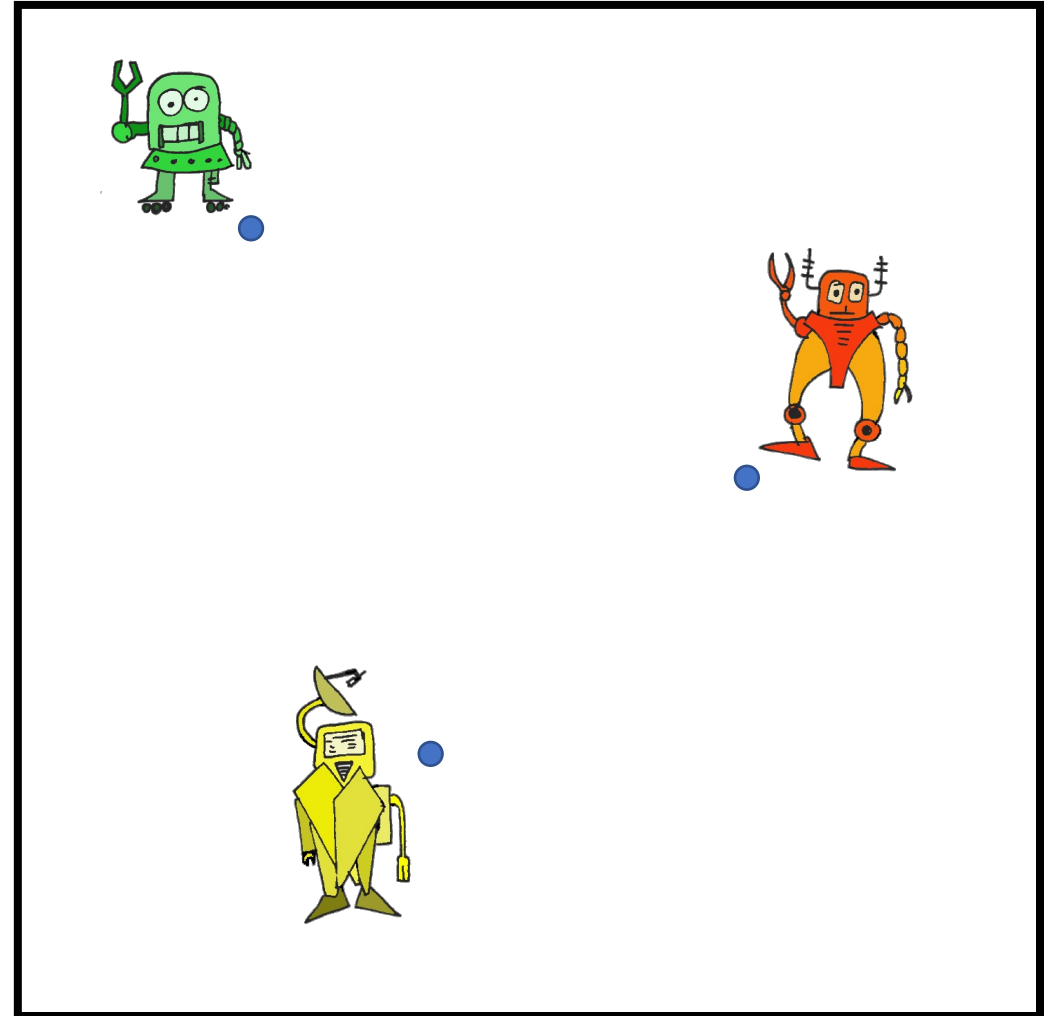
Balanced Trees





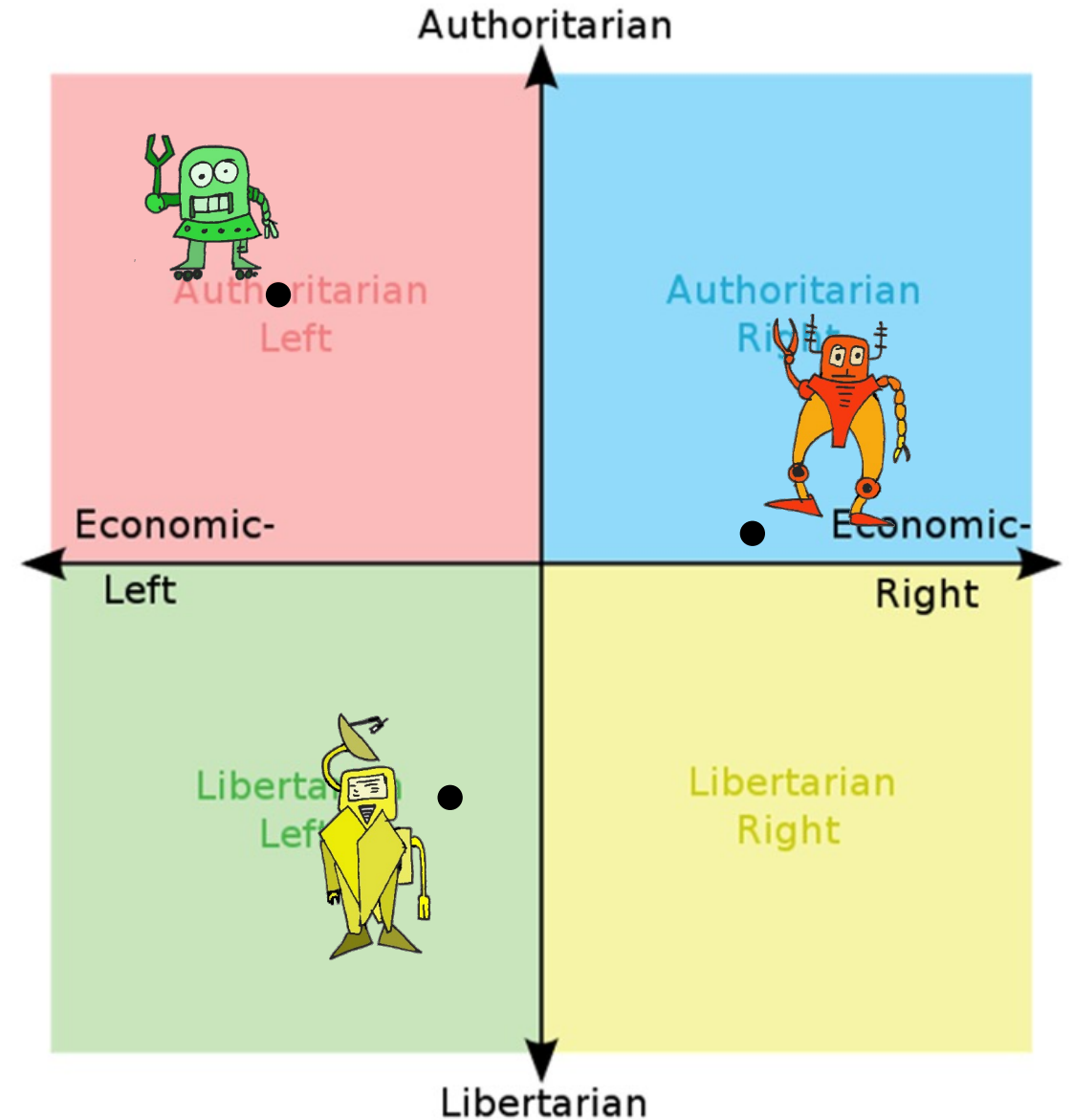
# Euclidean Preferences

**Euclidean Model:** Choose **points** for the voters and candidates from **Euclidean space**  $\mathbb{R}^t$ . Voter  $v$  prefers candidate  $x$  to  $y$  if  $x$ 's point is **closer** to  $v$  than  $y$ 's.



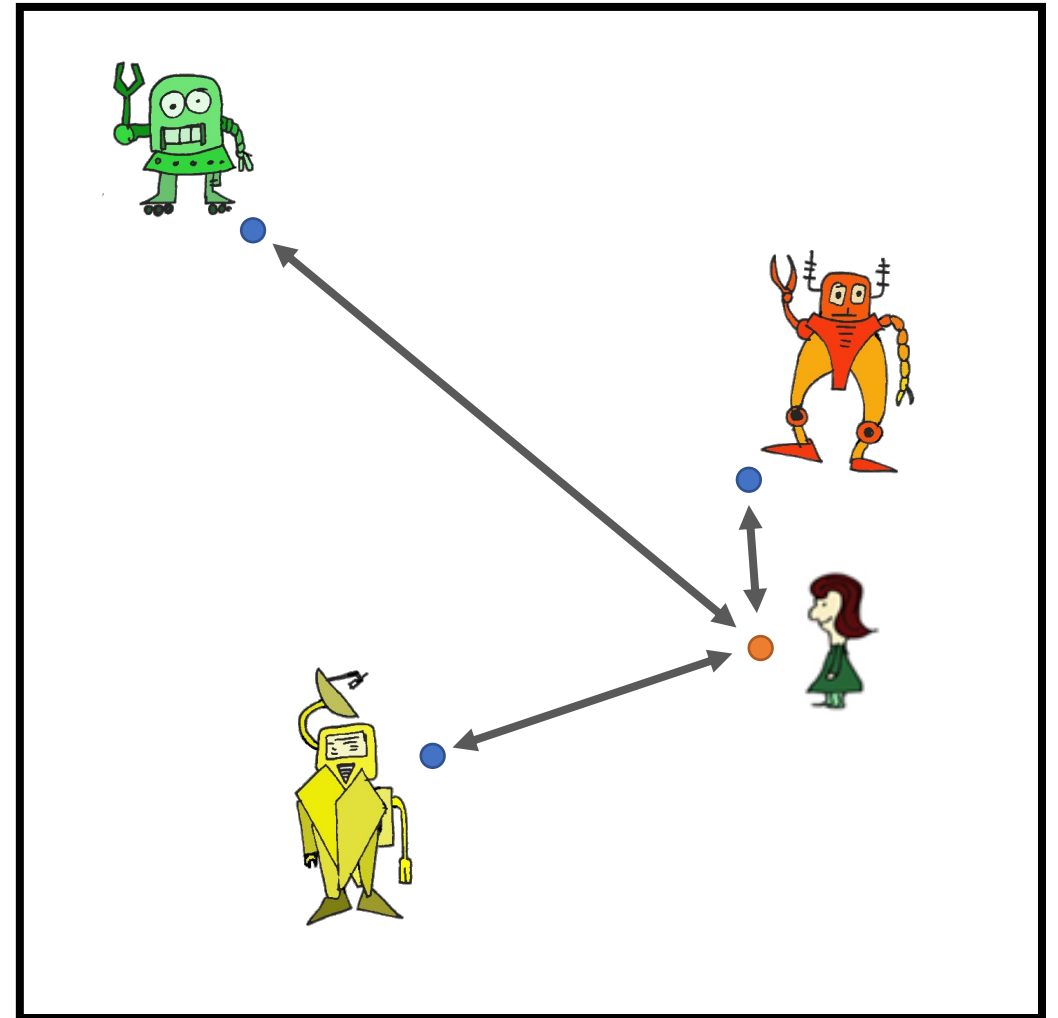
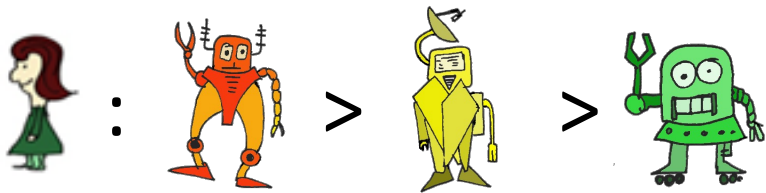
# Euclidean Preferences

**Euclidean Model:** Choose **points** for the voters and candidates from **Euclidean space**  $\mathbb{R}^t$ . Voter  $v$  prefers candidate  $x$  to  $y$  if  $x$ 's point is **closer** to  $v$  than  $y$ 's.



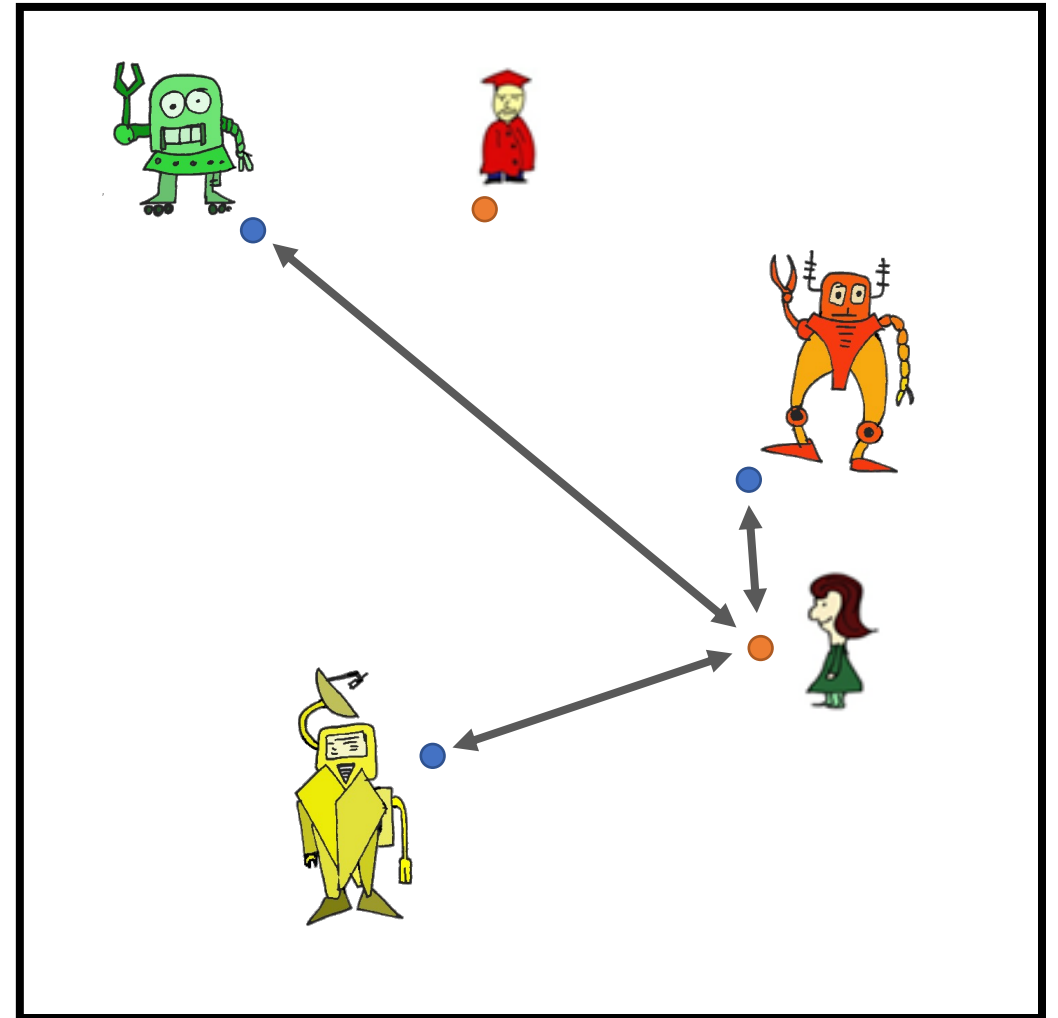
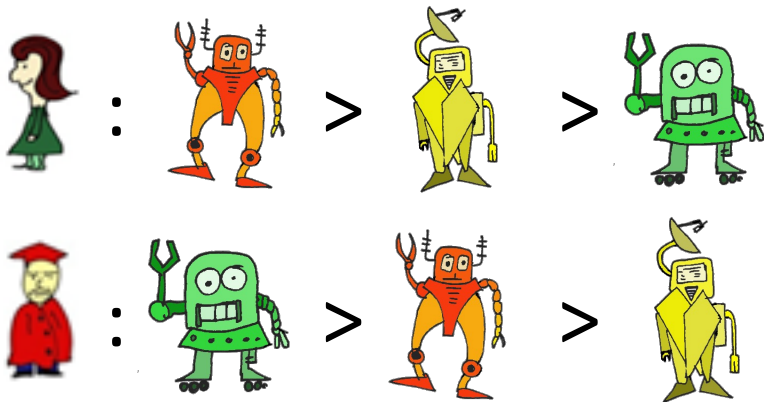
# Euclidean Preferences

**Euclidean Model:** Choose **points** for the voters and candidates from **Euclidean space**  $\mathbb{R}^t$ . Voter  $v$  prefers candidate  $x$  to  $y$  if  $x$ 's point is **closer** to  $v$  than  $y$ 's.



# Euclidean Preferences

**Euclidean Model:** Choose **points** for the voters and candidates from **Euclidean space**  $\mathbb{R}^t$ . Voter  $v$  prefers candidate  $x$  to  $y$  if  $x$ 's point is **closer** to  $v$  than  $y$ 's.



# Microscope of Structured Domains

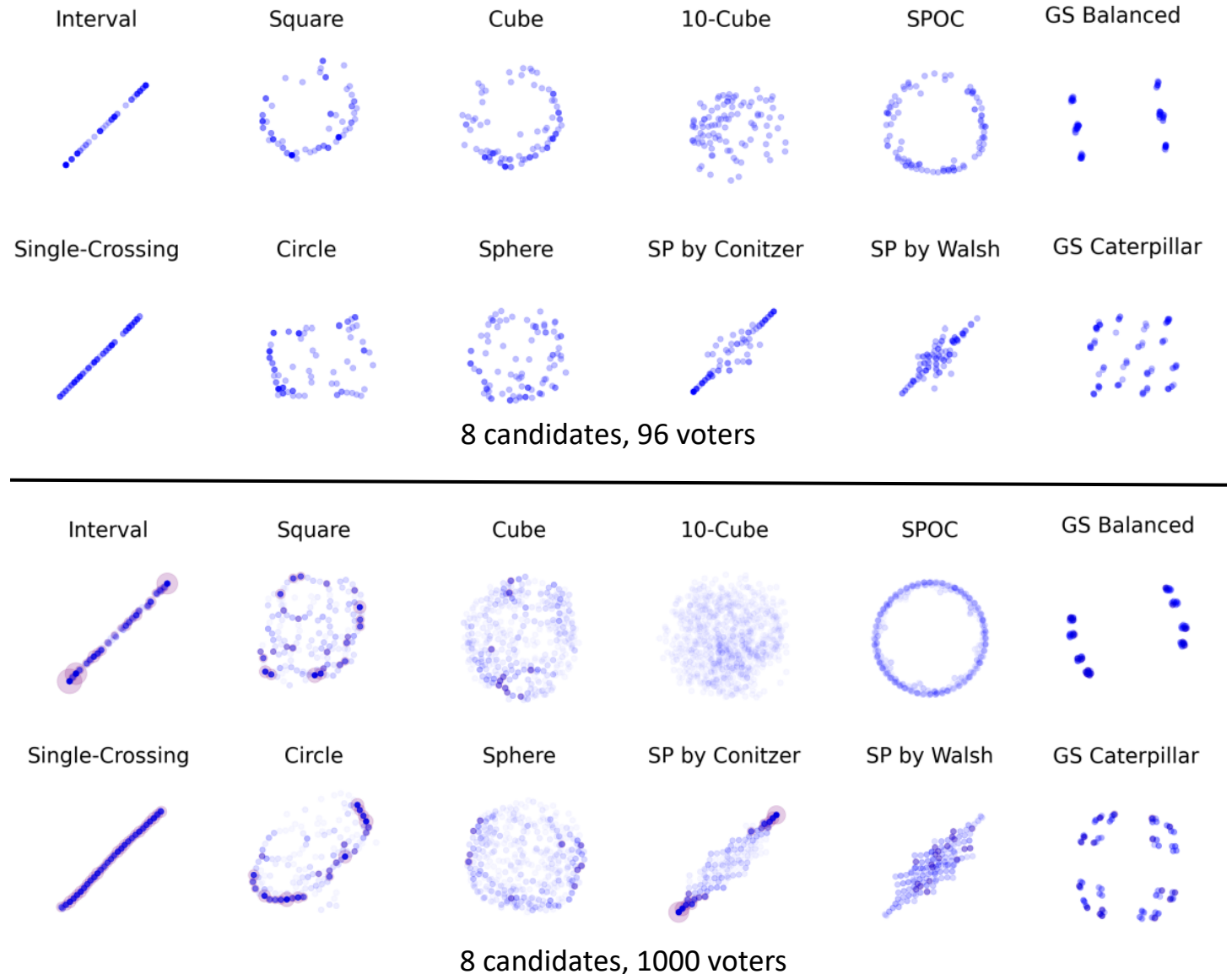
**Euclidean Model:** Choose **points** for the voters and candidates from **Euclidean space**  $\mathbb{R}^t$ . Voter  $v$  prefers candidate  $x$  to  $y$  if  $x$ 's point is **closer** to  $v$  than  $y$ 's.

**Single-Peaked:** There is societal axis (order of the of the candidates). Every single-peaked vote for this axis satisfies the property that „for each  $t$ , the top  $t$  candidates form an interval on the axis”.

**SPOC:** Like SP, but the axis is cyclic

**Single-Crossing:** It is possible to order the voters so that as we go along this order, the relative ranking of two candidates changes at most once

**Group-Separable:** Trees, trees everywhere!



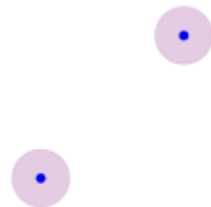
Impartial Culture



Identity



Antagonism



Stratific. 0.5



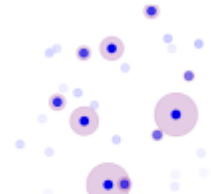
Stratific. 0.25



Urn 0.05



Urn 0.2



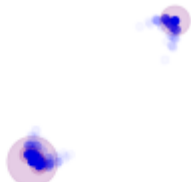
Urn 1



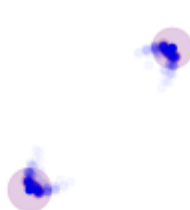
N-Mal. 0.05



0.25-N-Mal. 0.05



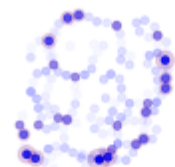
0.5-N-Mal. 0.05



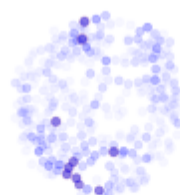
Interval



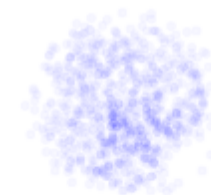
Square



Cube



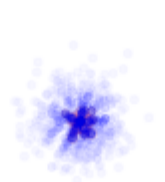
10-Cube



SPOC



N-Mal. 0.2



0.25-N-Mal. 0.2



0.5-N-Mal. 0.2



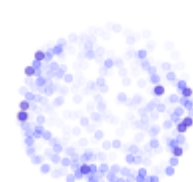
Single-Crossing



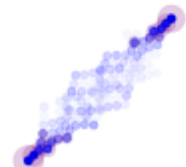
Circle



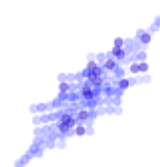
Sphere



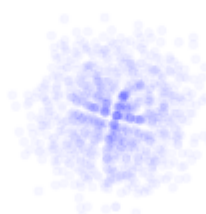
SP by Conitzer



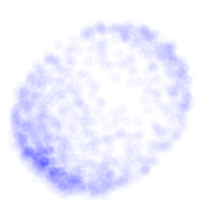
SP by Walsh



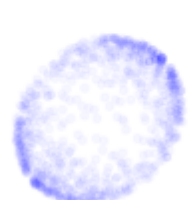
N-Mal. 0.5



0.25-N-Mal. 0.5



0.5-N-Mal. 0.5



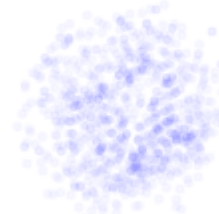
GS Balanced



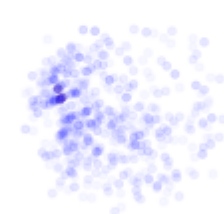
GS Caterpillar



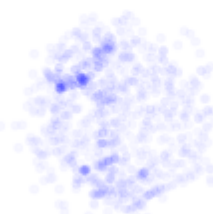
Sushi



Grenoble



Irish





Guide to Numerical Experiments on Elections in Computational Social Choice, Boehmer, Faliszewski, Janeczko, Kaczmarczyk, Lisowski, Pierczyński, Rey, Stolicki, Szufa, Wąs, arXiv 2024



10 minutes

# What's Used?

# Collecting the Data

## Papers

- AAI, AAMAS, IJCAI
- 2010—2023
- Downloaded all the papers using the XML file from DBLP (September 2023)

## Screening Process

- Automated script looking for election- and experiment-related keywords
  - `election, vote, ballot`
  - `experiment, empirical, simulation`
- Manual check of the shortlist
- E.g., IJCAI-23:
  - 846 papers
  - Script shortlisted 41
  - Manual check retained 7

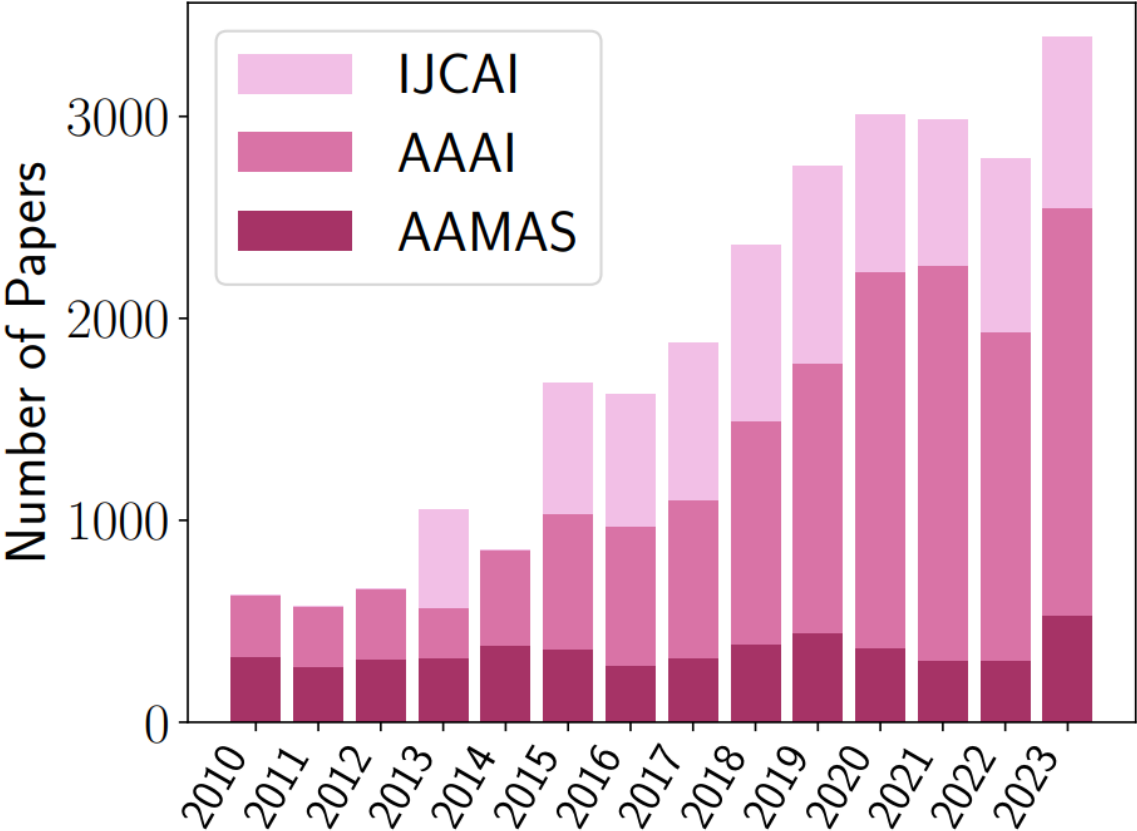


# Basic Statistics

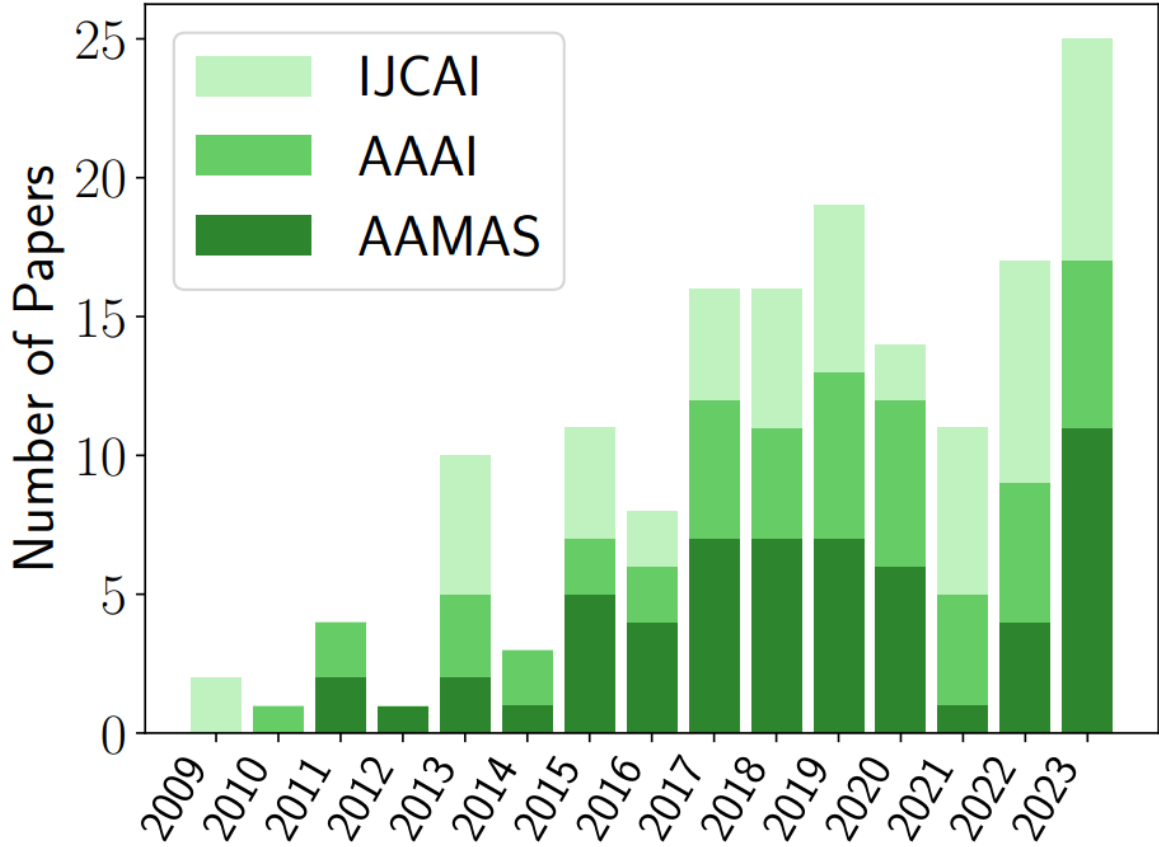
- Papers: 163
  - 130 ordinal
  - 35 approval
  - **Puzzle?**
- Experiments: 257
  - 211 ordinal
  - 46 approval
- Authors: 273 (+/-)

P. Faliszewski --> 26 paper(s) (18 ordinal, 8 approval)  
P. Skowron --> 14 paper(s) (8 ordinal, 6 approval)  
N. Talmon --> 14 paper(s) (11 ordinal, 3 approval)  
M. Lackner --> 12 paper(s) (3 ordinal, 9 approval)  
S. Szufa --> 11 paper(s) (8 ordinal, 3 approval)  
A. Procaccia --> 8 paper(s) (ordinal)  
A. Slinko --> 8 paper(s) (7 ordinal, 1 approval)  
N. Boehmer --> 7 paper(s) (ordinal)  
N. Mattei --> 7 paper(s) (5 ordinal, 2 approval)  
N. Shah --> 7 paper(s) (6 ordinal, 1 approval)  
L. Xia --> 7 paper(s) (ordinal)  
C. Boutilier --> 6 paper(s) (ordinal)  
U. Endriss --> 6 paper(s) (4 ordinal, 2 approval)  
J. Lang --> 6 paper(s) (3 ordinal, 3 approval)  
O. Lev --> 6 paper(s) (ordinal)  
D. Peters --> 6 paper(s) (4 ordinal, 2 approval)  
T. Walsh --> 6 paper(s) (ordinal)  
R. Brederick --> 5 paper(s) (4 ordinal, 1 approval)  
M. Brill --> 5 paper(s) (2 ordinal, 3 approval)  
E. Elkind --> 5 paper(s) (3 ordinal, 2 approval)  
R. Meir --> 5 paper(s) (3 ordinal, 3 approval)  
R. Niedermeier --> 5 paper(s) (4 ordinal, 1 approval)  
J. Rosenschein --> 5 paper(s) (ordinal)  
F. Rossi --> 5 paper(s) (ordinal)  
H. Aziz --> 4 paper(s) (ordinal)  
F. Brandt --> 4 paper(s) (ordinal)  
I. Caragiannis --> 4 paper(s) (ordinal)  
S. Kraus --> 4 paper(s) (ordinal)  
Y. Lewenberg --> 4 paper(s) (ordinal)  
S. Nath --> 4 paper(s) (ordinal)  
K. Sornat --> 4 paper(s) (2 ordinal, 2 approval)  
A. Wilczynski --> 4 paper(s) (ordinal)

# Experiments on Elections in COMSOC

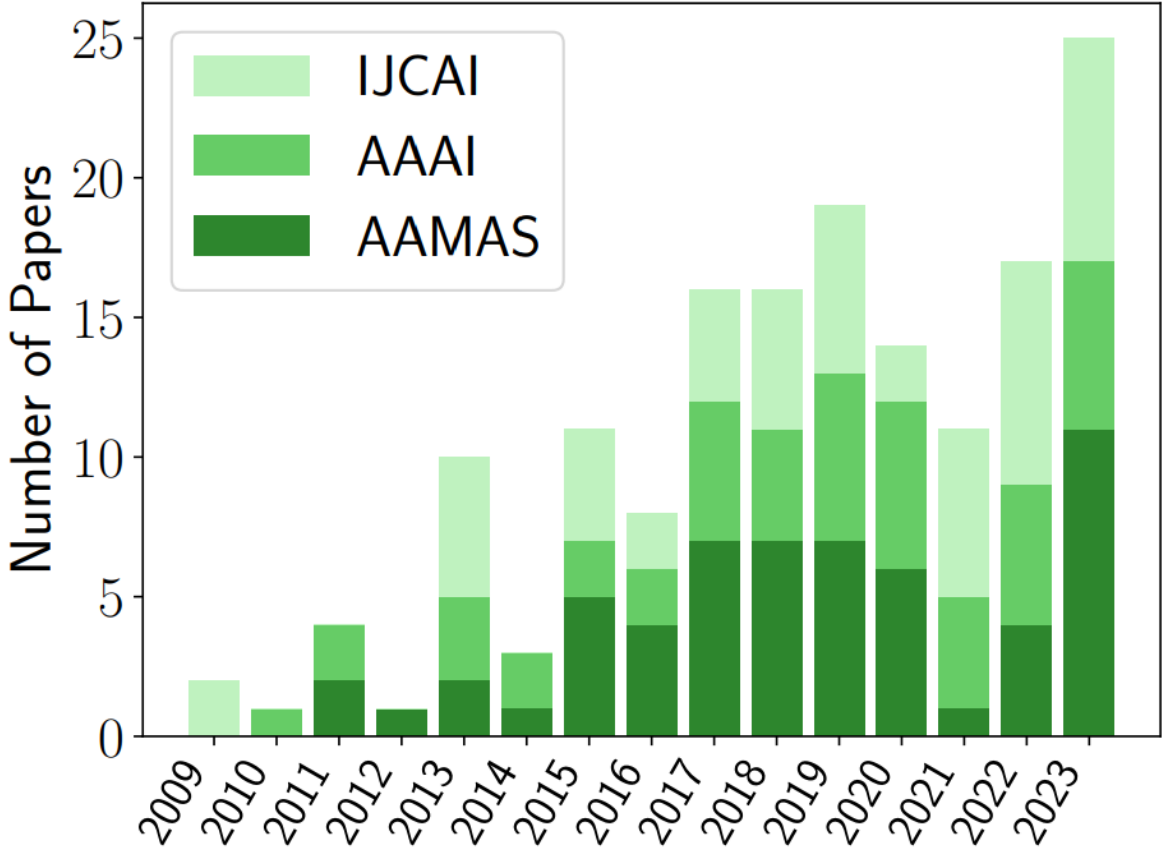


Papers in recent AI conferences

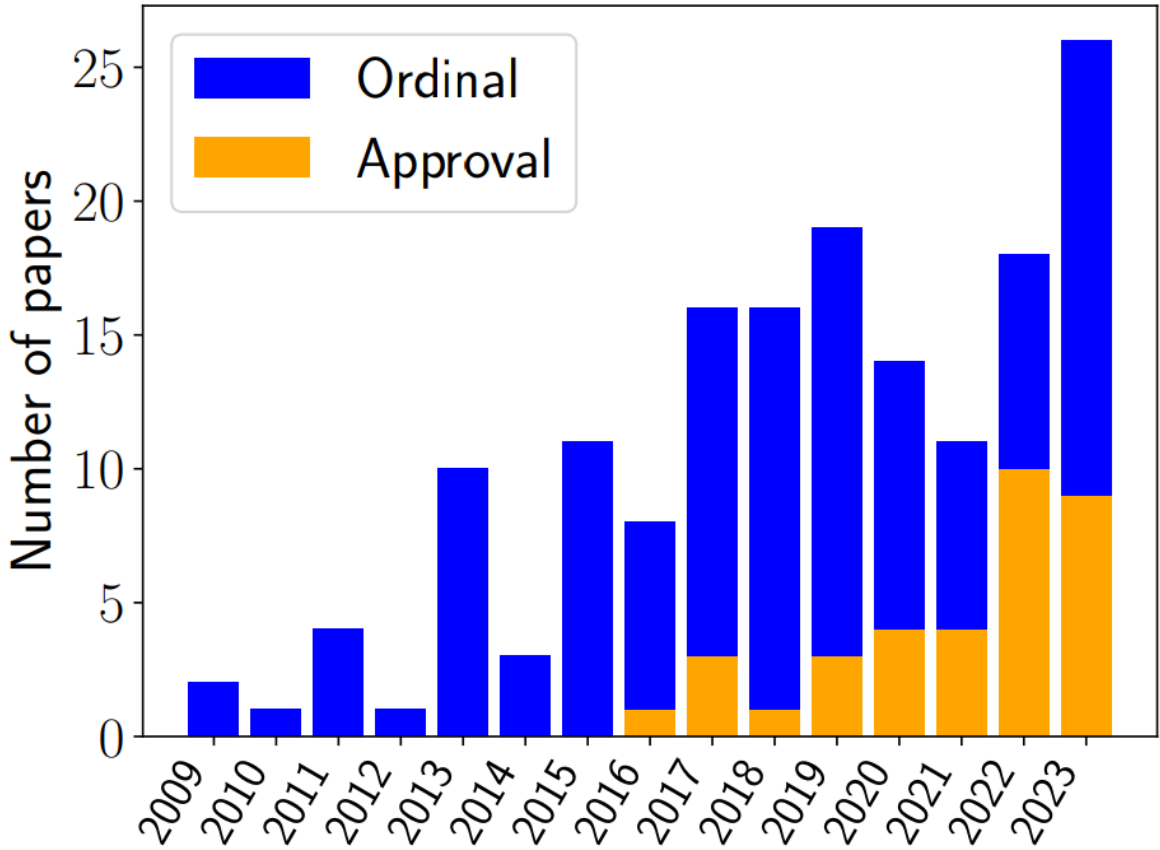


Papers in recent AI conferences that include experiments on elections\*

# Experiments on Elections in COMSOC



Papers in recent AI conferences that include experiments on elections\*



Ordinal preferences versus approval (as covered in the papers)

# What Elections to Study?

Structure of the preference orders?

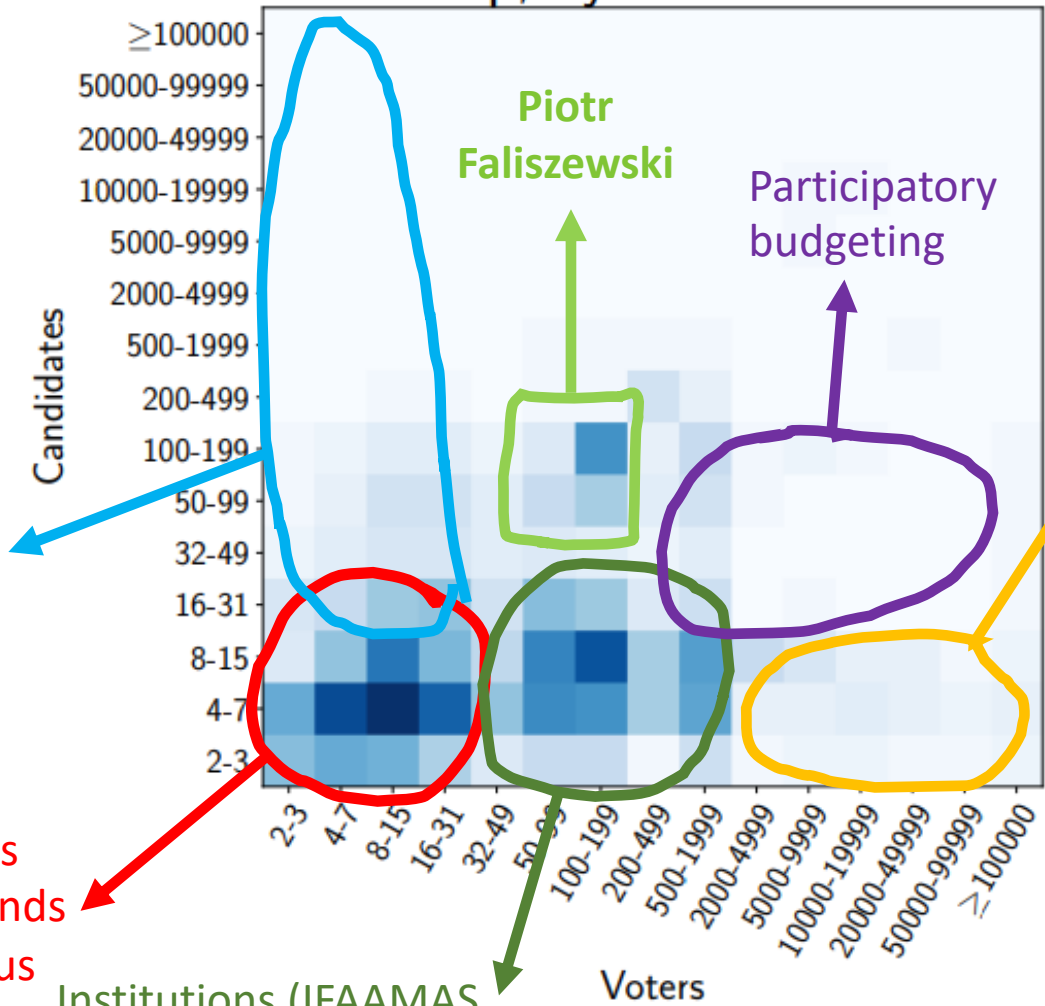
Reasonable numbers of candidates and voters?

Ground-truth search (sporting events, meta-search engines, recommendation systems, etc.)

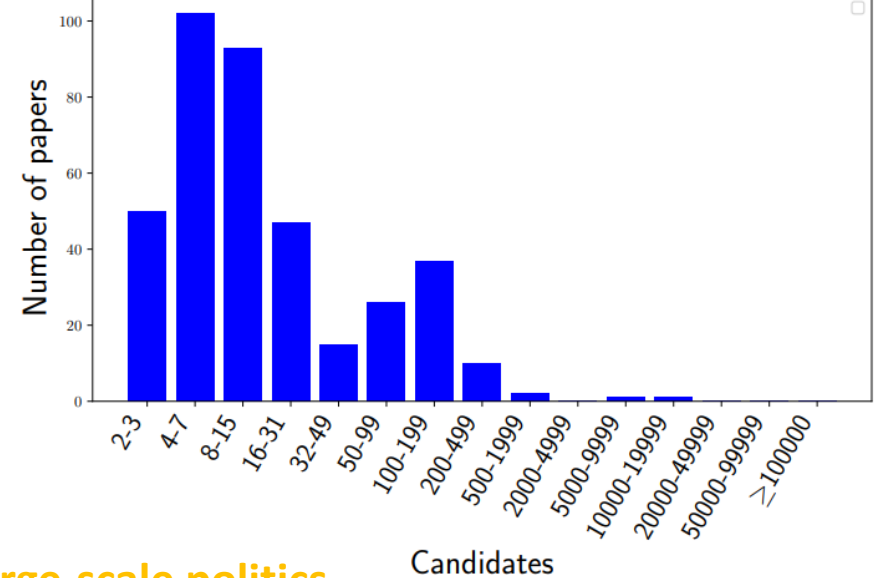
Small committees (e.g., hiring), friends voting on frivolous stuff, „usual life”

Institutions (IFAAMAS board elections, choosing electors at universities, etc.)

## Heatmap, Synthetic Elections

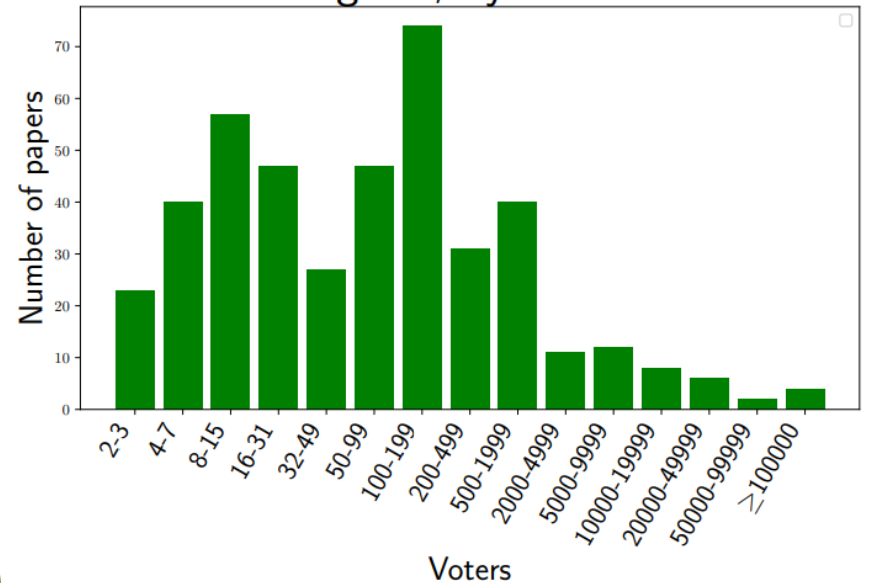


## Candidate Histogram, Synthetic Elections



Large-scale politics

## Voter Histogram, Synthetic Elections



# What Elections to Study?

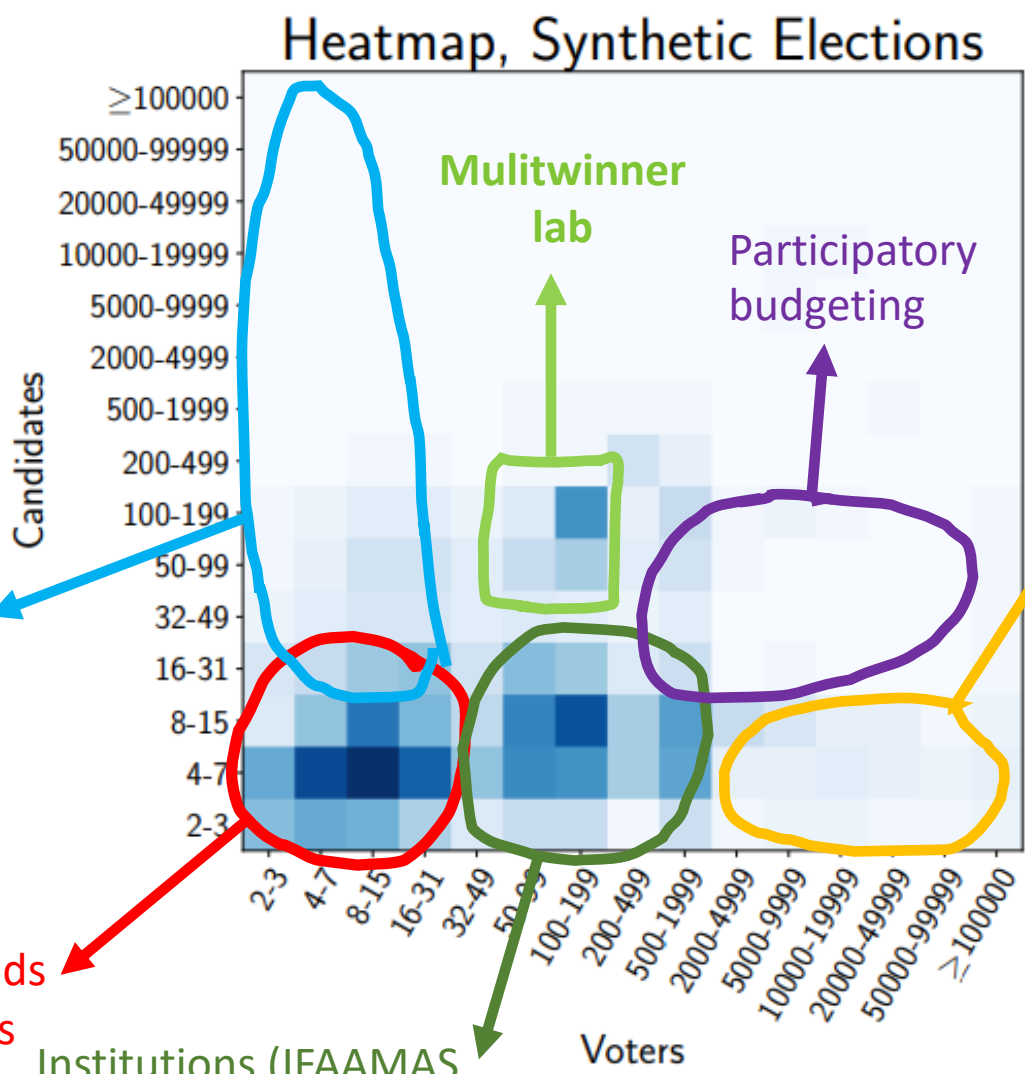
Structure of the preference orders?

Reasonable numbers of candidates and voters?

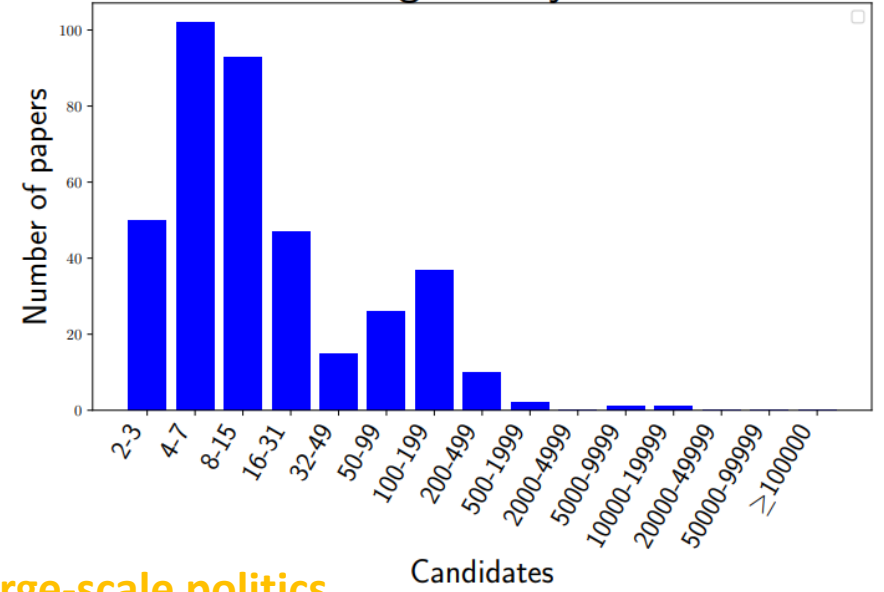
Ground-truth search (sporting events, meta-search engines, recommendation systems, etc.)

Small committees (e.g., hiring), friends voting on frivolous stuff, „usual life”

Institutions (IFAAMAS board elections, choosing electors at universities, etc.)

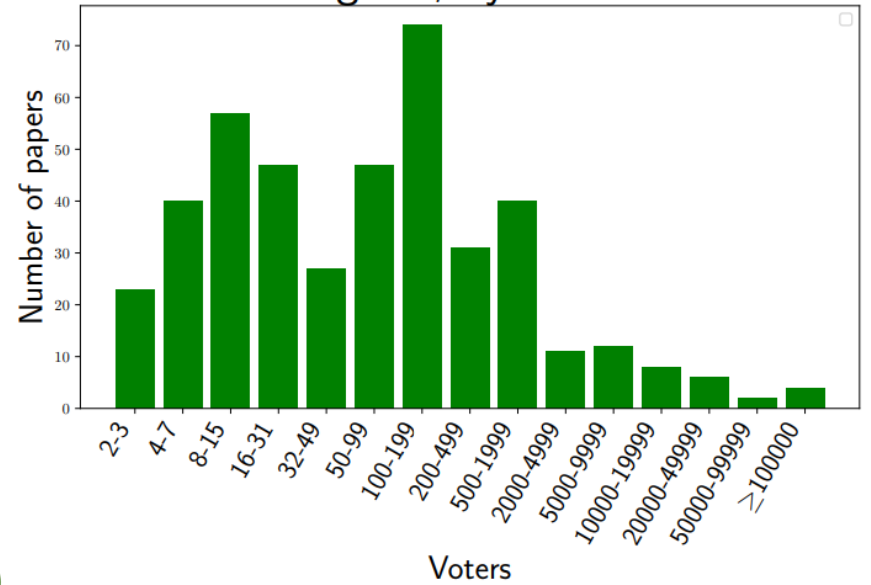


Candidate Histogram, Synthetic Elections

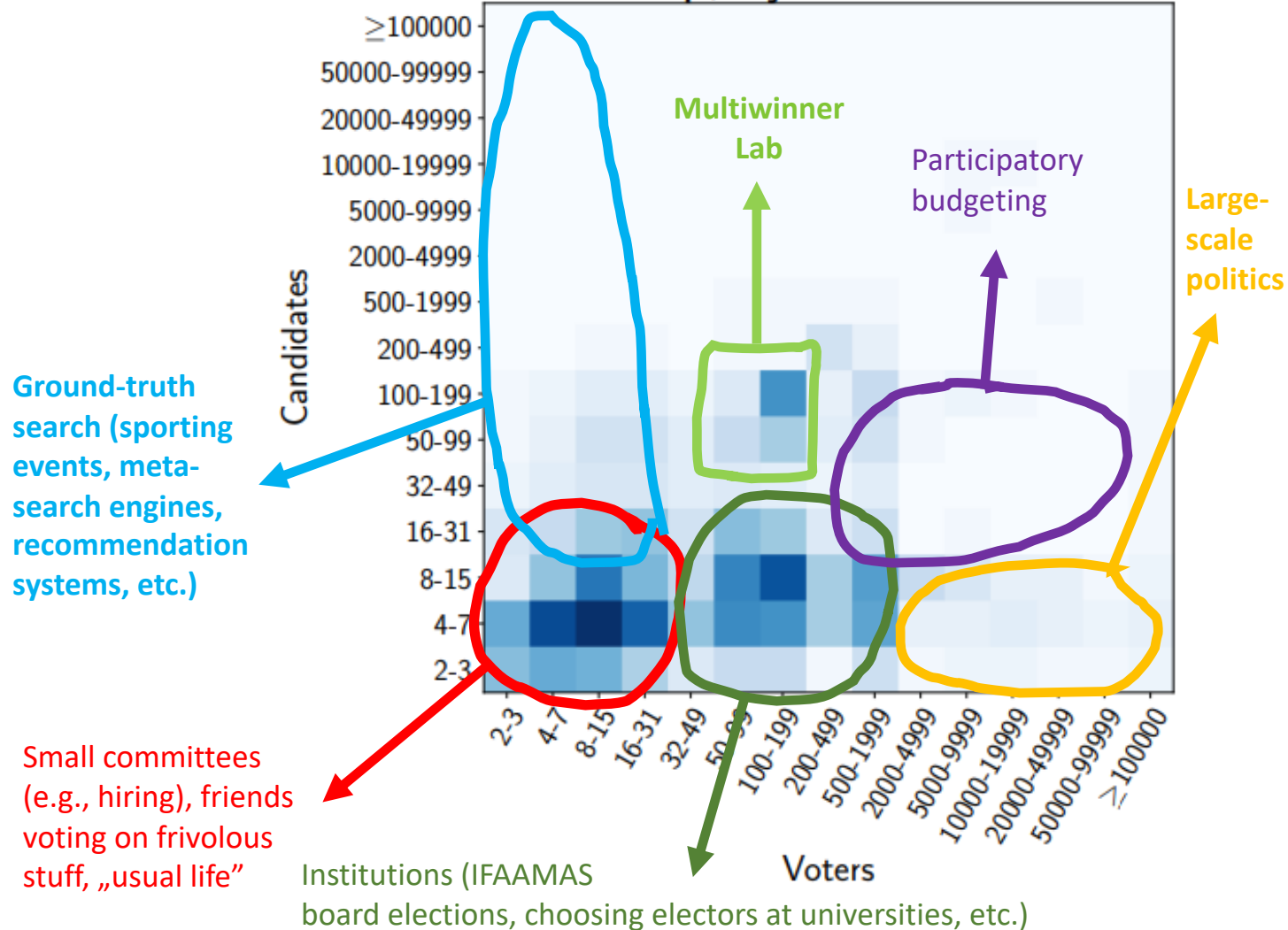


Large-scale politics

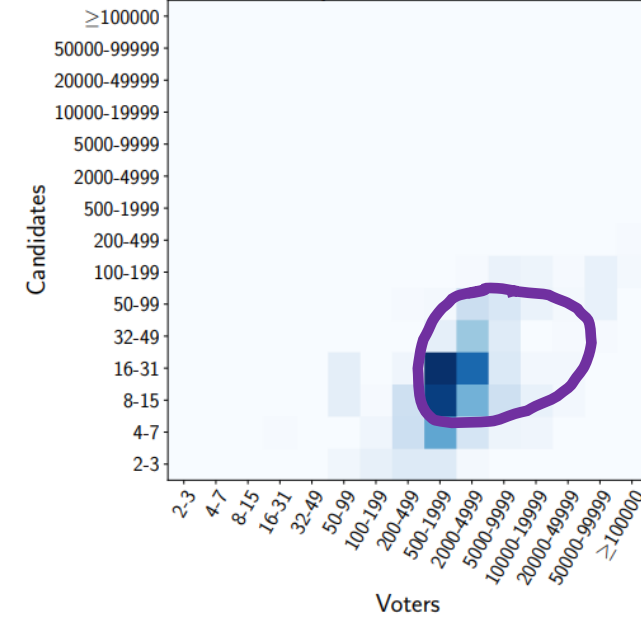
Voter Histogram, Synthetic Elections



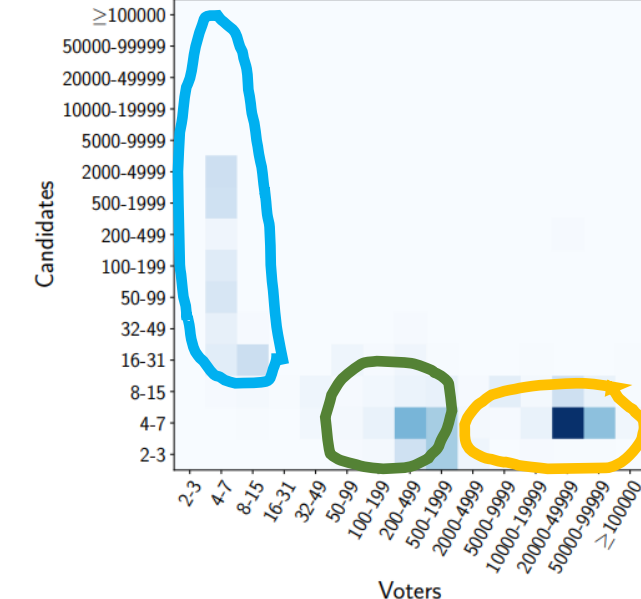
### Heatmap, Synthetic Elections



### Heatmap, Pabulib Elections



### Heatmap, Preflib Elections



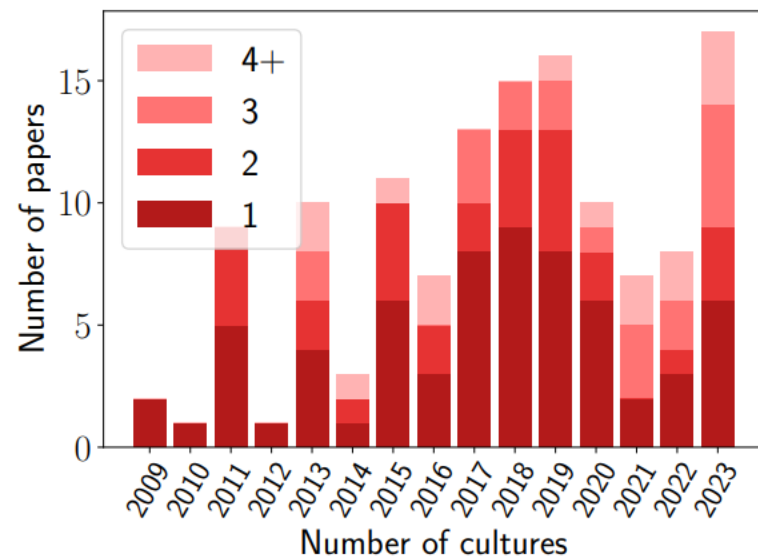
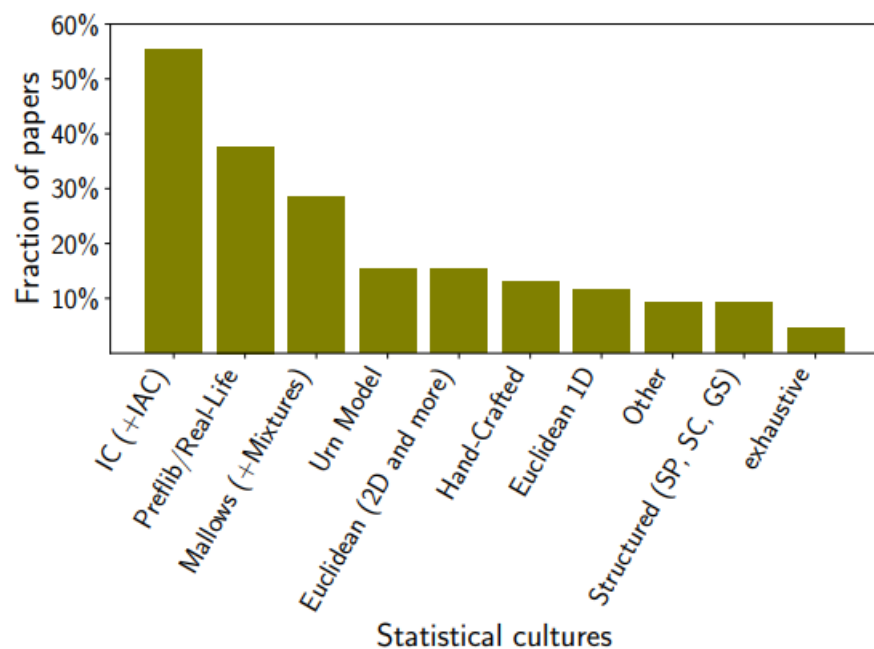
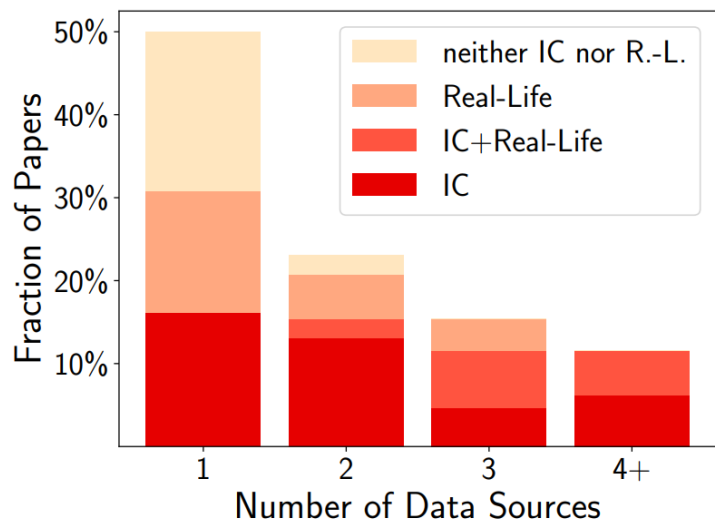
# What Elections to Study?



Reasonable  
numbers of  
candidates and  
voters?

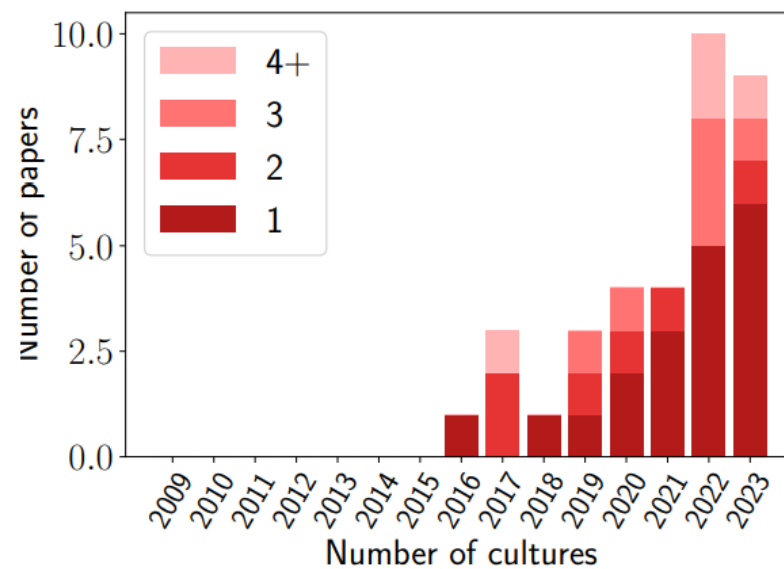
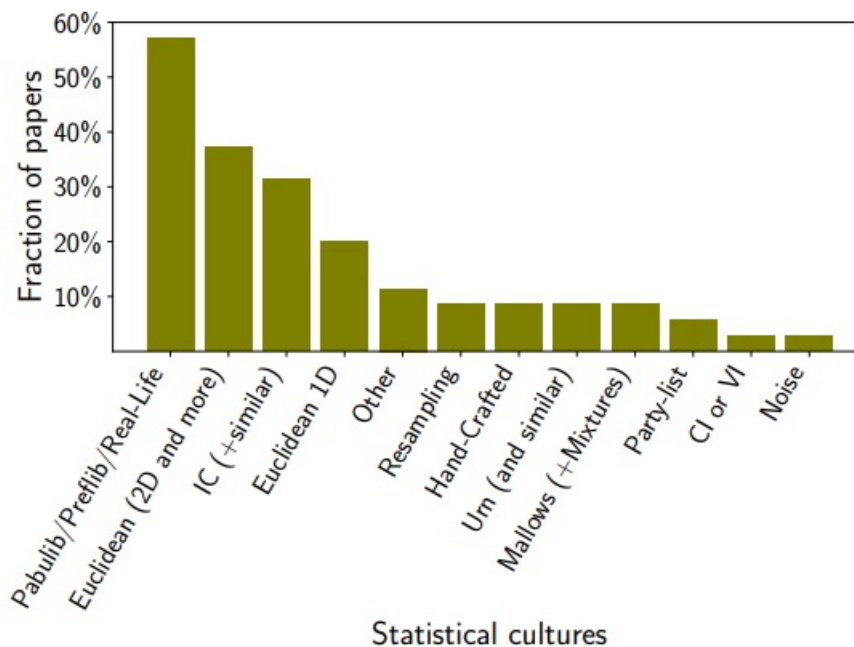


Structure of the preference  
orders?



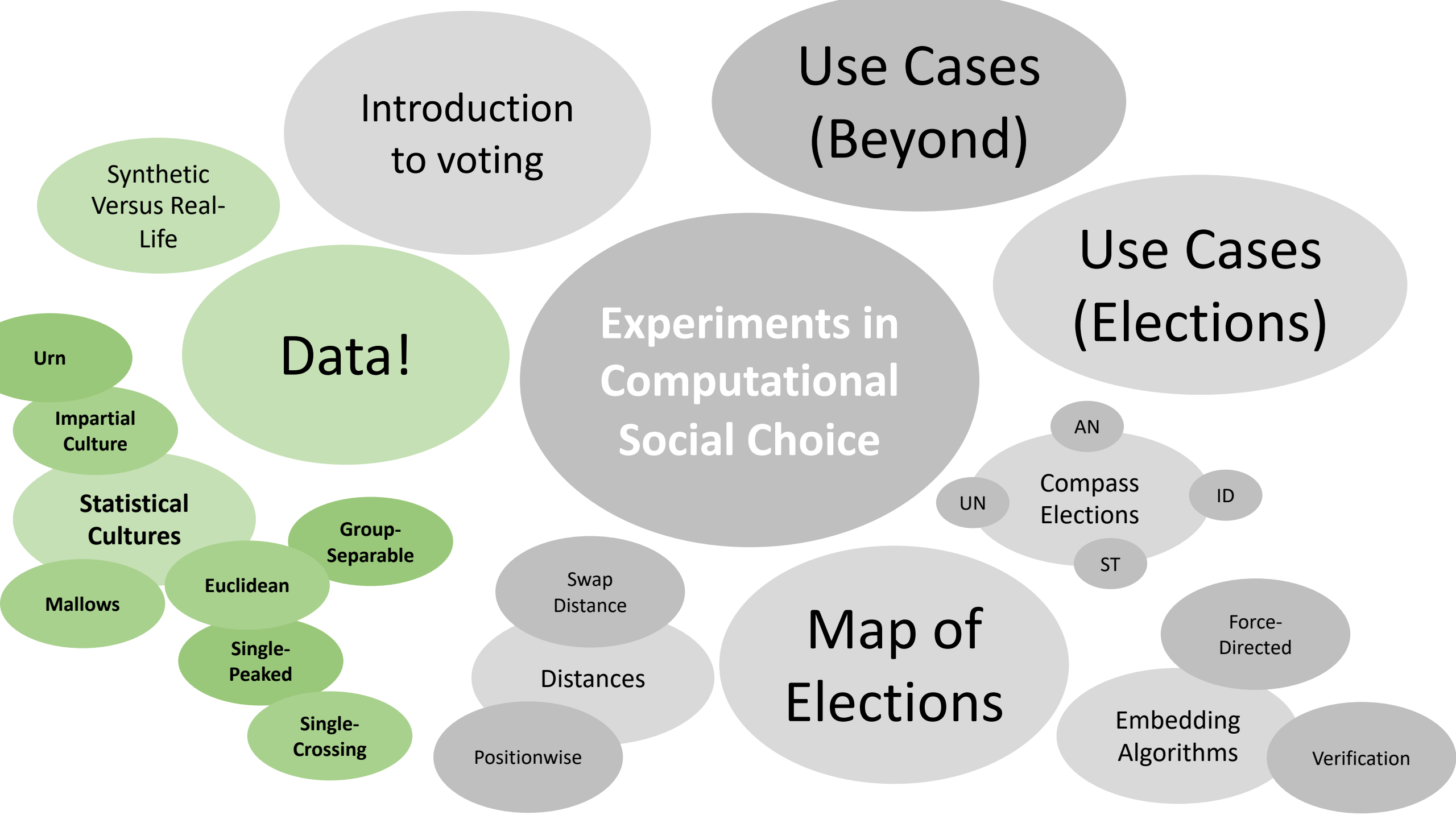
Ordinal

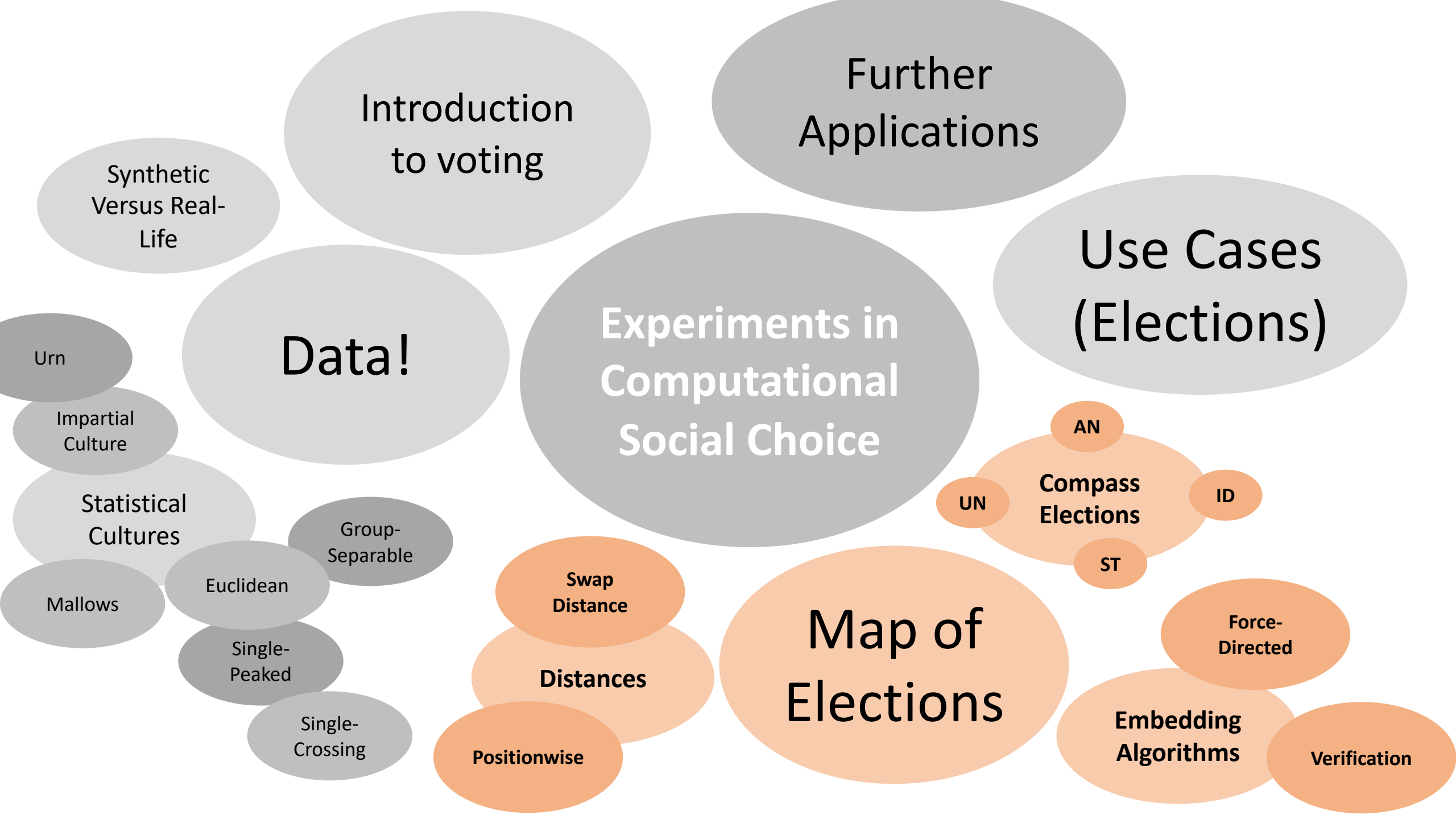
Approval











Further Applications

Use Cases (Elections)

Data!

Experiments in Computational Social Choice

Map of Elections

Compass Elections

Distances

Embedding Algorithms

Introduction to voting

Synthetic Versus Real-Life

Urn

Impartial Culture

Statistical Cultures

Mallows

Euclidean

Single-Peaked

Single-Crossing

Group-Separable

Swap Distance

Positionwise

Force-Directed

Verification

AN

UN

ID

ST



30 minutes

# Map of Elections

# How different?

- V<sub>1</sub>: 🐼 > 🐳 > 🐱
- V<sub>2</sub>: 🐼 > 🐱 > 🐳
- V<sub>3</sub>: 🐱 > 🐳 > 🐼
- V<sub>4</sub>: 🐱 > 🐼 > 🐳
- V<sub>5</sub>: 🐳 > 🐼 > 🐱
- V<sub>6</sub>: 🐳 > 🐱 > 🐼

all possible preference orders

**uniformity**

●  
UN

Count the number of swaps that make the elections isomorphic (i.e., identical up to renaming the candidates and reordering the voters)

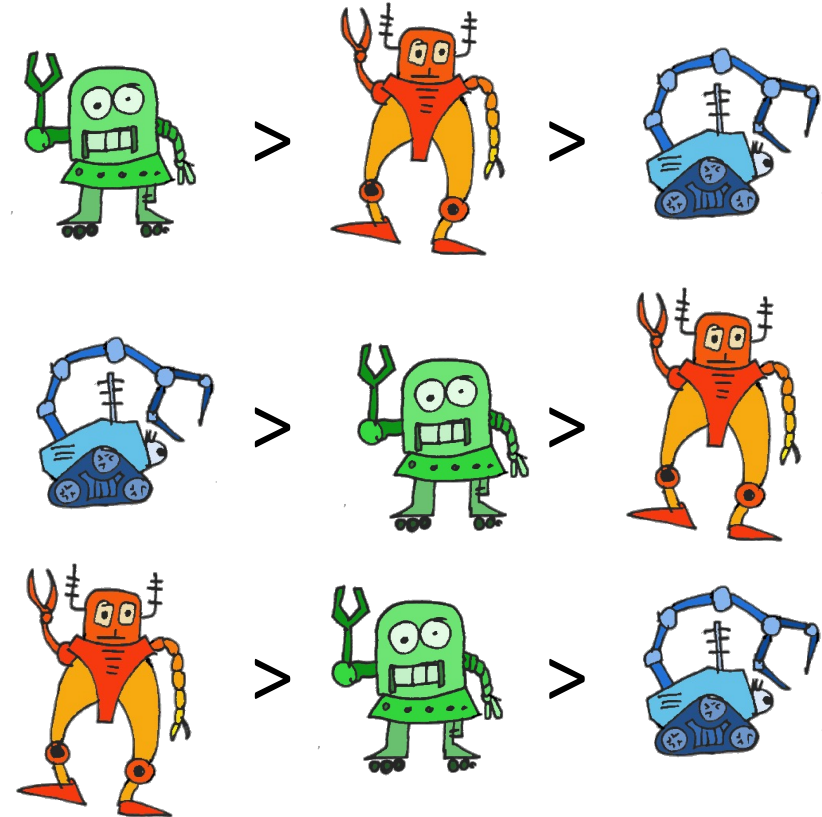
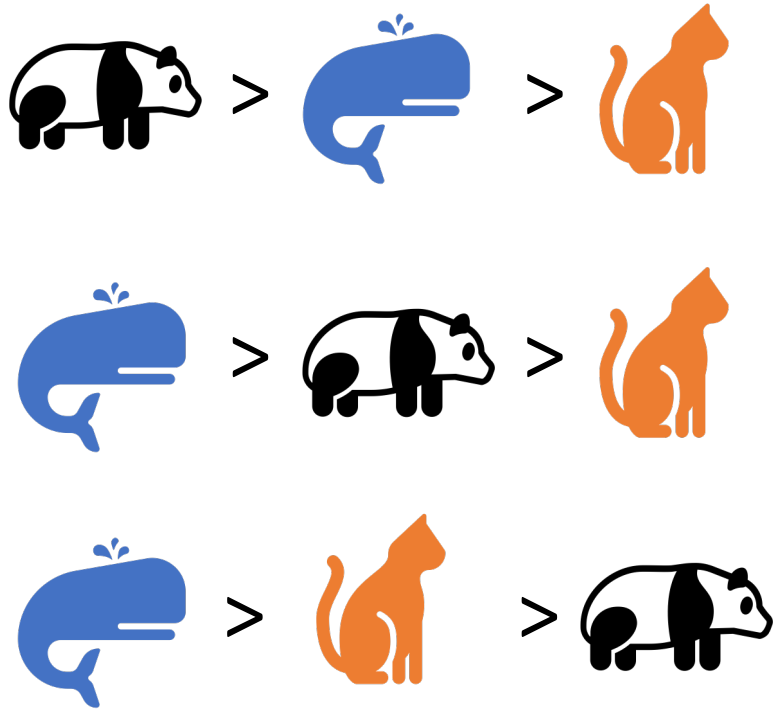
●  
ID

- V<sub>1</sub>: 👧 > 👨 > 👨
- V<sub>2</sub>: 👧 > 👨 > 👨
- V<sub>3</sub>: 👧 > 👨 > 👨
- V<sub>4</sub>: 👧 > 👨 > 👨
- V<sub>5</sub>: 👧 > 👨 > 👨
- V<sub>6</sub>: 👧 > 👨 > 👨

Identical preference orders

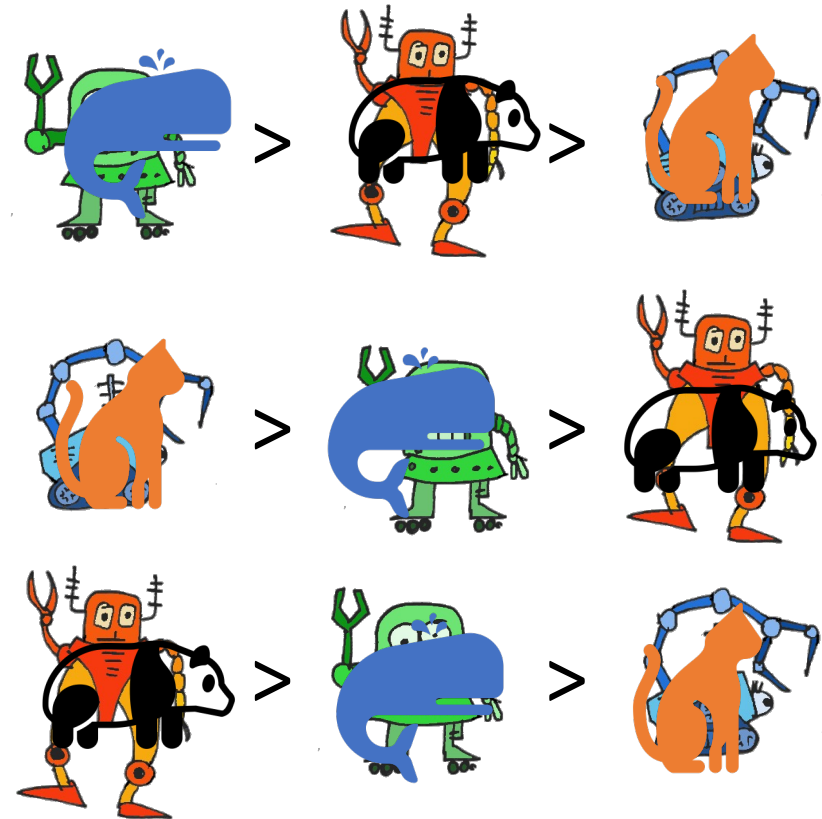
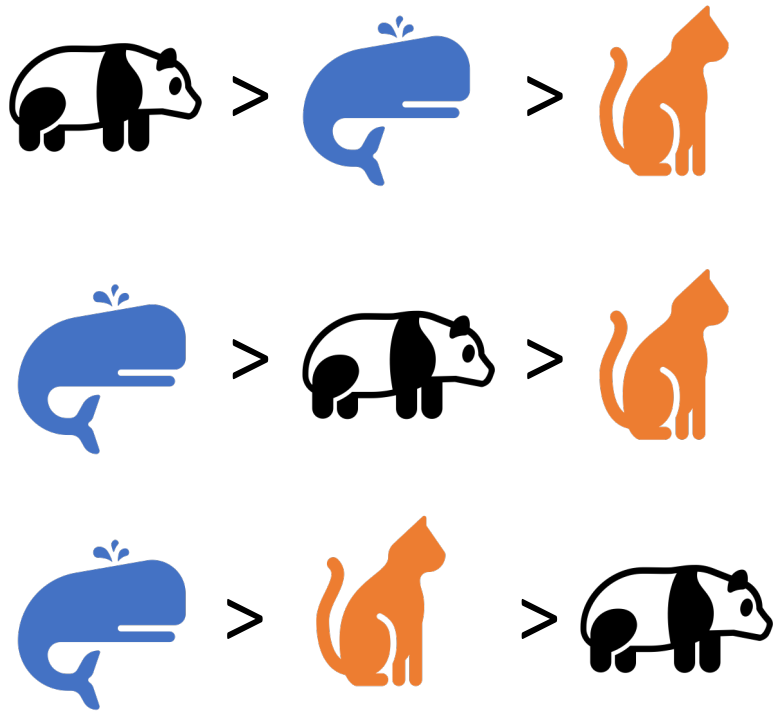
**identity**

# Isomorphic Swap Distance



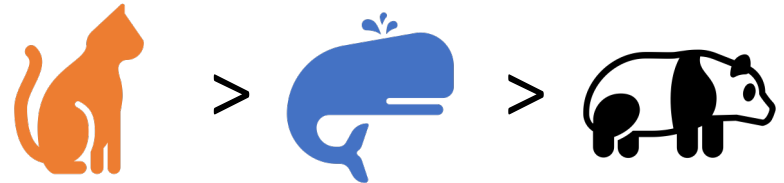
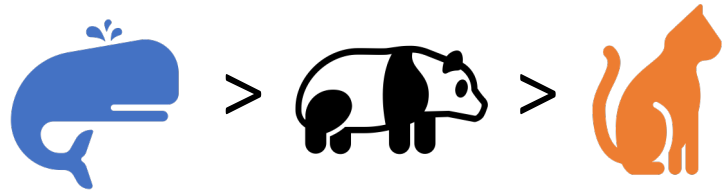
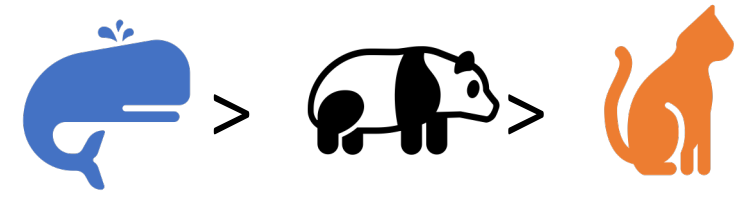
1. Match the candidates
2. Match the voters
3. Count the swaps

# Isomorphic Swap Distance



1. Match the candidates
2. Match the voters
3. Count the swaps

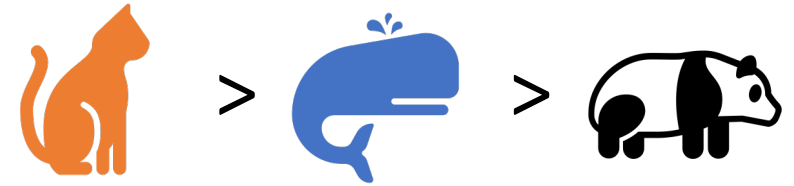
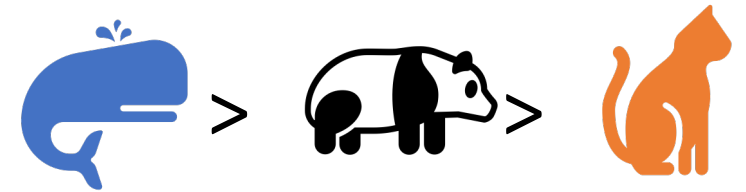
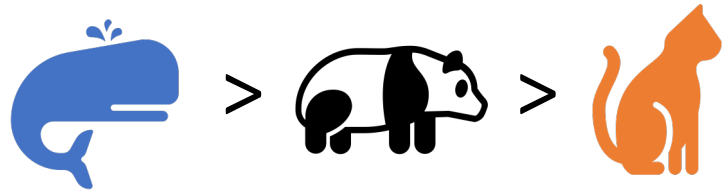
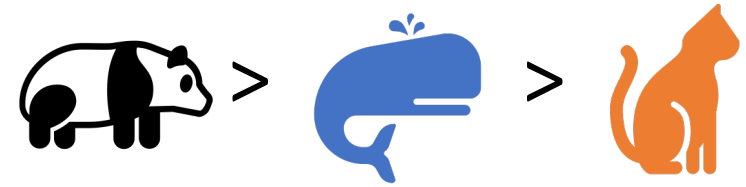
# Isomorphic Swap Distance



1. Match the candidates
2. Match the voters
3. Count the swaps

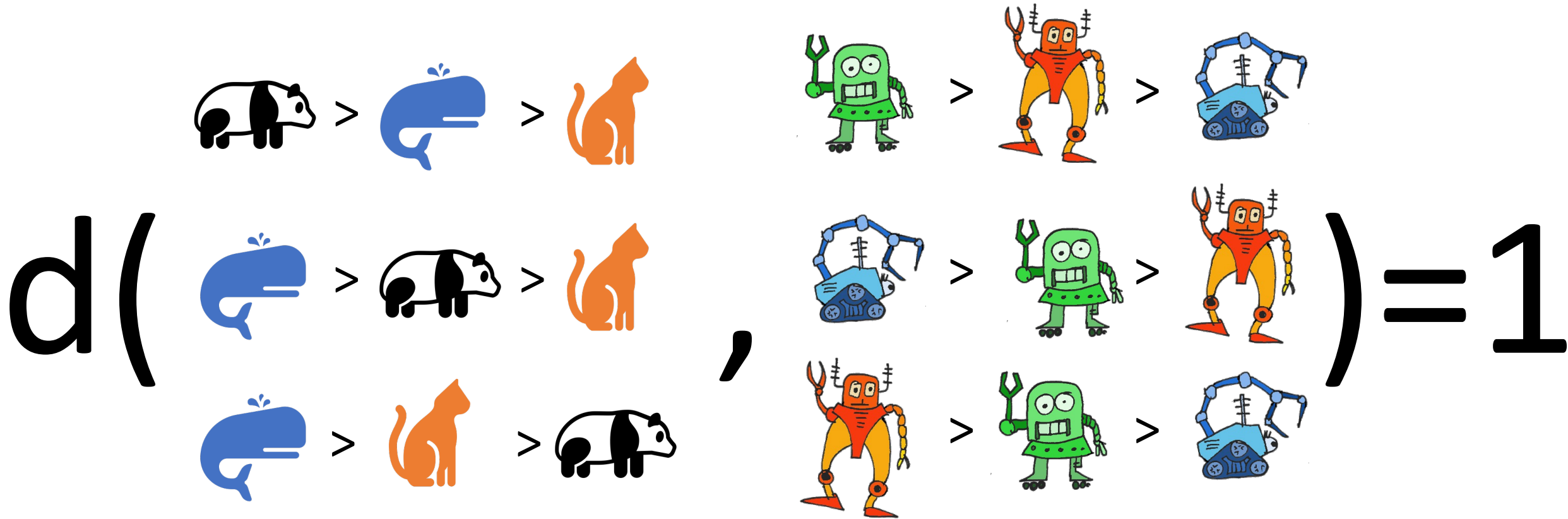


# Isomorphic Swap Distance



1. Match the candidates
2. Match the voters
3. Count the swaps

# Isomorphic Swap Distance



1. Match the candidates
2. Match the voters
3. Count the swaps

Thm. In an election with  $m$  candidates and  $n = t \cdot m!$  votes, every two elections are at distance at most  $\frac{1}{4} n(m^2 - m)$ .

- V<sub>1</sub>: 🐼 > 🐳 > 🐱
- V<sub>2</sub>: 🐼 > 🐱 > 🐳
- V<sub>3</sub>: 🐱 > 🐳 > 🐼
- V<sub>4</sub>: 🐱 > 🐼 > 🐳
- V<sub>5</sub>: 🐳 > 🐼 > 🐱
- V<sub>6</sub>: 🐳 > 🐱 > 🐼

all possible preference orders

**uniformity**



UN

$$\frac{1}{4} n(m^2 - m)$$



ID

Count the number of swaps that make the elections isomorphic (i.e., identical up to renaming the candidates and reordering the voters)

- V<sub>1</sub>: 👩 > 👨 > 👨
- V<sub>2</sub>: 👩 > 👨 > 👨
- V<sub>3</sub>: 👩 > 👨 > 👨
- V<sub>4</sub>: 👩 > 👨 > 👨
- V<sub>5</sub>: 👩 > 👨 > 👨
- V<sub>6</sub>: 👩 > 👨 > 👨

Identical preference orders

**identity**

- V<sub>1</sub>: 🐼 > 🐋 > 🐱
- V<sub>2</sub>: 🐼 > 🐱 > 🐋
- V<sub>3</sub>: 🐱 > 🐋 > 🐼
- V<sub>4</sub>: 🐱 > 🐼 > 🐋
- V<sub>5</sub>: 🐋 > 🐼 > 🐱
- V<sub>6</sub>: 🐋 > 🐱 > 🐼

all possible preference orders

**uniformity**



UN

1



ID

- V<sub>1</sub>: 👩 > 👨 > 👴
- V<sub>2</sub>: 👩 > 👨 > 👴
- V<sub>3</sub>: 👩 > 👨 > 👴
- V<sub>4</sub>: 👩 > 👨 > 👴
- V<sub>5</sub>: 👩 > 👨 > 👴
- V<sub>6</sub>: 👩 > 👨 > 👴

Identical preference orders

**identity**

two reverse orders  
antagonism



- V<sub>1</sub>: > > >
- V<sub>2</sub>: > > >
- V<sub>3</sub>: > > >
- V<sub>4</sub>: > > >
- V<sub>5</sub>: > > >
- V<sub>6</sub>: > > >

- V<sub>1</sub>: > > >
- V<sub>2</sub>: > > >
- V<sub>3</sub>: > > >
- V<sub>4</sub>: > > >
- V<sub>5</sub>: > > >
- V<sub>6</sub>: > > >



UN

1



ID

- V<sub>1</sub>: > > >
- V<sub>2</sub>: > > >
- V<sub>3</sub>: > > >
- V<sub>4</sub>: > > >
- V<sub>5</sub>: > > >
- V<sub>6</sub>: > > >



ST

two groups of candidates,  
each voter prefers members  
of one group to the other

**stratification**

- V<sub>1</sub>: > > >
- V<sub>2</sub>: > > >
- V<sub>3</sub>: > > >
- V<sub>4</sub>: > > >
- V<sub>5</sub>: > > >
- V<sub>6</sub>: > > >

two reverse orders  
antagonism

● AN

- V<sub>1</sub>: > > >
- V<sub>2</sub>: > > >
- V<sub>3</sub>: > > >
- V<sub>4</sub>: > > >
- V<sub>5</sub>: > > >
- V<sub>6</sub>: > > >

- V<sub>1</sub>: > > >
- V<sub>2</sub>: > > >
- V<sub>3</sub>: > > >
- V<sub>4</sub>: > > >
- V<sub>5</sub>: > > >
- V<sub>6</sub>: > > >

● UN

≈ 1

1

1

● ID

- V<sub>1</sub>: > > >
- V<sub>2</sub>: > > >
- V<sub>3</sub>: > > >
- V<sub>4</sub>: > > >
- V<sub>5</sub>: > > >
- V<sub>6</sub>: > > >

- V<sub>1</sub>: > > >
- V<sub>2</sub>: > > >
- V<sub>3</sub>: > > >
- V<sub>4</sub>: > > >
- V<sub>5</sub>: > > >
- V<sub>6</sub>: > > >

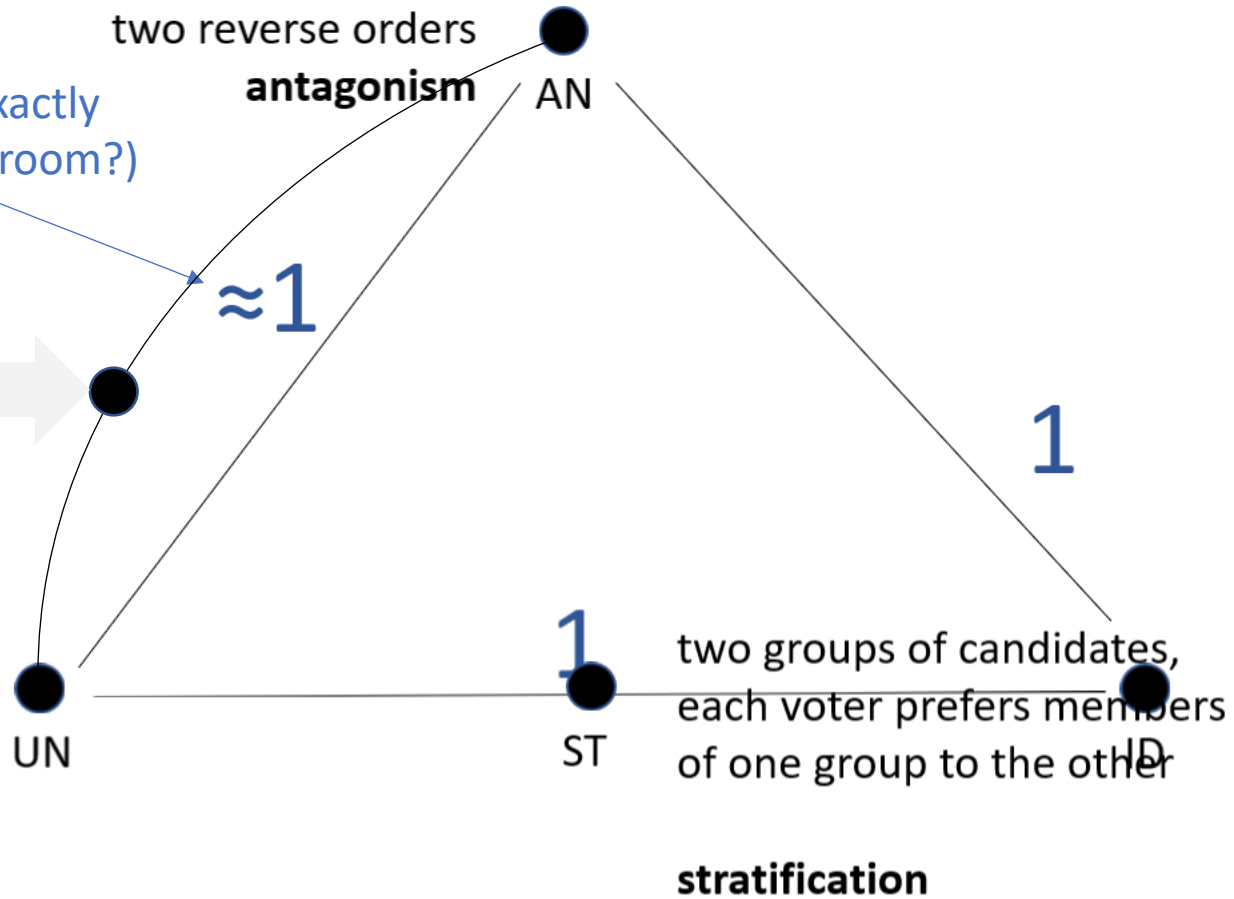
● ST

two groups of candidates,  
each voter prefers members  
of one group to the other

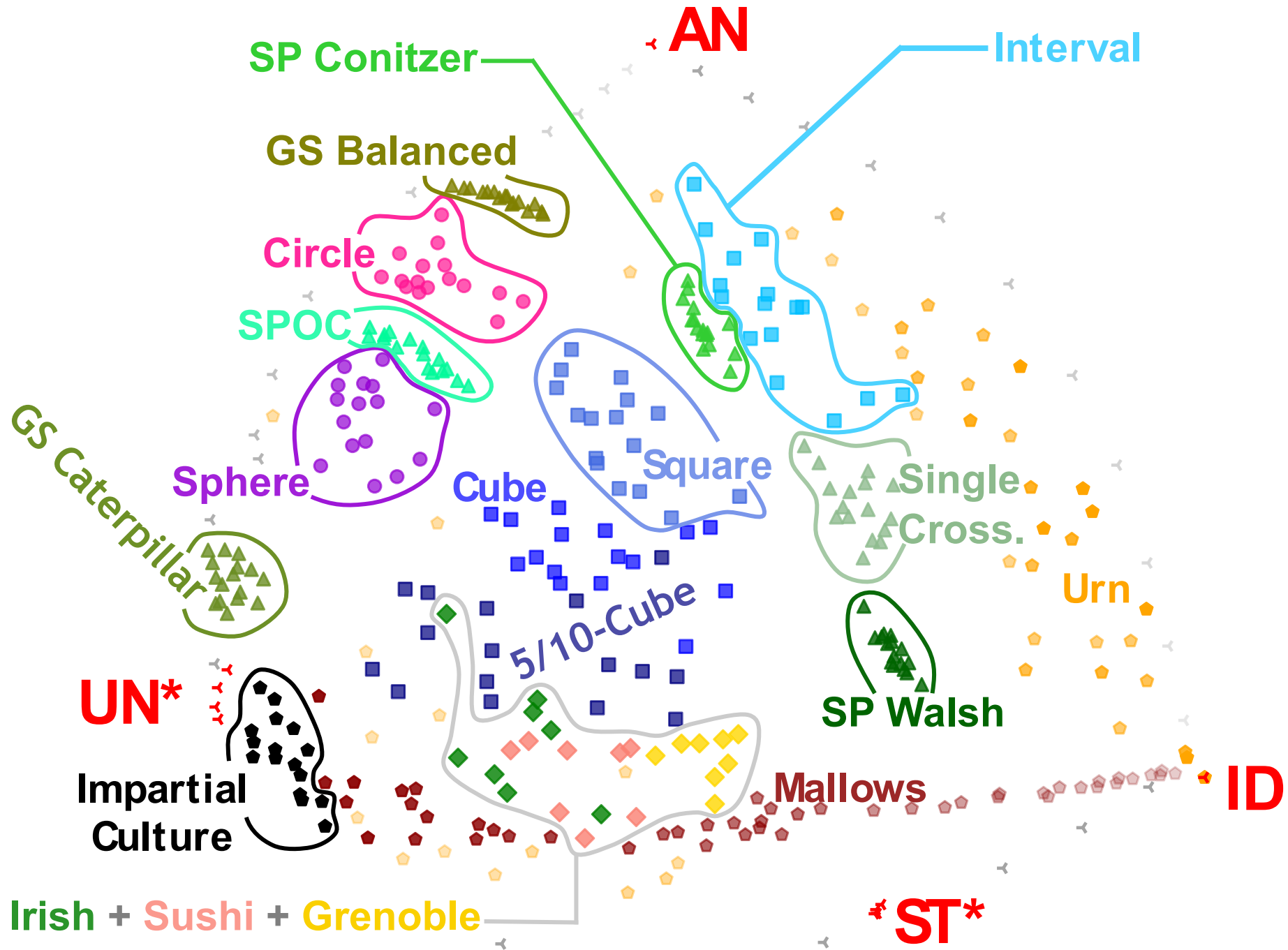
**stratification**

Somehow difficult to compute exactly  
(is there a mathematician in the room?)

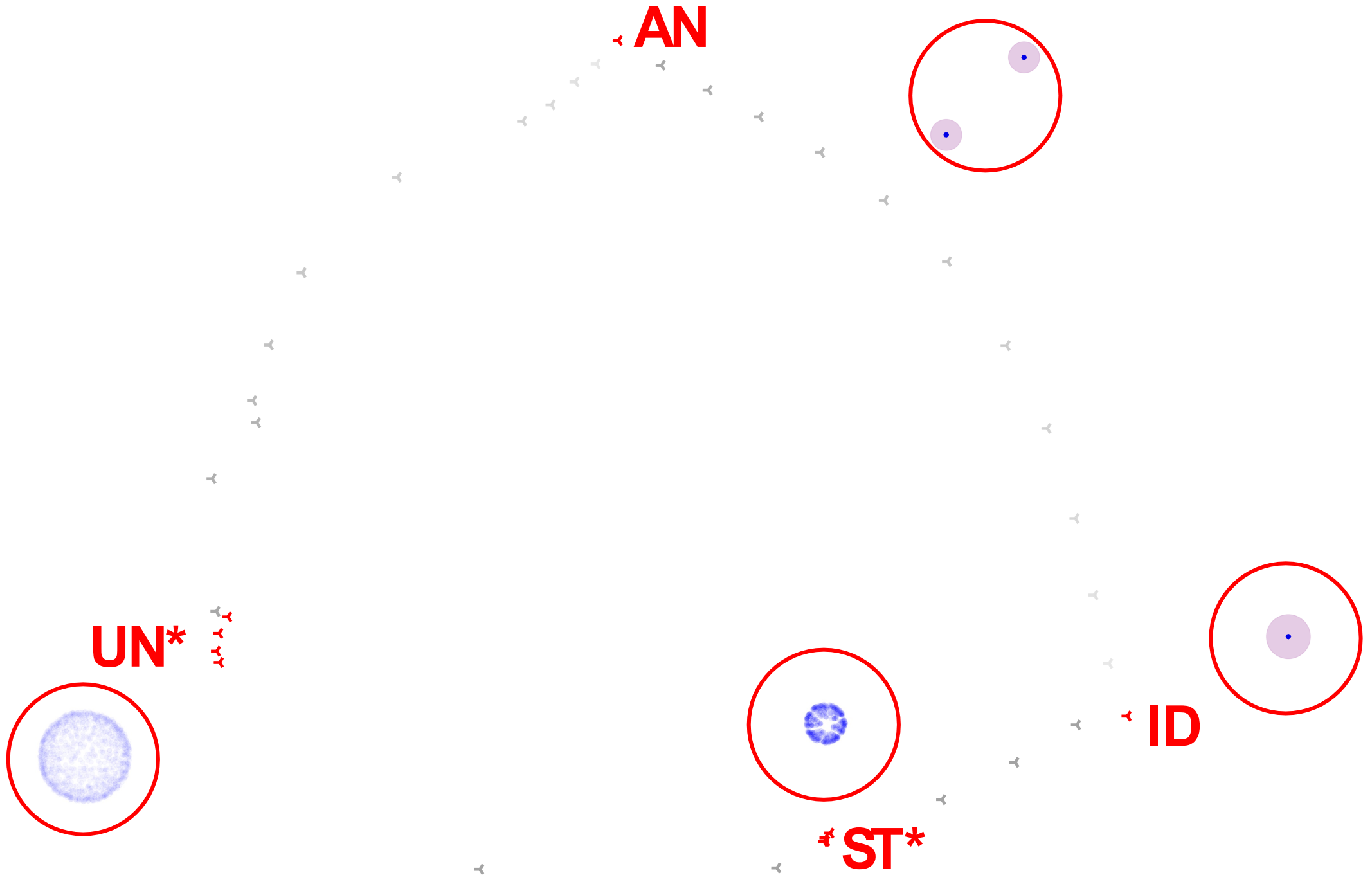
**Thm.** An election is at  
(normalized) distance 1  
from ID if and only if for  
all pairs of candidates  $a$   
and  $b$ , half of the voters  
prefer  $a$  to  $b$ , and half of  
the voters prefer  $b$  to  $a$ .

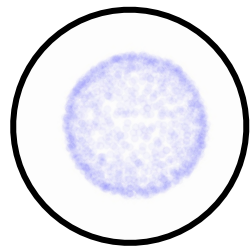


$v_6$ : > > >



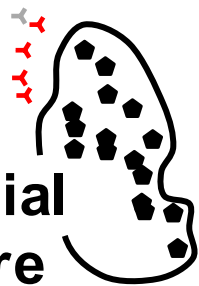






**UN\***

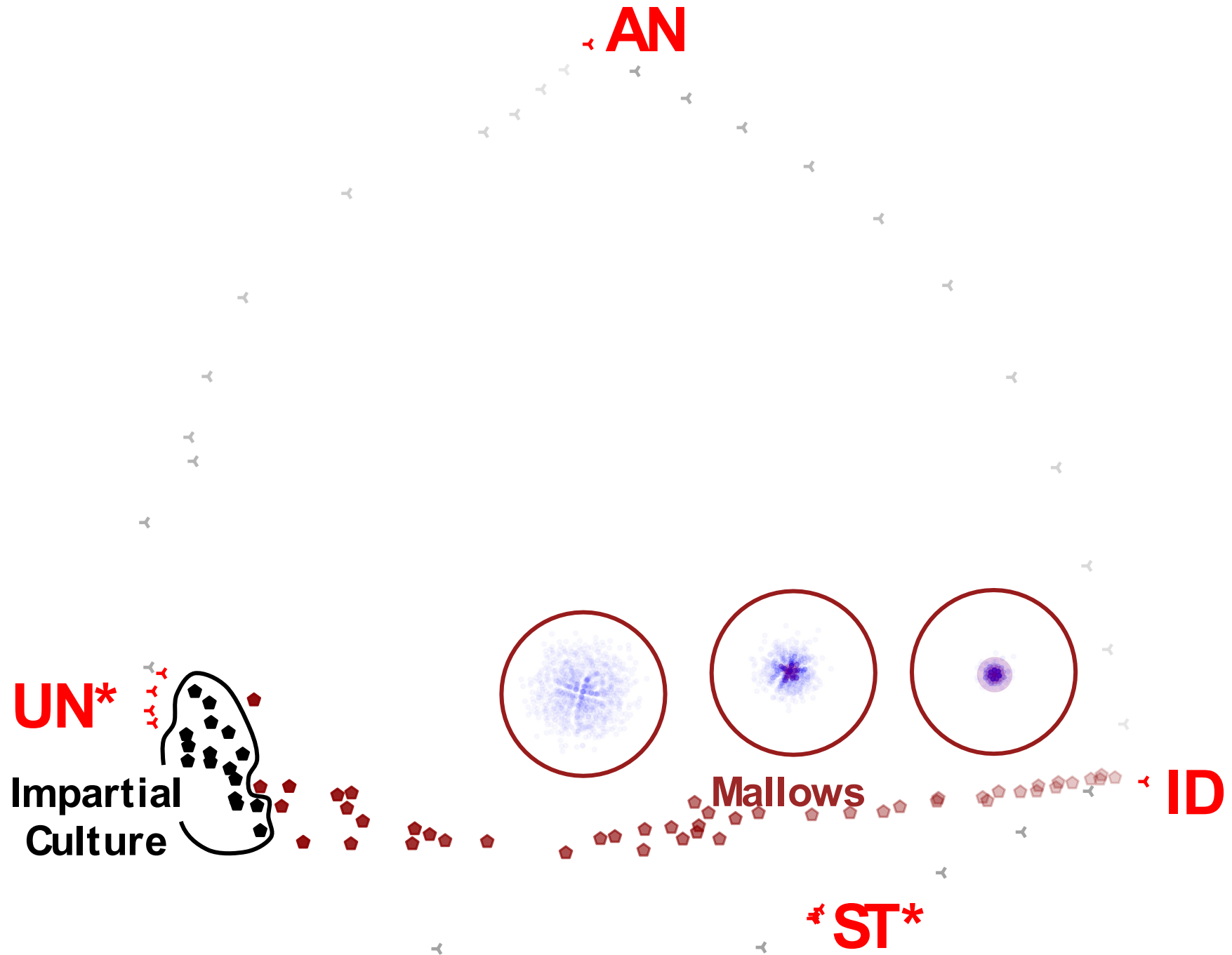
**Impartial  
Culture**

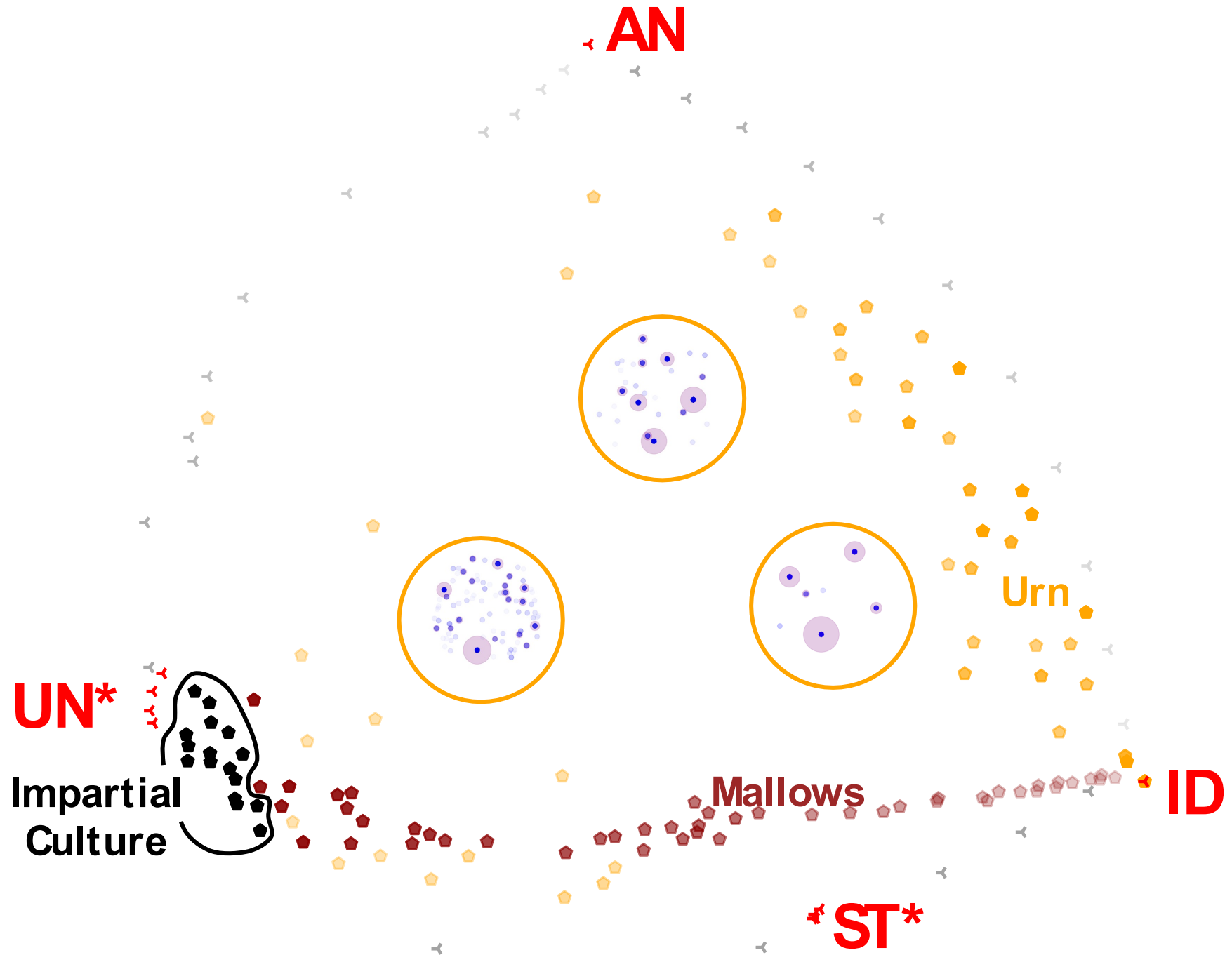


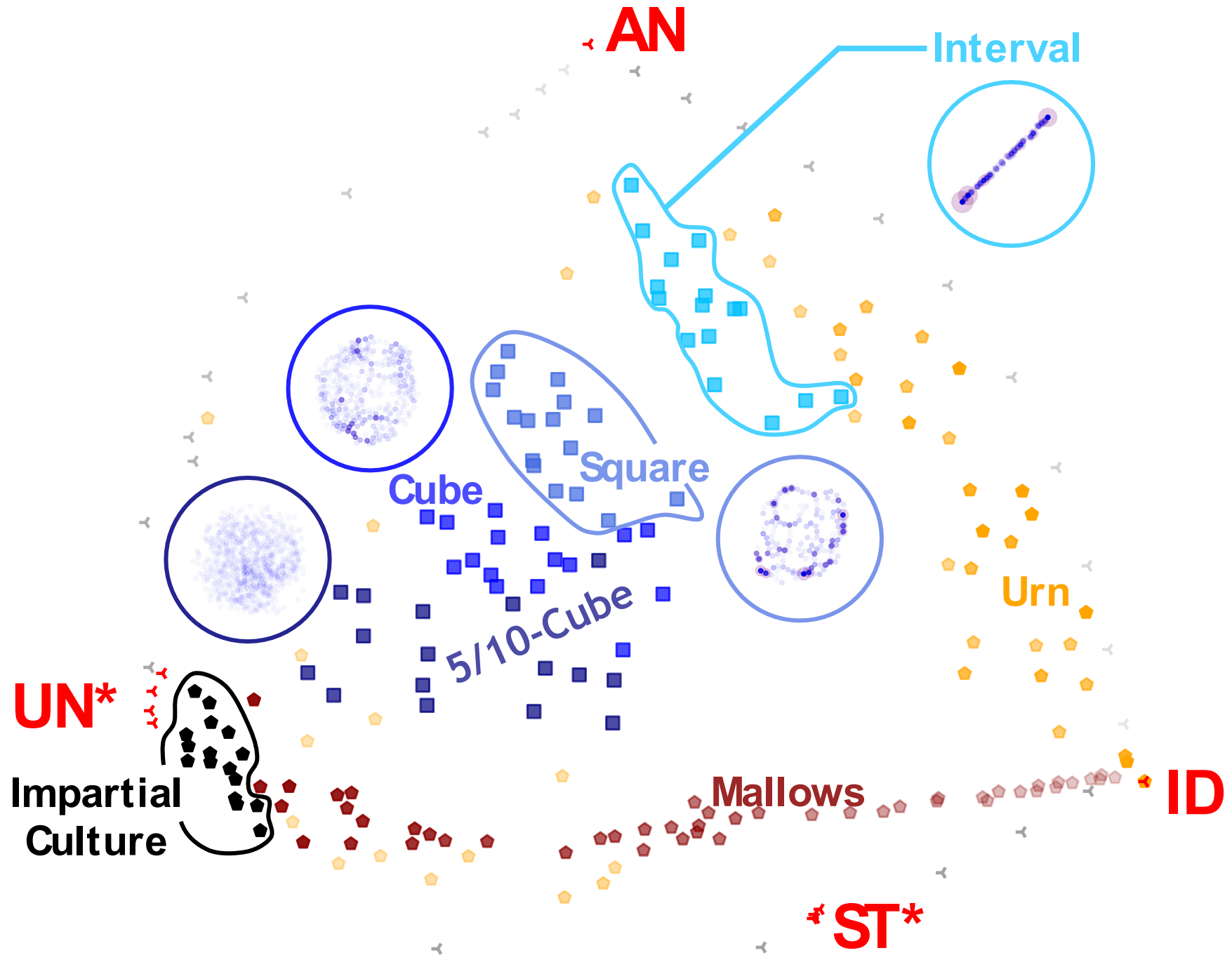
**AN**

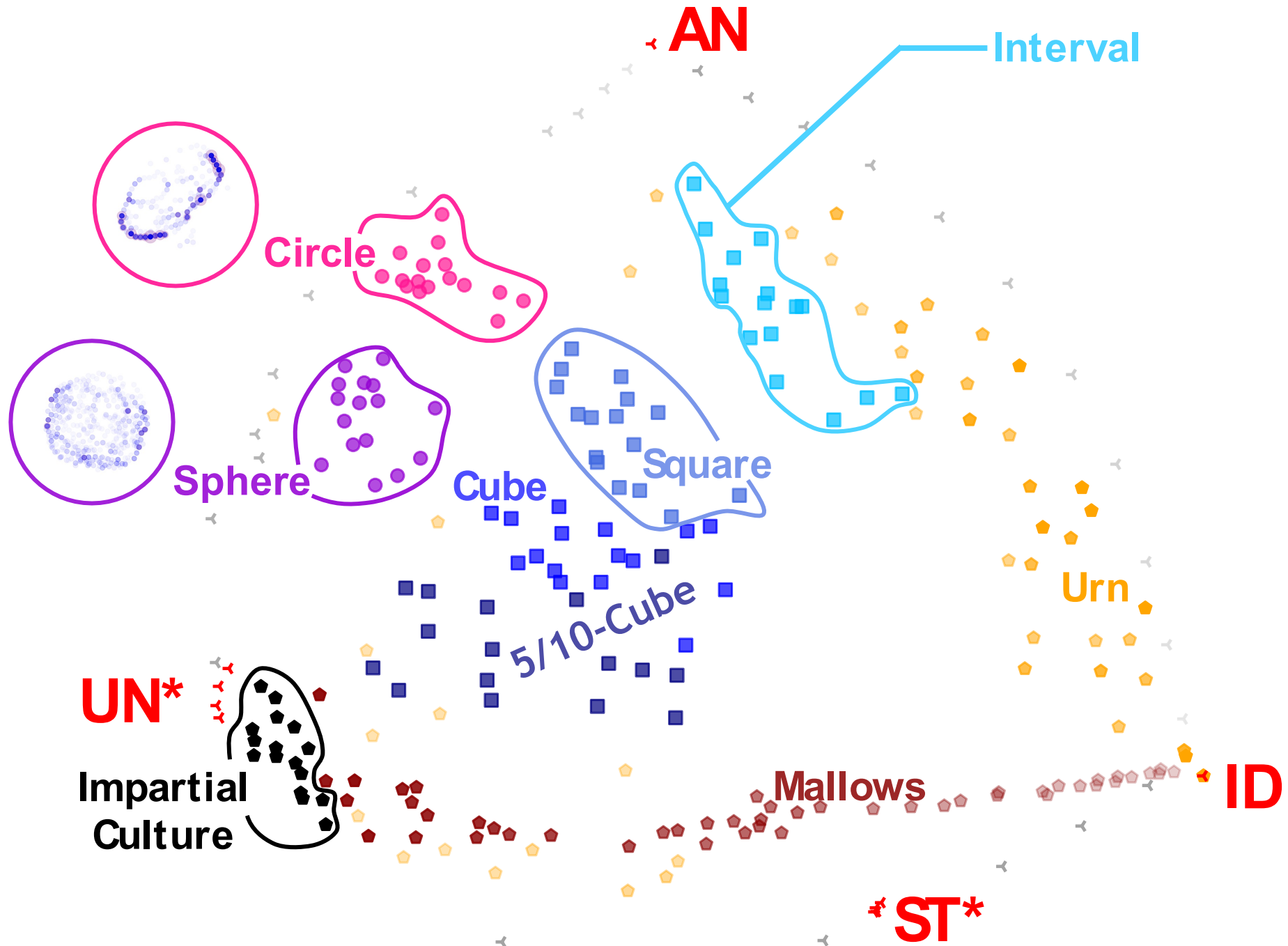
**\*ST\***

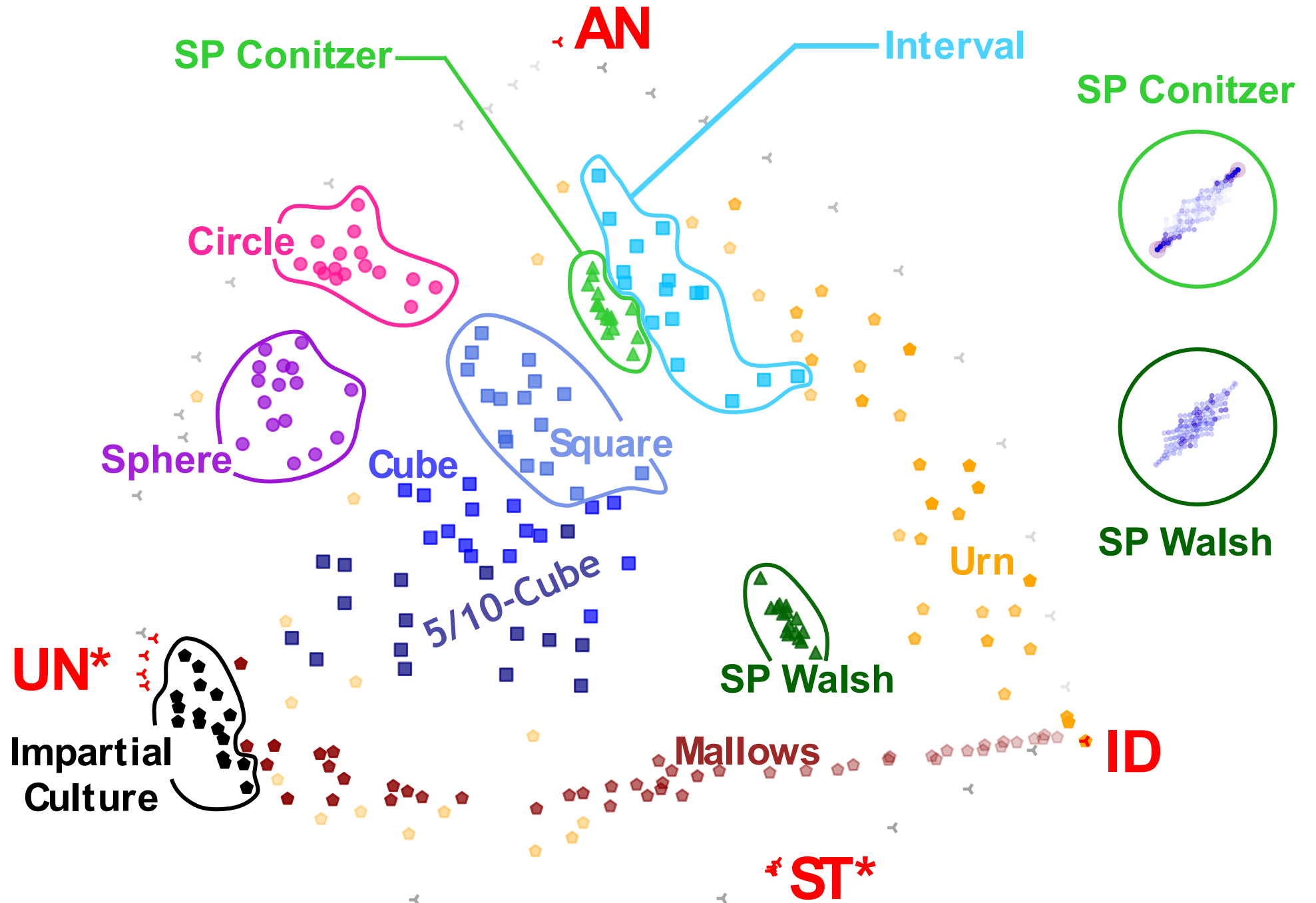
**ID**

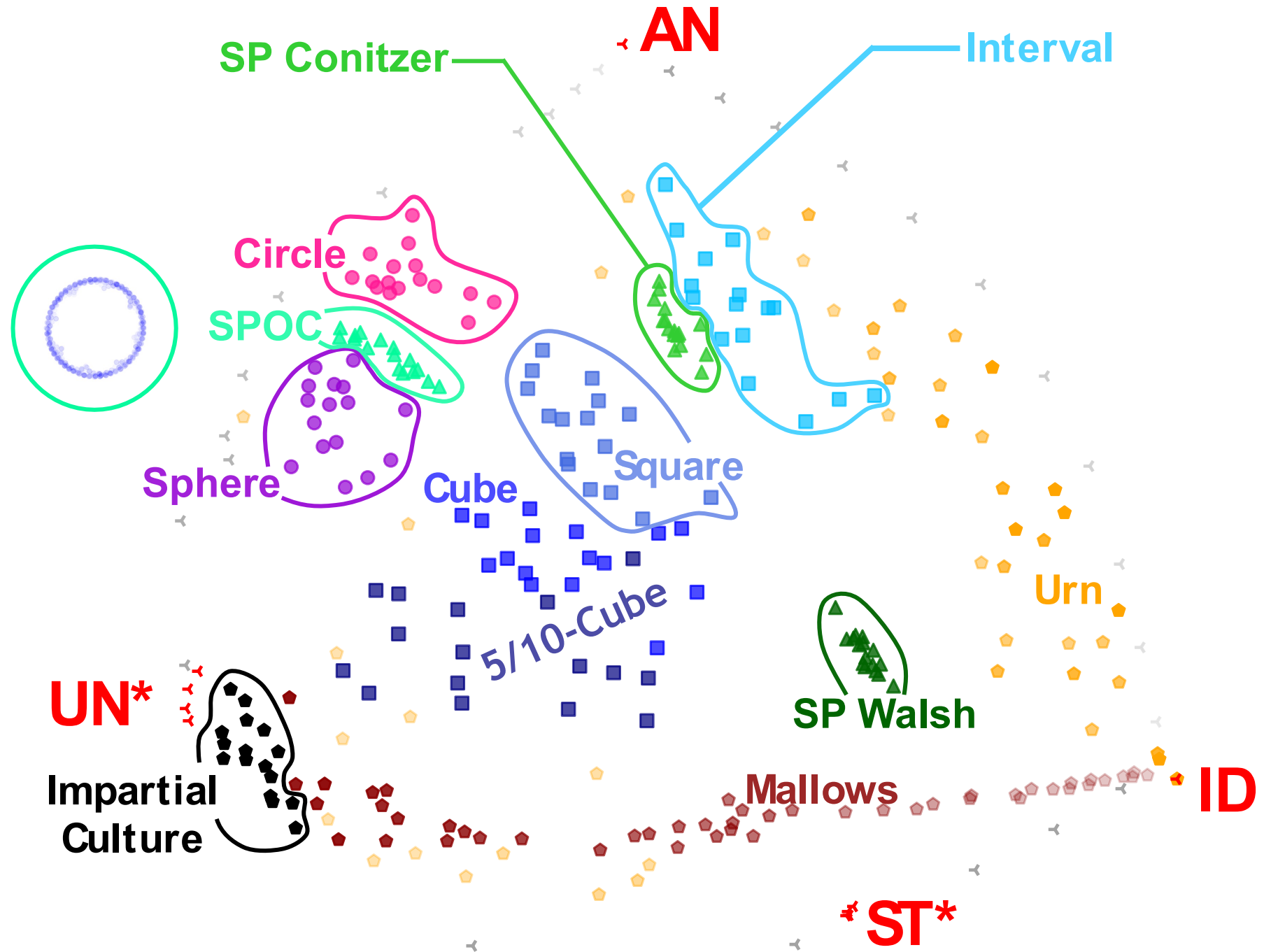




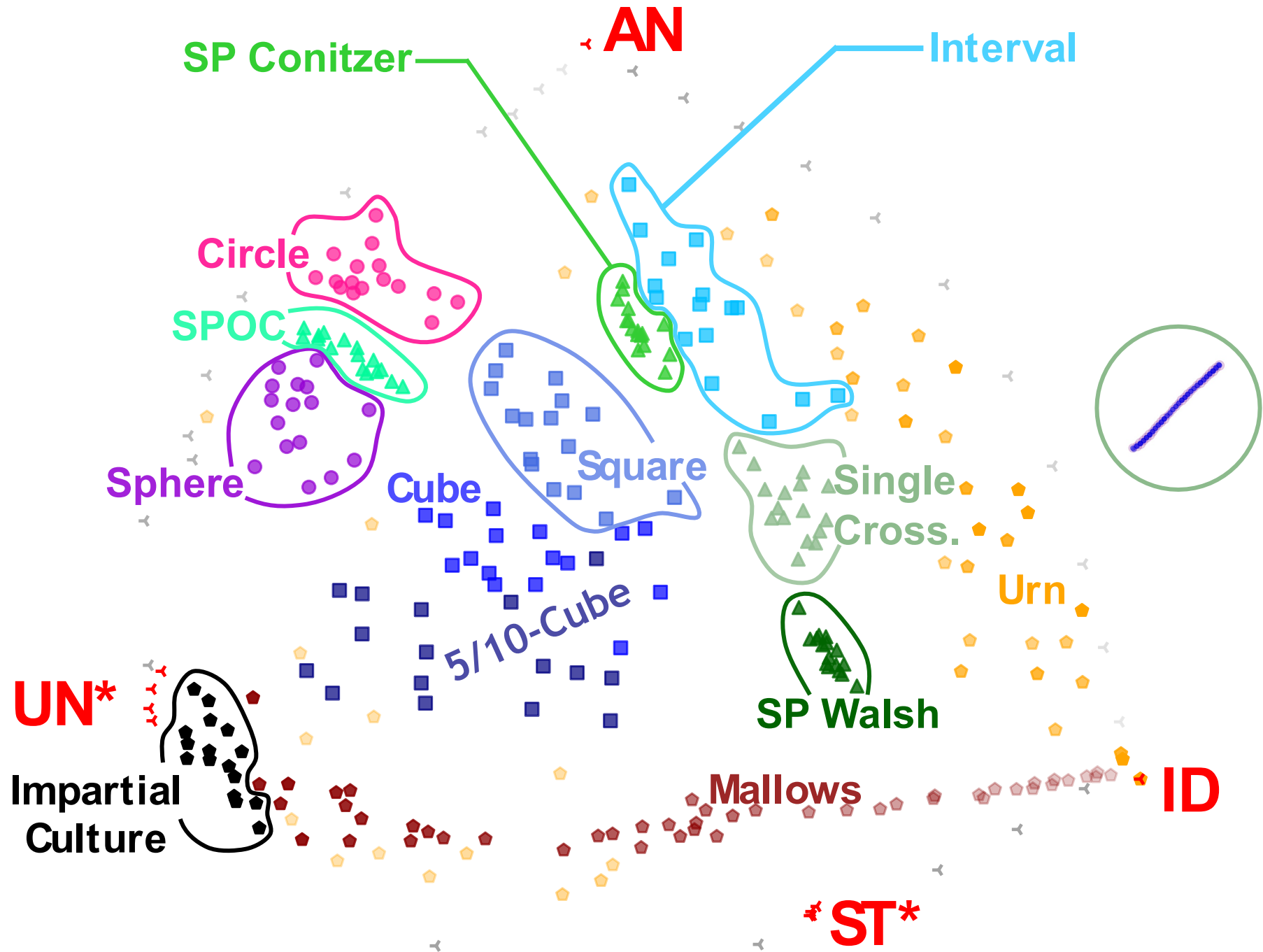


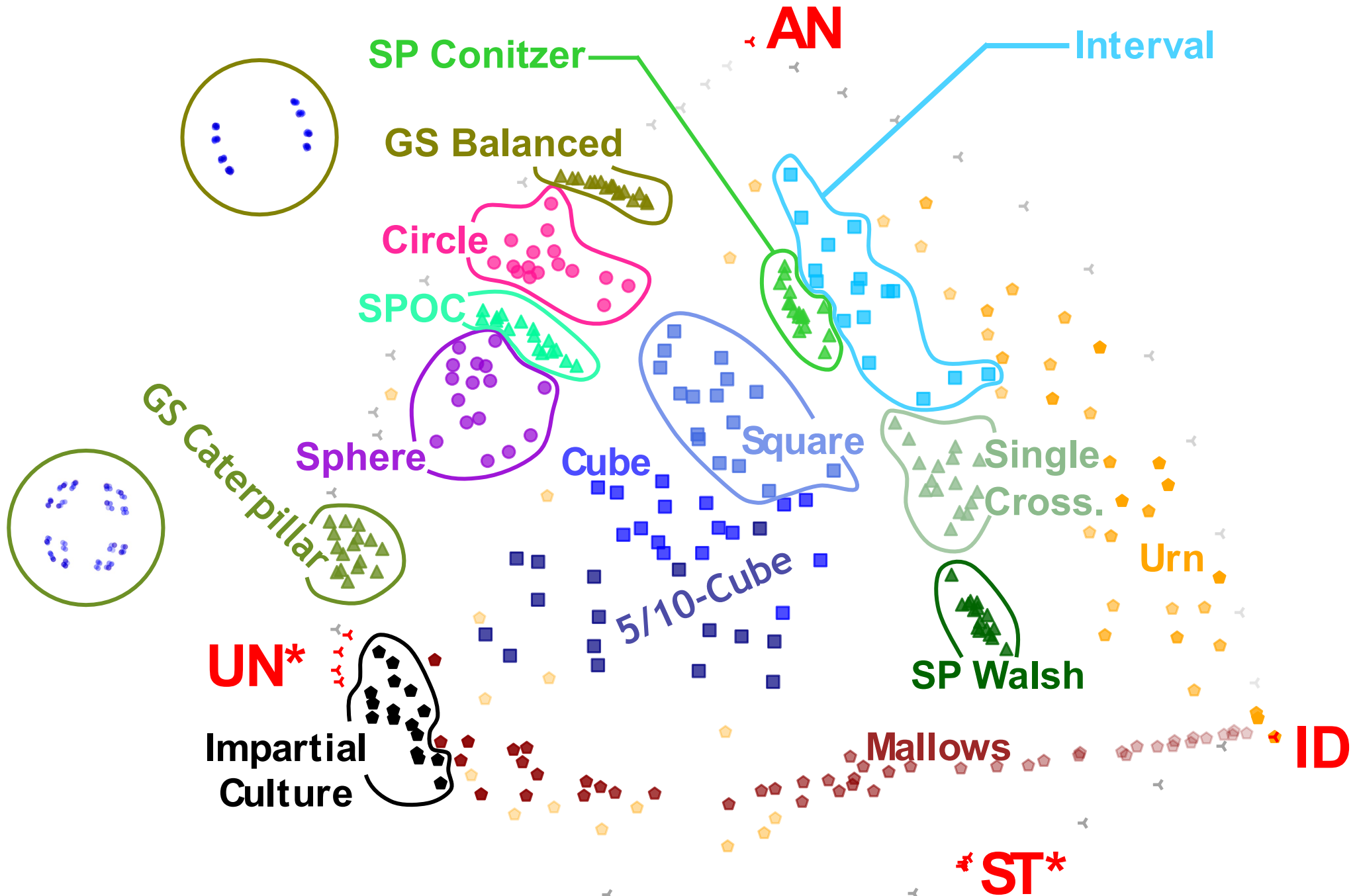


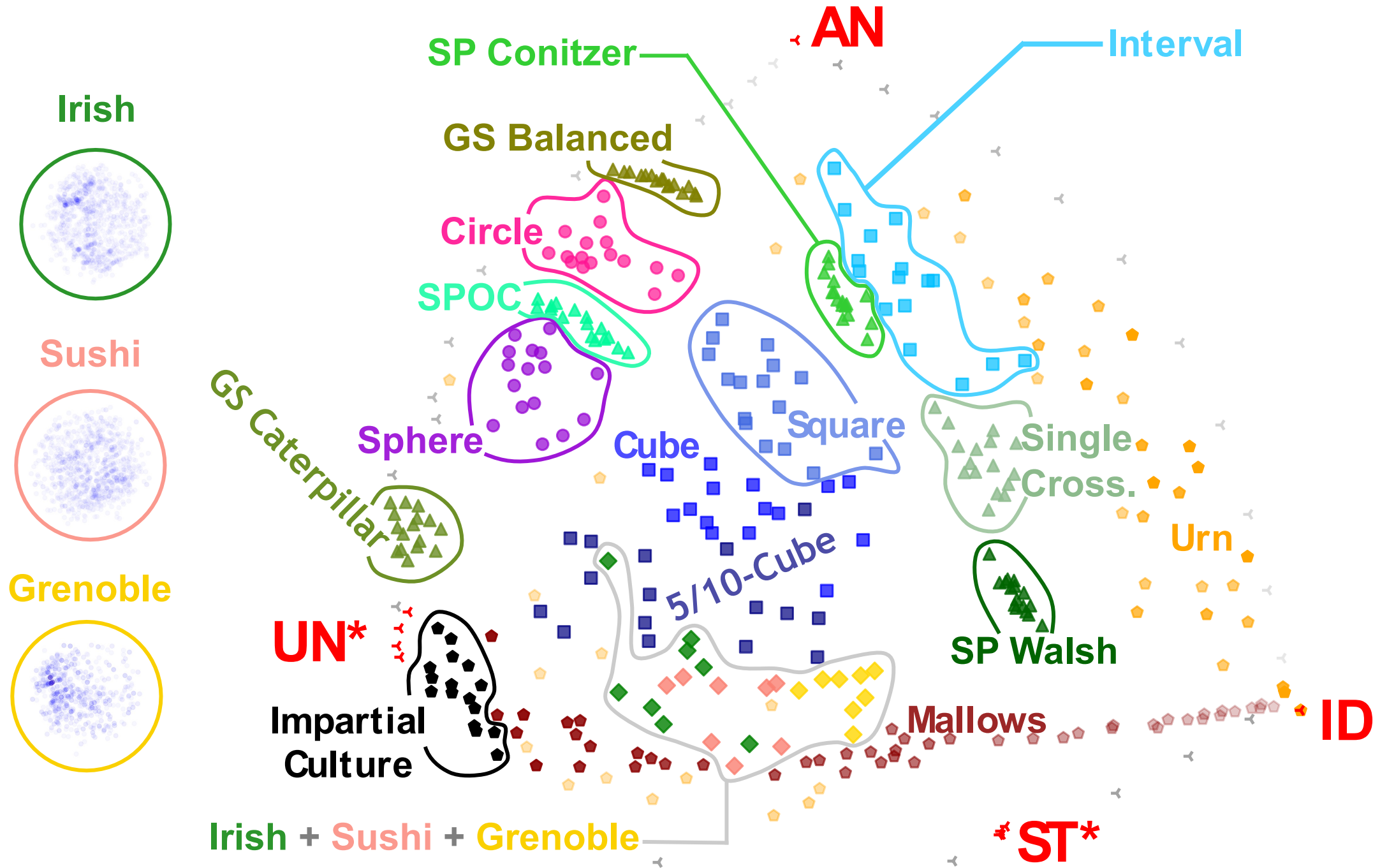










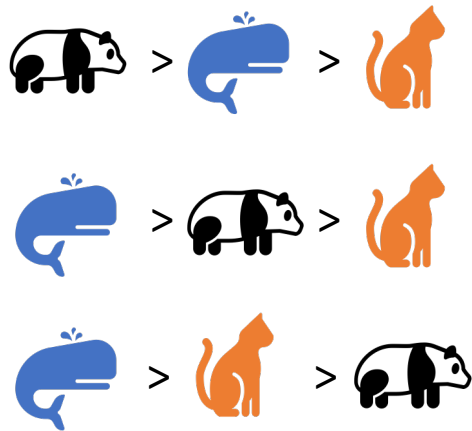


## Computing Isomorphic Swap distance is:

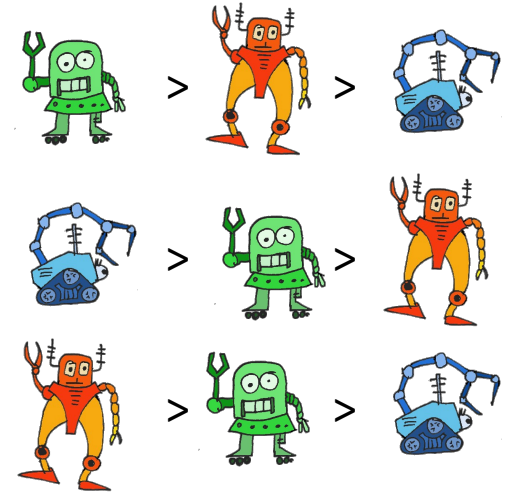
- NP-hard
- Hard to approximate
  - $O(m)$ -approx. and no better
- FPT-computable, but impractical
- Infeasible using ILP
- Just plain tough!
- Bruteforce works up to 10x50 elections, if you have hundreds of cores and plenty of time...




# How to Go Around Isomorphic Swap Distance?

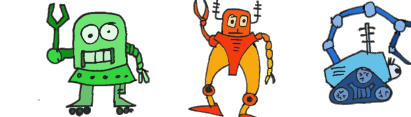


1. Match the candidates
2. Match the voters
3. Count the swaps



# How to Go Around Isomorphic Swap Distance?

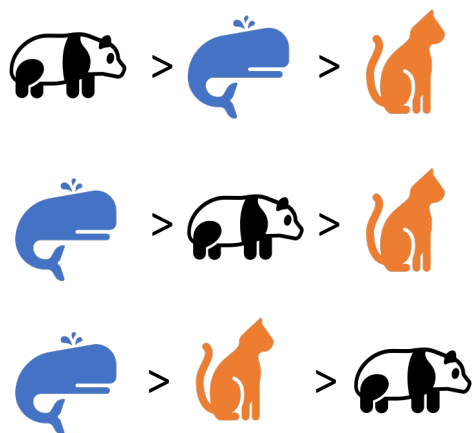


$$\begin{matrix}
 1. & \begin{bmatrix} 2 & 1 & 0 \end{bmatrix} \\
 2. & \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \\
 3. & \begin{bmatrix} 0 & 1 & 2 \end{bmatrix}
 \end{matrix}$$


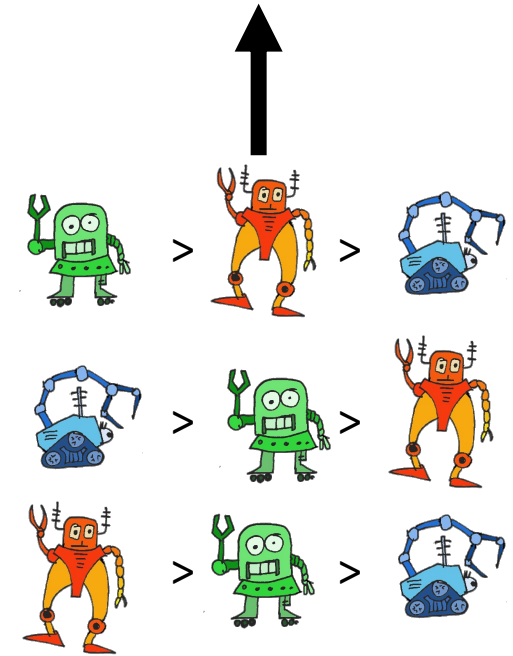
$$\begin{matrix}
 1. & \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \\
 2. & \begin{bmatrix} 2 & 1 & 0 \end{bmatrix} \\
 3. & \begin{bmatrix} 0 & 1 & 2 \end{bmatrix}
 \end{matrix}$$


2. Match the candidates  
3. Compute the distance

1. Compute Position Matrix



1. Match the candidates  
2. Match the voters  
3. Count the swaps



# Distance Between Vectors

$$\begin{bmatrix} 3 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & 0 & 1 & 0 & 1 \end{bmatrix}$$

# Distance Between Vectors

$\ell_1$ -distance

$$\begin{bmatrix} 3 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & 0 & 1 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 2 & 0 & 0 & 1 & 1 \end{bmatrix} \quad 6$$



# Distance Between Vectors

$\ell_1$ -distance

$$\begin{bmatrix} 3 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 0 & 3 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 & 0 & 2 & 1 & 1 \end{bmatrix} \quad 6$$

# Distance Between Vectors

$\ell_1$ -distance

$$\begin{bmatrix} 3 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

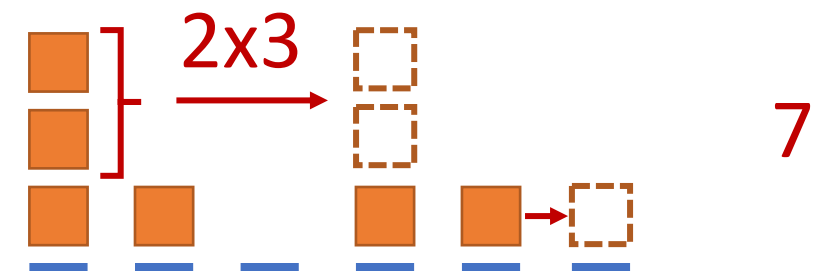
$$\begin{bmatrix} 1 & 1 & 0 & 3 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 & 0 & 2 & 1 & 1 \end{bmatrix} \quad 6$$

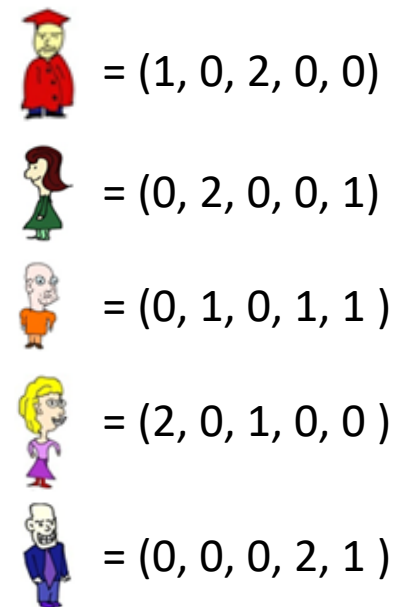
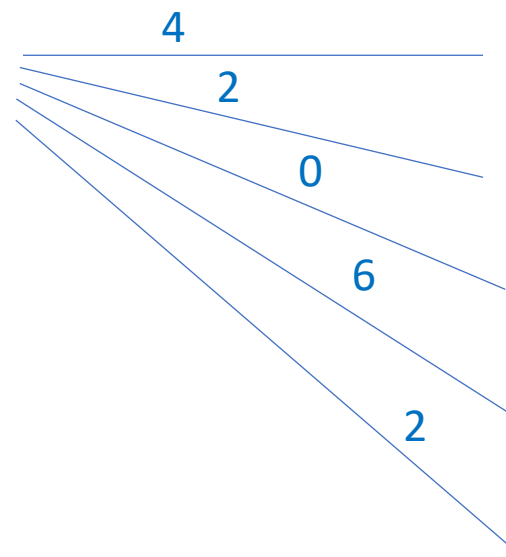
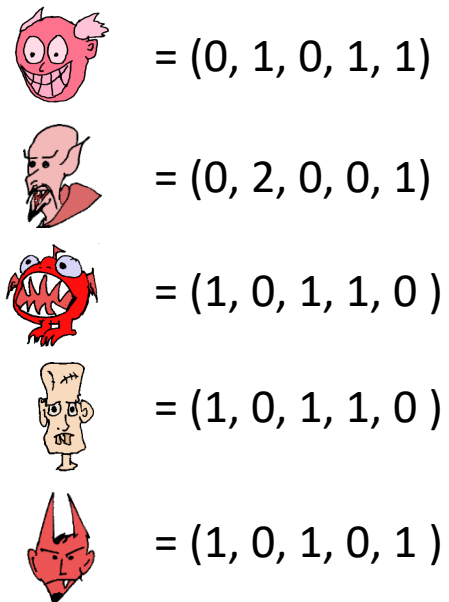
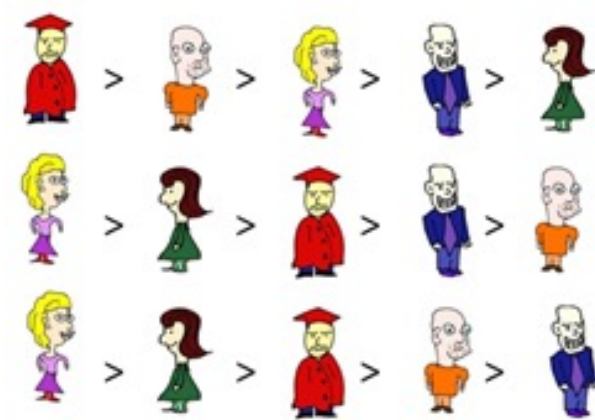
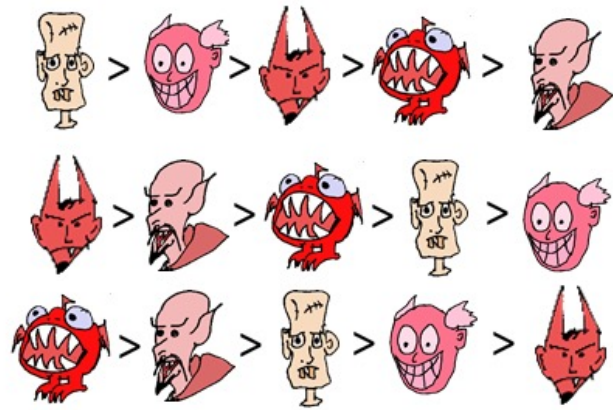
Earth Mover's Distance (EMD)

$$\begin{bmatrix} 3 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 0 & 3 & 0 & 1 \end{bmatrix}$$

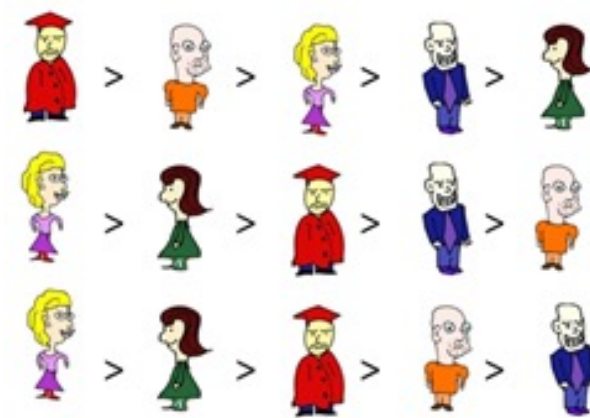
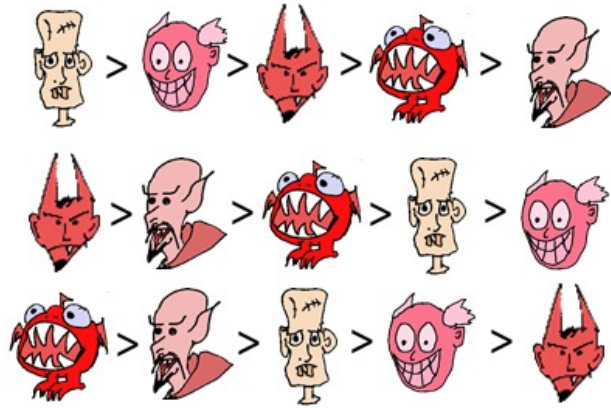



# Positionwise Distance





Earth mover distances


# Positionwise Distance




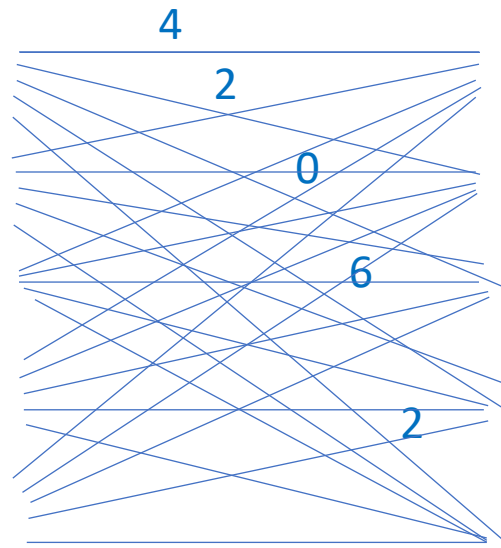
 = (0, 1, 0, 1, 1)


 = (0, 2, 0, 0, 1)


 = (1, 0, 1, 1, 0)


 = (1, 0, 1, 1, 0)


 = (1, 0, 1, 0, 1)




 = (1, 0, 2, 0, 0)

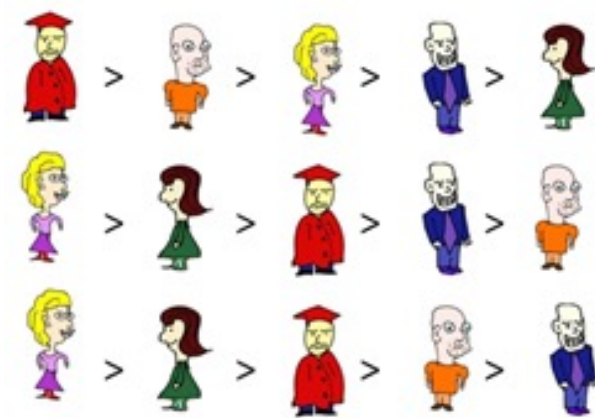
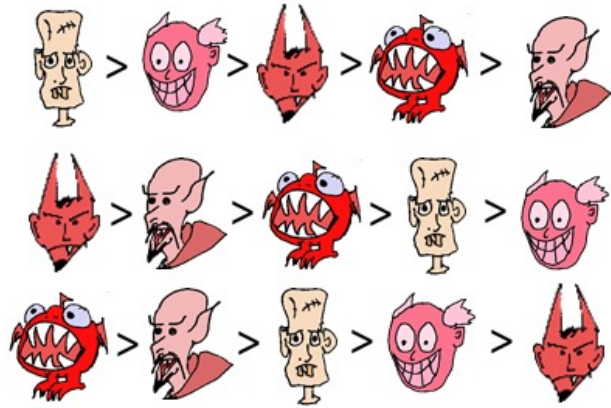
 = (0, 2, 0, 0, 1)


 = (0, 1, 0, 1, 1)


 = (2, 0, 1, 0, 0)


 = (0, 0, 0, 2, 1)


# Positionwise Distance





 = (0, 1, 0, 1, 1)


 = (0, 2, 0, 0, 1)


 = (1, 0, 1, 1, 0)


 = (1, 0, 1, 1, 0)


 = (1, 0, 1, 0, 1)

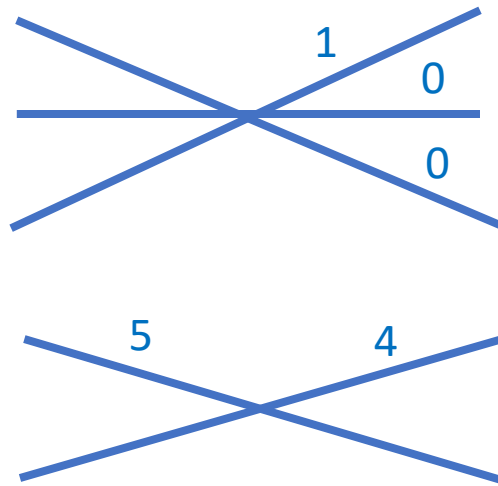
 = (1, 0, 2, 0, 0)

 = (0, 2, 0, 0, 1)

 = (0, 1, 0, 1, 1)

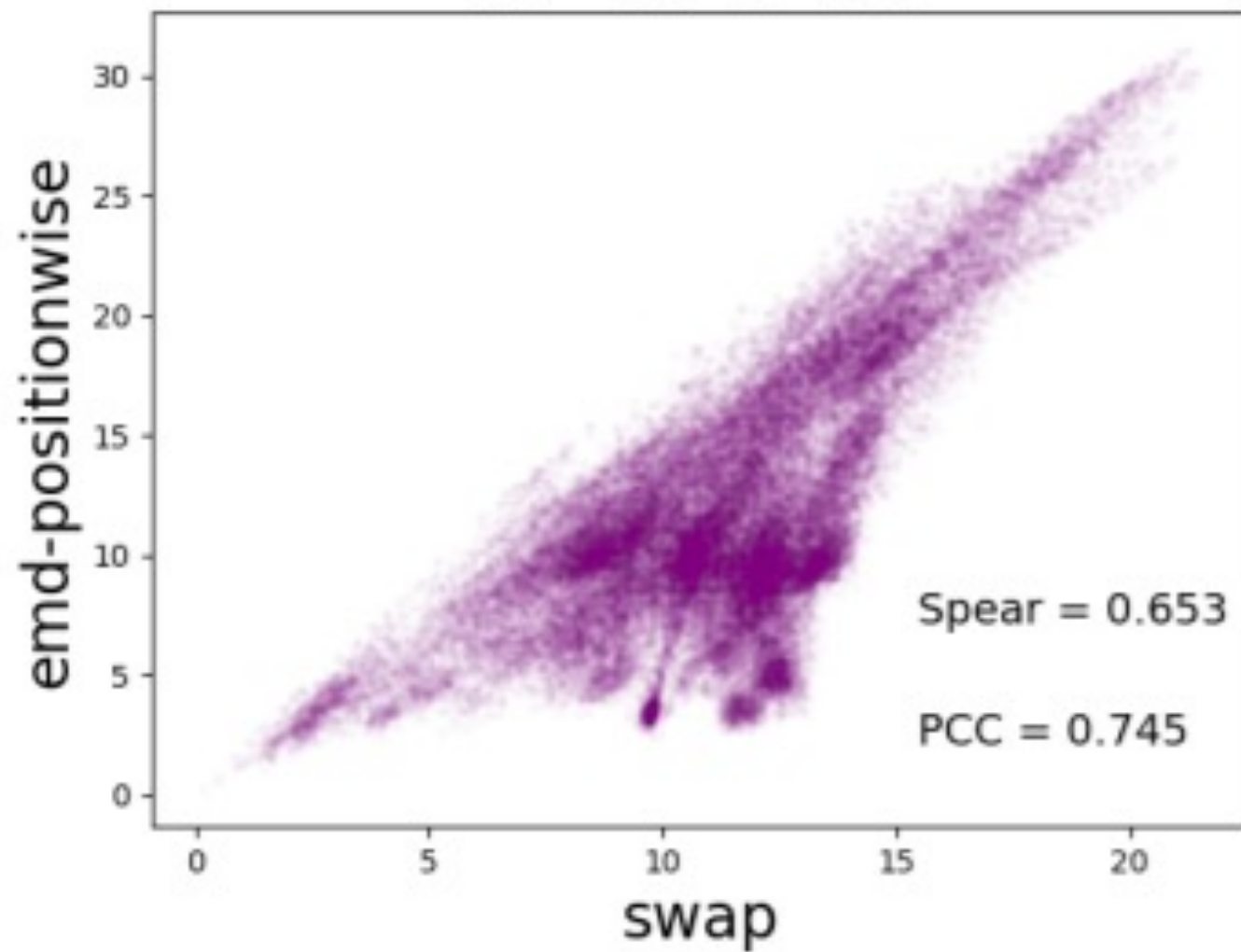
 = (2, 0, 1, 0, 0)

 = (0, 0, 0, 2, 1)



distance = 1+0+0+5+4 = 10

m=10 n=50



two reverse orders  
**antagonism**



AN

- V<sub>1</sub>: > > >
- V<sub>2</sub>: > > >
- V<sub>3</sub>: > > >
- V<sub>4</sub>: > > >
- V<sub>5</sub>: > > >
- V<sub>6</sub>: > > >

- V<sub>1</sub>: > > >
- V<sub>2</sub>: > > >
- V<sub>3</sub>: > > >
- V<sub>4</sub>: > > >
- V<sub>5</sub>: > > >
- V<sub>6</sub>: > > >



UN

1



ID

- V<sub>1</sub>: > > >
- V<sub>2</sub>: > > >
- V<sub>3</sub>: > > >
- V<sub>4</sub>: > > >
- V<sub>5</sub>: > > >
- V<sub>6</sub>: > > >

- V<sub>1</sub>: > > >
- V<sub>2</sub>: > > >
- V<sub>3</sub>: > > >
- V<sub>4</sub>: > > >
- V<sub>5</sub>: > > >
- V<sub>6</sub>: > > >



ST

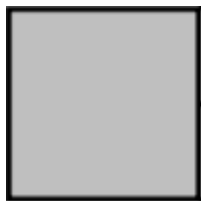
two groups of candidates,  
each voter prefers members  
of one group to the other

**stratification**



- V<sub>1</sub>: > > >
- V<sub>2</sub>: > > >
- V<sub>3</sub>: > > >
- V<sub>4</sub>: > > >
- V<sub>5</sub>: > > >
- V<sub>6</sub>: > > >

- V<sub>1</sub>: > > >
- V<sub>2</sub>: > > >
- V<sub>3</sub>: > > >
- V<sub>4</sub>: > > >
- V<sub>5</sub>: > > >
- V<sub>6</sub>: > > >



1

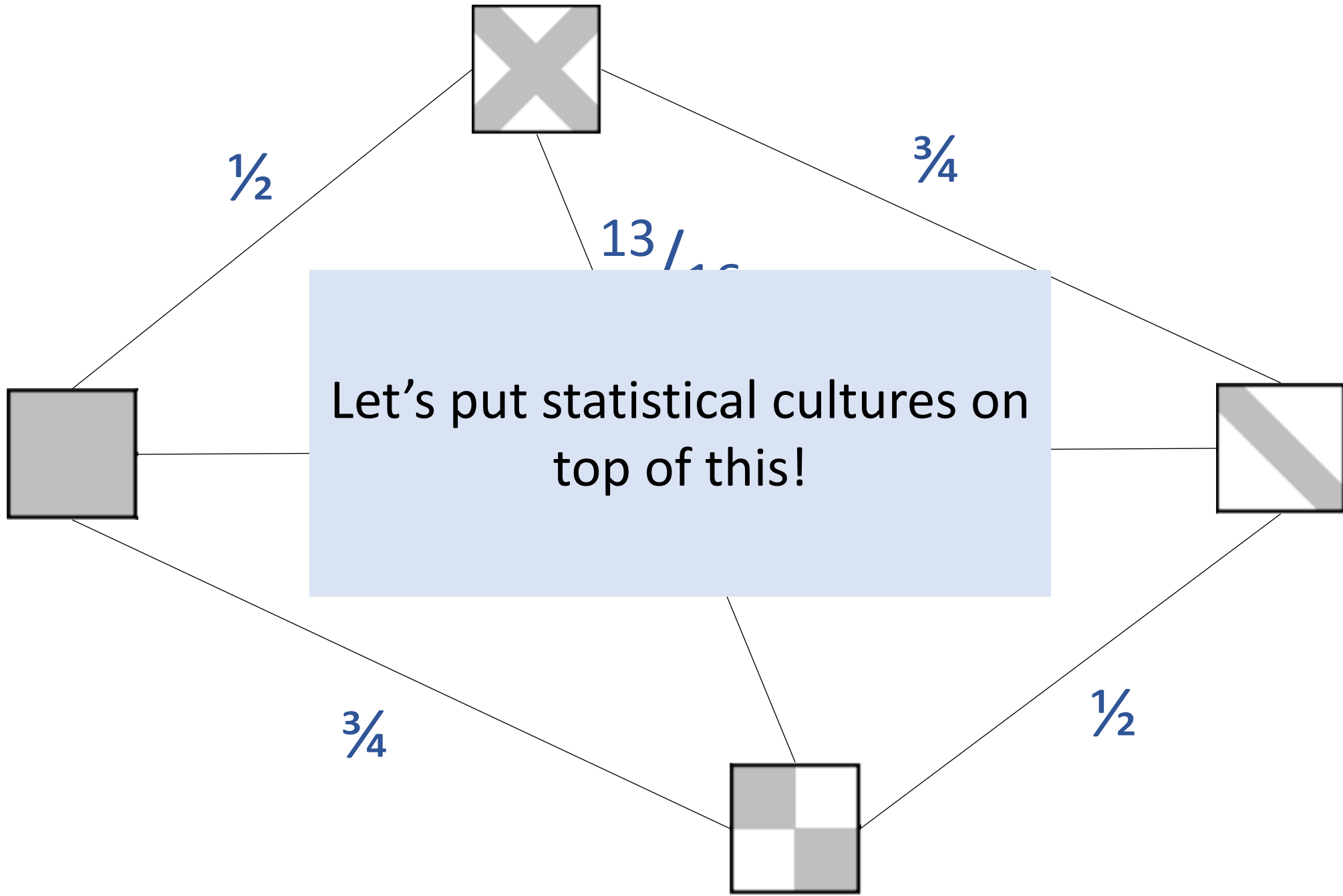


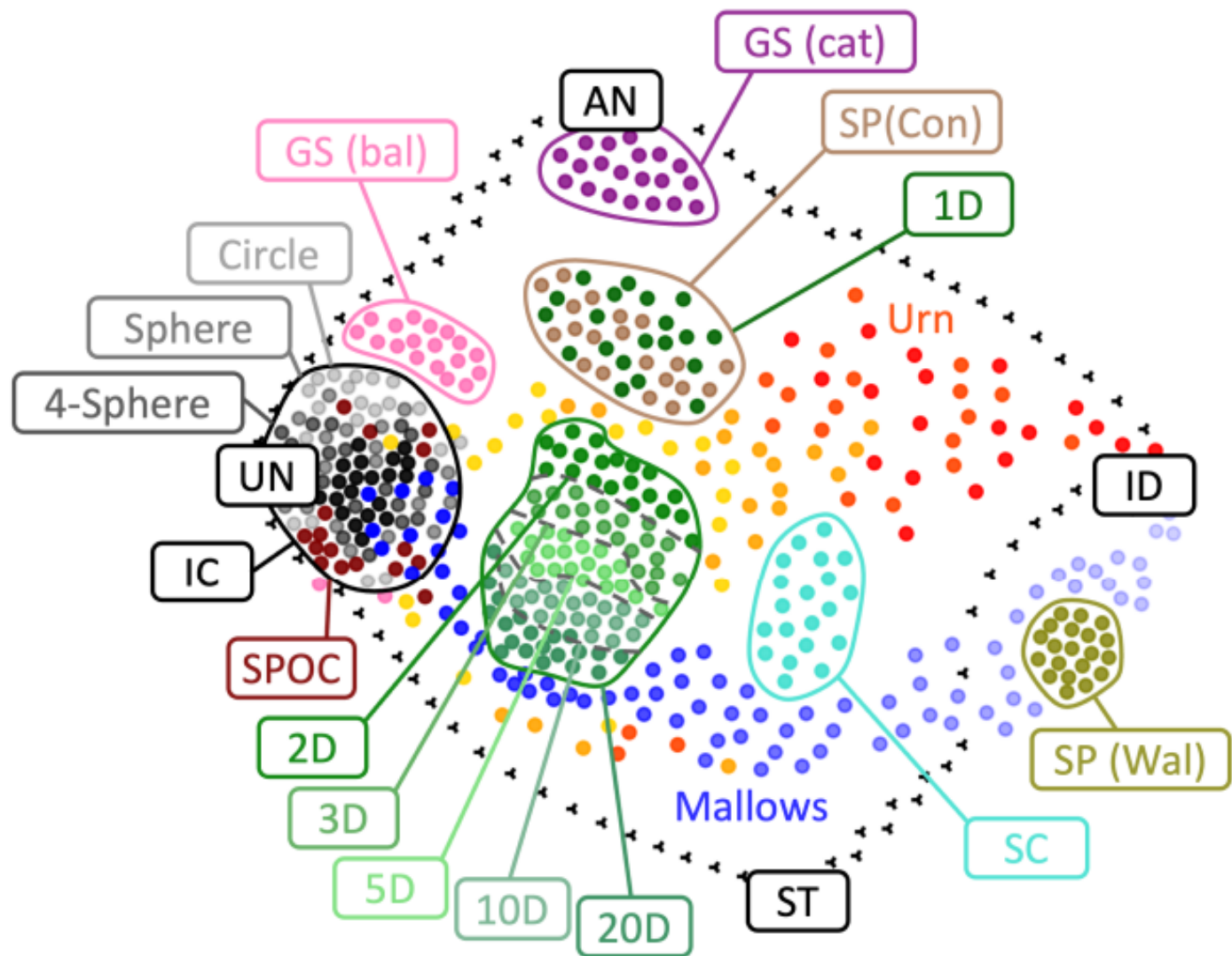
- V<sub>1</sub>: > > >
- V<sub>2</sub>: > > >
- V<sub>3</sub>: > > >
- V<sub>4</sub>: > > >
- V<sub>5</sub>: > > >
- V<sub>6</sub>: > > >

- V<sub>1</sub>: > > >
- V<sub>2</sub>: > > >
- V<sub>3</sub>: > > >
- V<sub>4</sub>: > > >
- V<sub>5</sub>: > > >
- V<sub>6</sub>: > > >

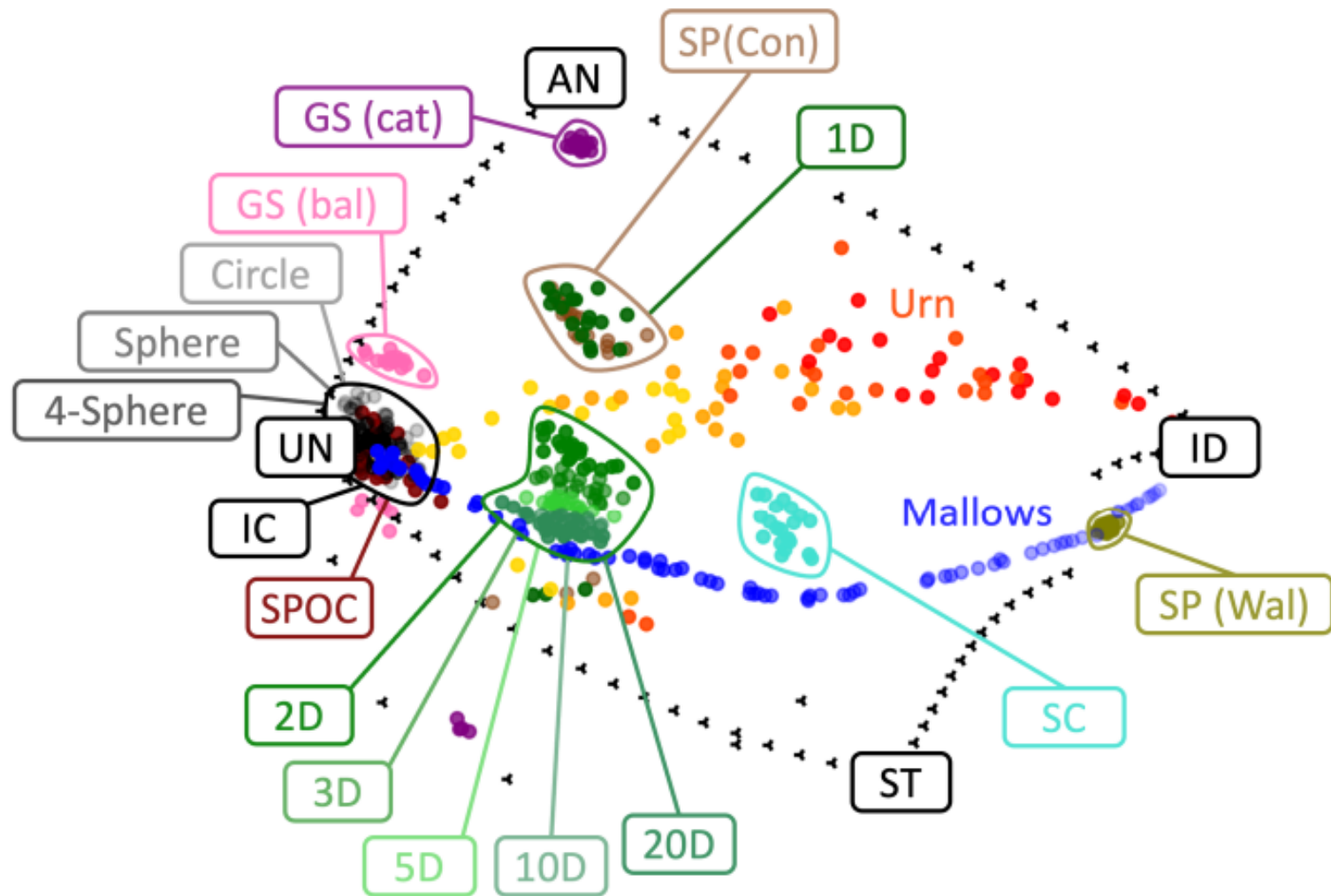




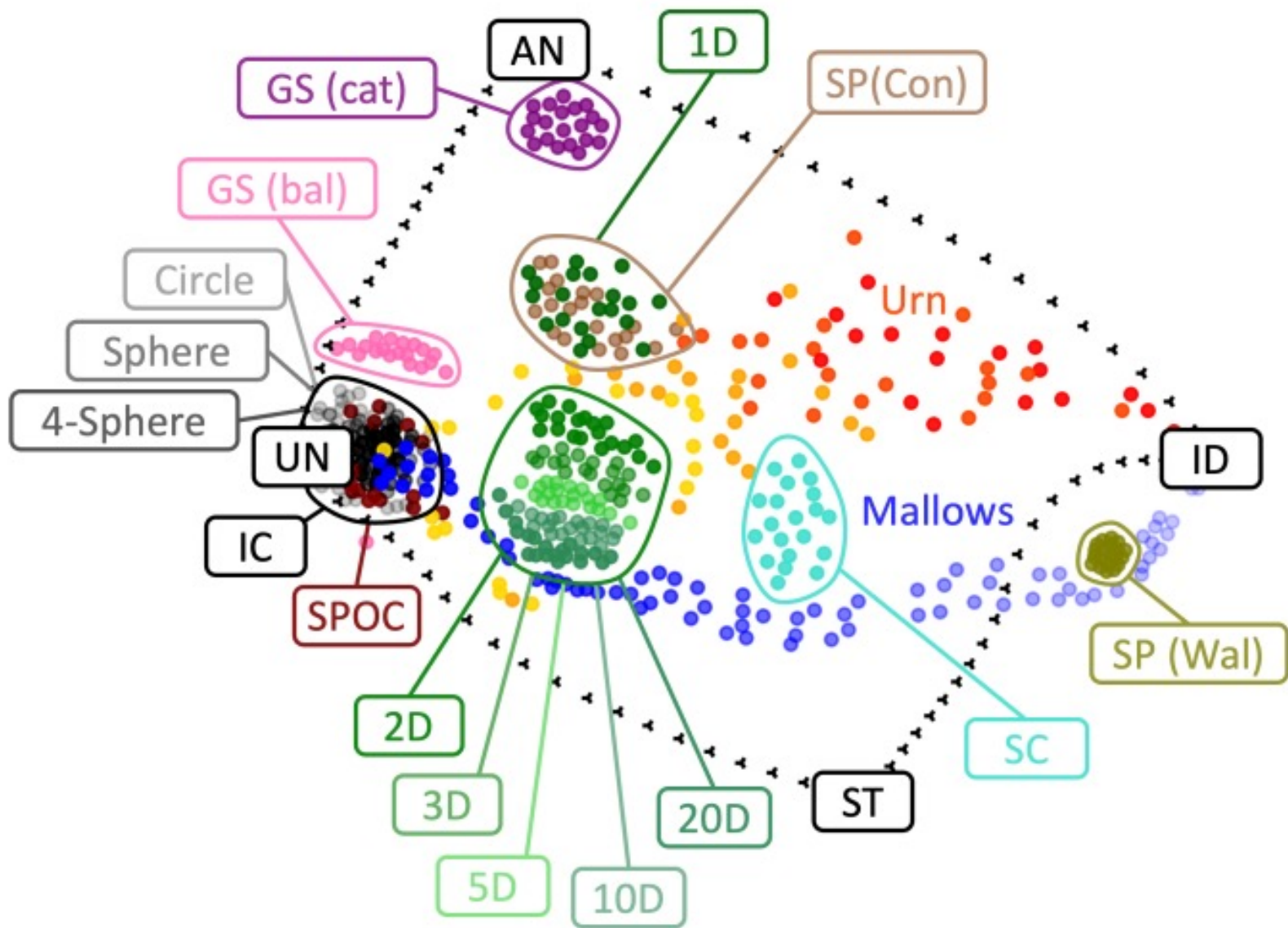




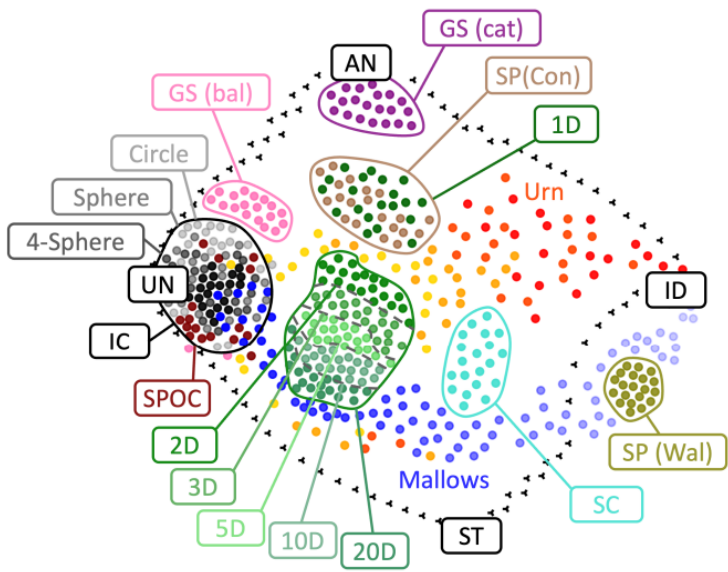
(a) FR



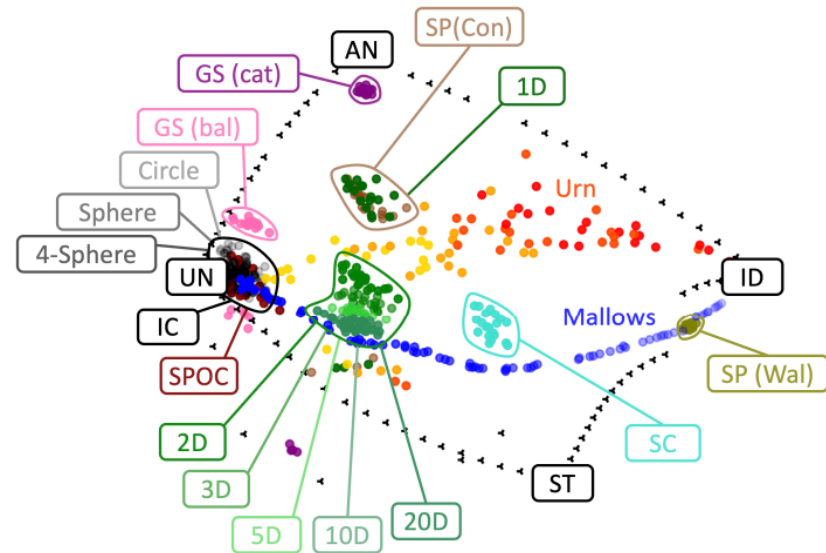
(b) MDS



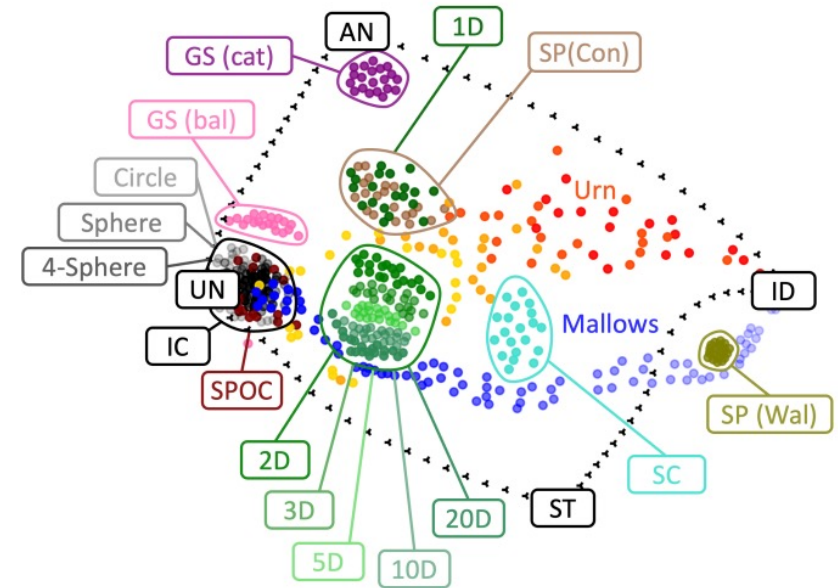
(c) KK



(a) FR



(b) MDS

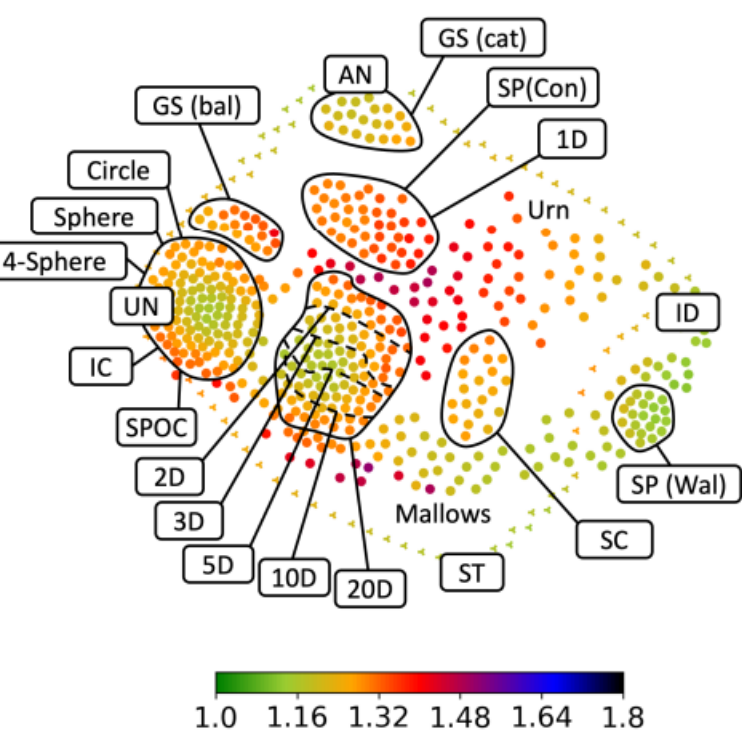


(c) KK

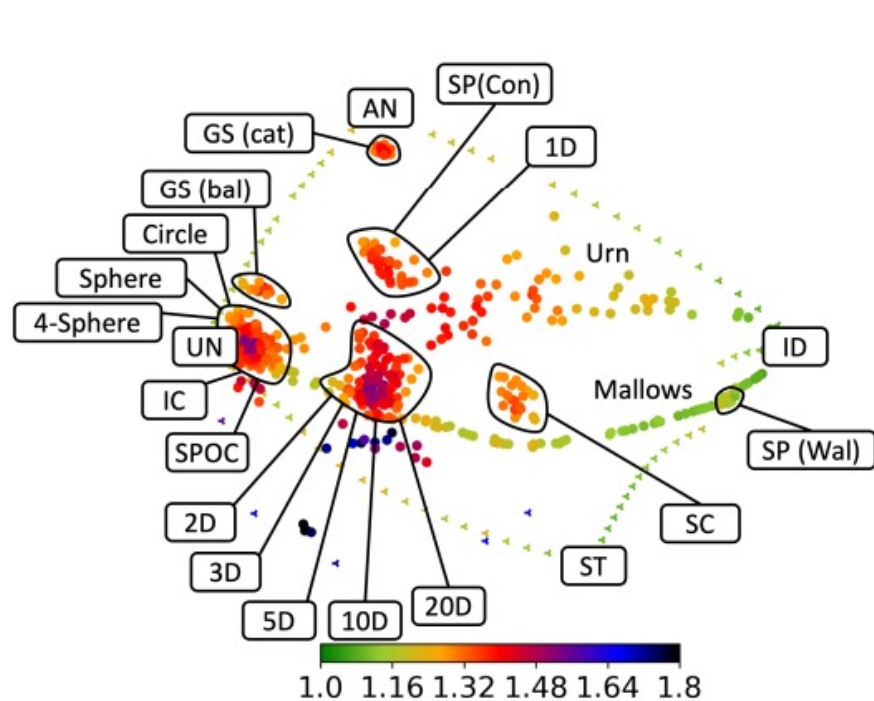
# Which embedding is best?

$$\text{MR}(X, Y) = \frac{\max(\bar{d}_{\text{Euc}}(X, Y), \bar{d}_{\mathcal{M}}(X, Y))}{\min(\bar{d}_{\text{Euc}}(X, Y), \bar{d}_{\mathcal{M}}(X, Y))},$$

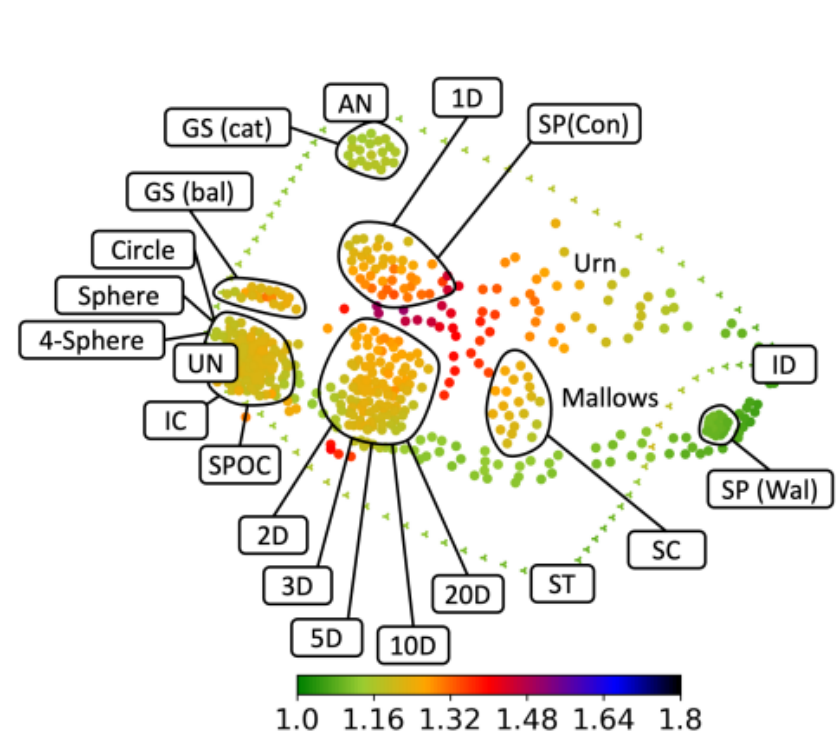




(a) FR



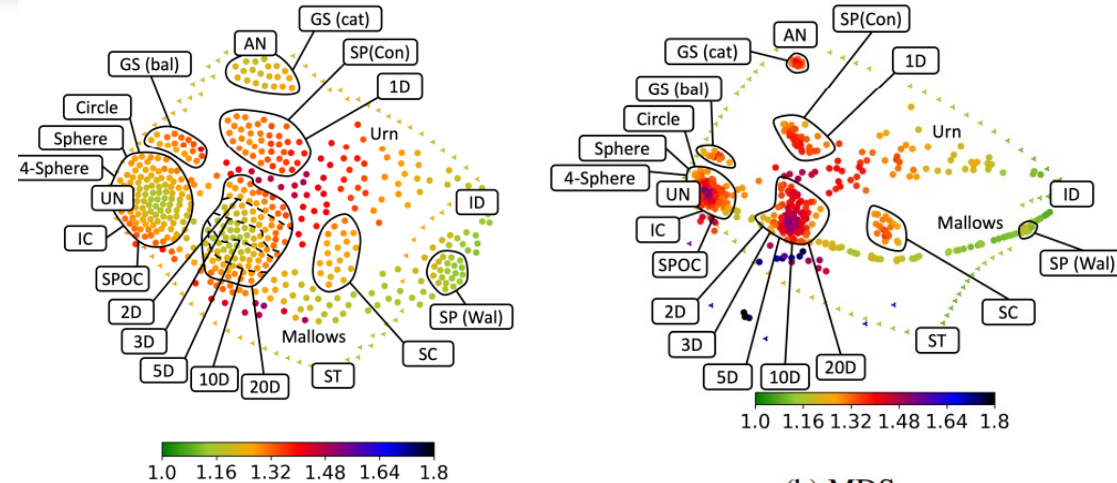
(b) MDS



(c) KK

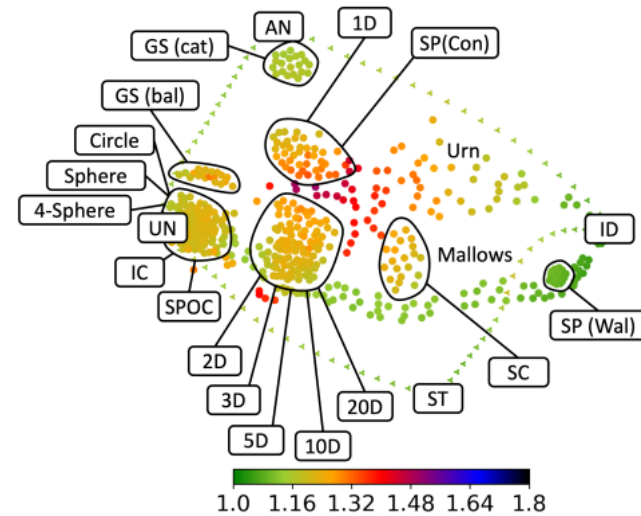
dataset	average total distortion values		
	FR	MDS	KK
$4 \times 100$	$1.3213 \pm 0.0157$	$1.3099 \pm 0.0076$	$1.2612 \pm 0.0158$
$10 \times 100$	$1.3119 \pm 0.0194$	$1.3531 \pm 0.0108$	$1.2625 \pm 0.0125$
$20 \times 100$	$1.2979 \pm 0.0195$	$1.3545 \pm 0.0126$	$1.2406 \pm 0.0060$
$100 \times 100$	$1.3006 \pm 0.0256$	$1.3225 \pm 0.0194$	$1.2119 \pm 0.0123$

Model	average total distortion values		
	FR	MDS	KK
Impartial Culture	1.145	1.087	1.07
Single-Peaked (Conitzer)	1.313	1.305	1.244
Single-Peaked (Walsh)	1.114	1.067	1.071
SPOC	1.223	1.094	1.081
Single-Crossing	1.256	1.298	1.225
Interval	1.321	1.3	1.233
Square	1.267	1.274	1.203
Cube	1.216	1.217	1.146
5-Cube	1.155	1.177	1.114
10-Cube	1.2	1.162	1.094
20-Cube	1.252	1.162	1.097
Circle	1.222	1.105	1.101
Sphere	1.187	1.09	1.077
4-Sphere	1.174	1.084	1.072
Group-Separable (Balanced)	1.302	1.298	1.204
Group-Separable (Caterpillar)	1.215	1.218	1.14
Urn	1.338	1.298	1.285
Mallows	1.195	1.121	1.094
All	1.241	1.198	1.159



(a) FR

(b) MDS



(c) KK

**Mapel**

Matchings

Further Applications

Approval Elections

Map of Rules

Data!

Introduction to voting

Experiments in Computational Social Choice

Preference Learning

Mallows

Real-Life Data

Map of Elections

Use Cases (Elections)

Swap Distance

Distances

Approximations

Force-Directed

Positionwise

Embedding Algorithms

UN

Compass Elections

ID

Winners

Election Results

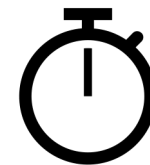
Verification

ST

Committees

Running Time





15 minutes

# Create your own map of elections!

Introduction to Mapel Software Package 1/2

Mapel

Matchings

Further Applications

Approval Elections

Map of Rules

Data!

Introduction to voting

Experiments in Computational Social Choice

Preference Learning

Mallows

Real-Life Data

Map of Elections

Use Cases (Elections)

Swap Distance

Approximations

Distances

Winners

Election Results

Positionwise

Embedding Algorithms

Force-Directed

UN

Compass Elections

ID

ST

Committees

Running Time

Verification

**Mapel**

Matchings

**Further Applications**

Approval Elections

Map of Rules

Data!

Introduction to voting

**Experiments in Computational Social Choice**

Preference Learning

Mallows

Real-Life Data

Swap Distance

Map of Elections

**Use Cases (Elections)**

Distances

Positionwise

Embedding Algorithms

Force-Directed

Verification

UN

Compass Elections

ST

ID

AN

Winners

Election Results

Approximations

Committees

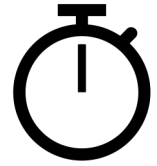
Running Time



Drawing a Map of Elections in the Space of Statistical Cultures, Szufa et al., AAMAS-20



Map of Elections, S. Szufa, PhD Thesis



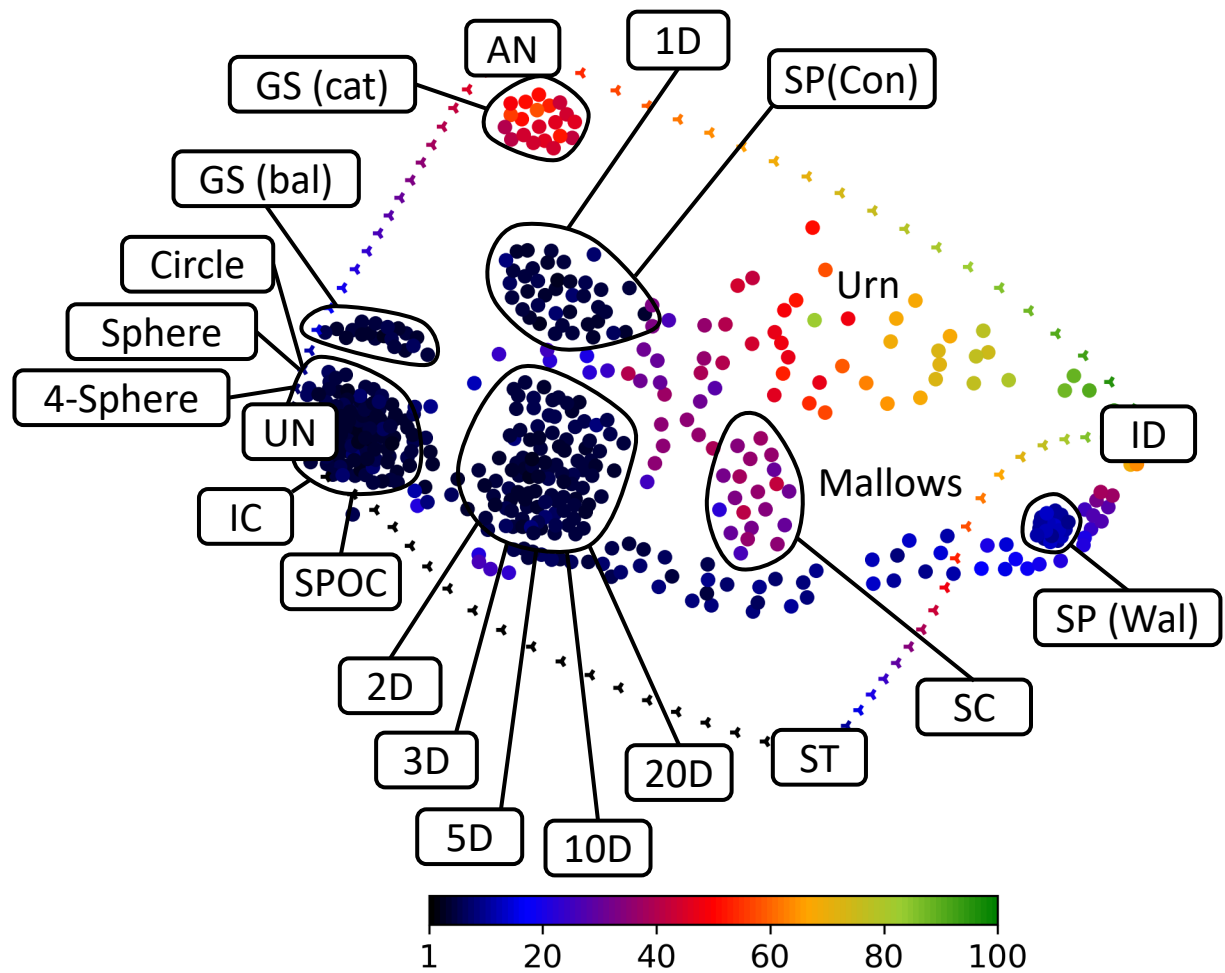
20 minutes

# Visualizing Experiment Results

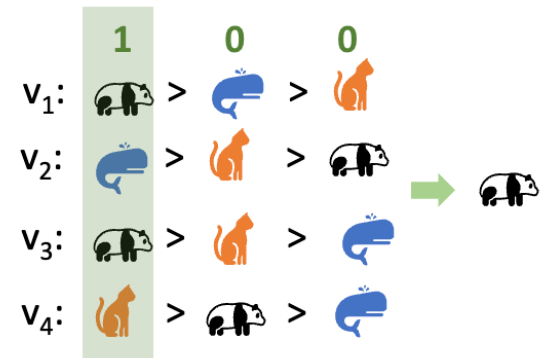
Use Cases

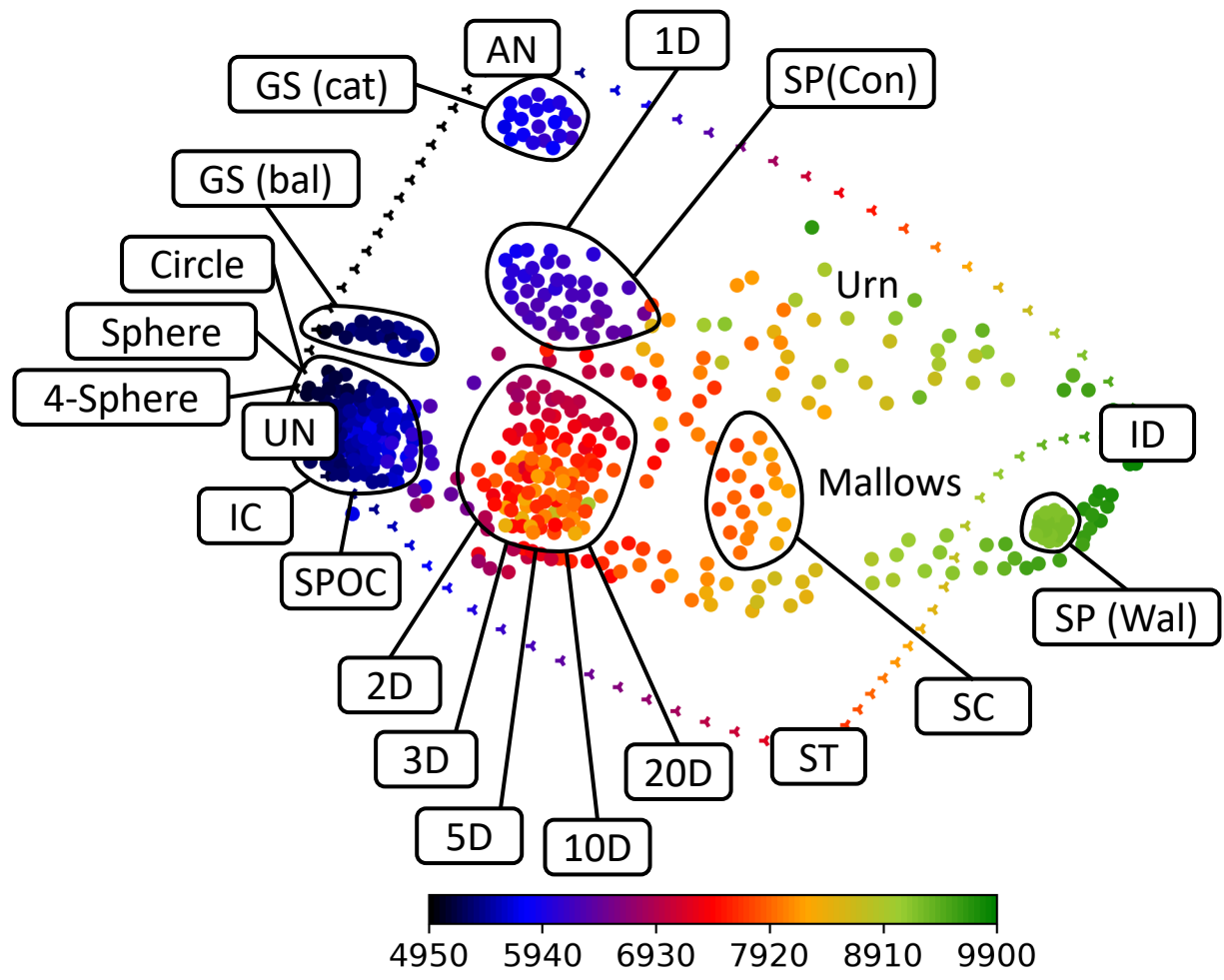
# Winner Score

Visualizing Experiment Results

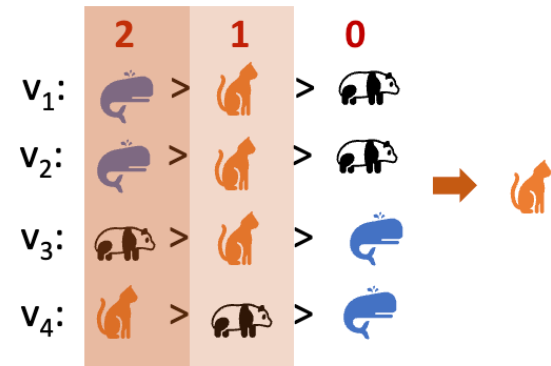


Highest Plurality Score

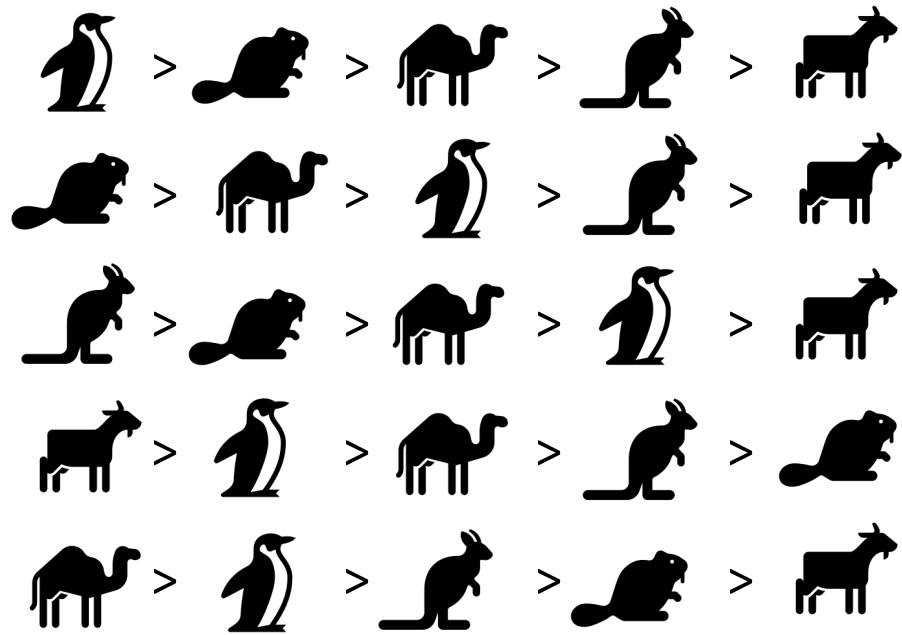




Highest Borda Score

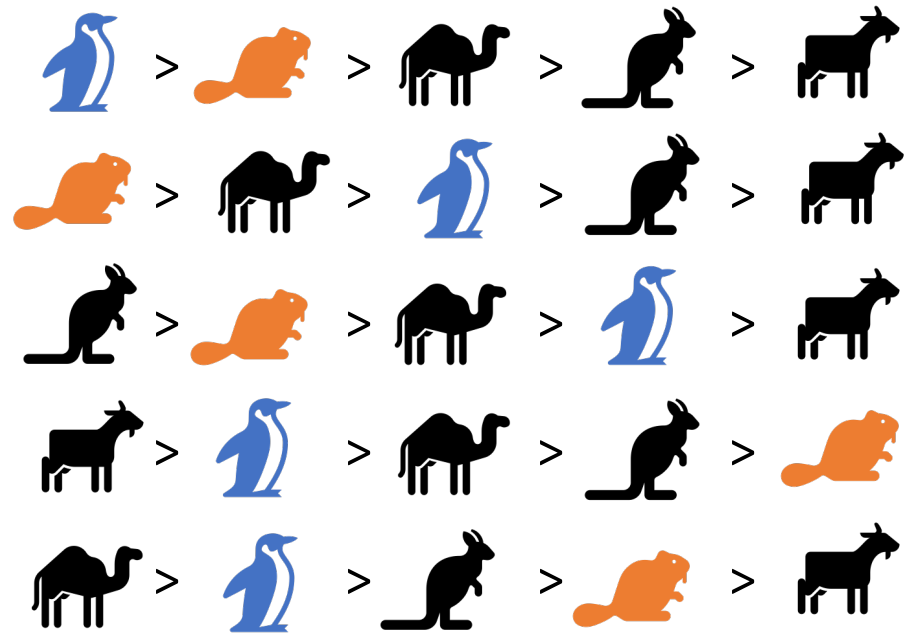


# Copeland Rule

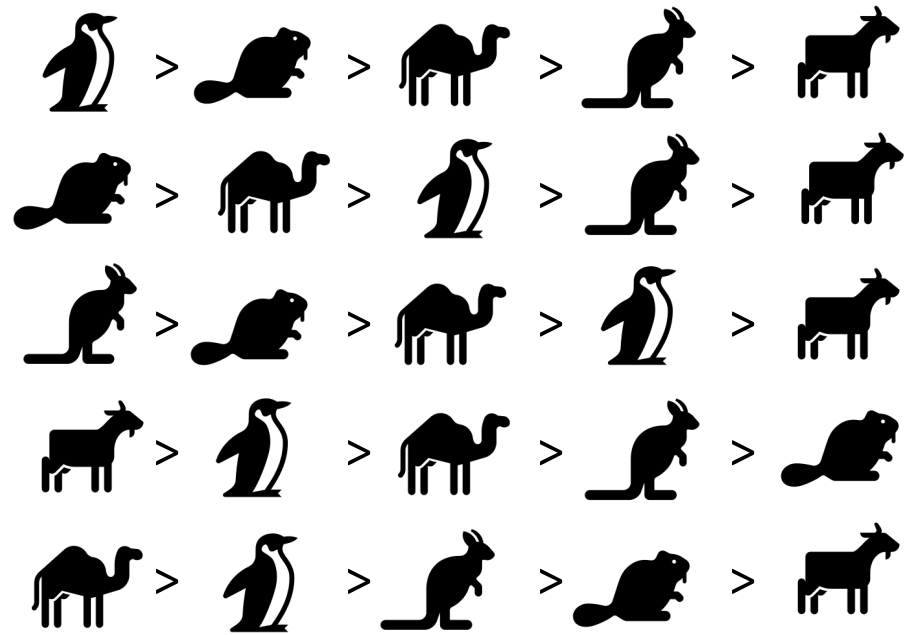




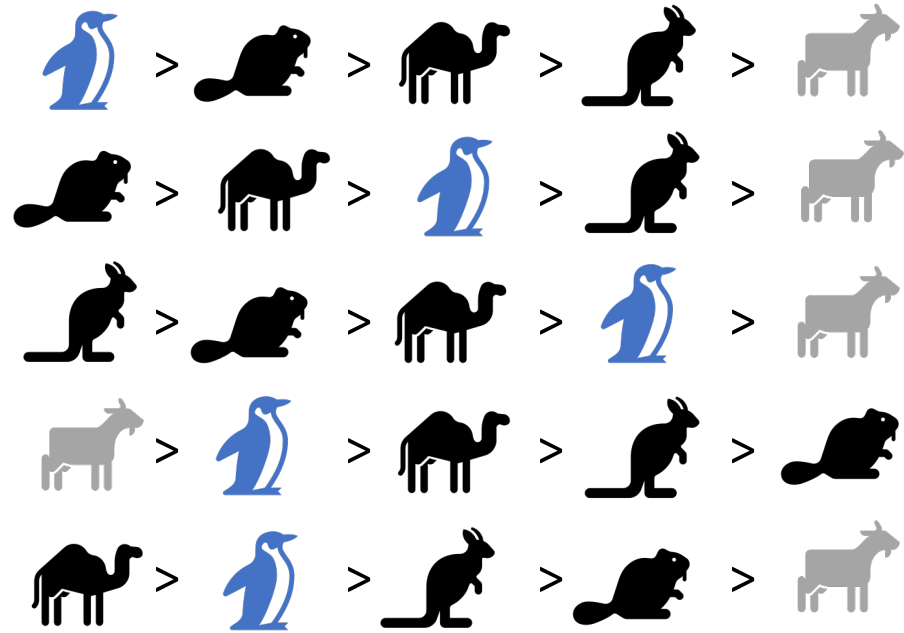
# Copeland Rule



# Copeland Rule



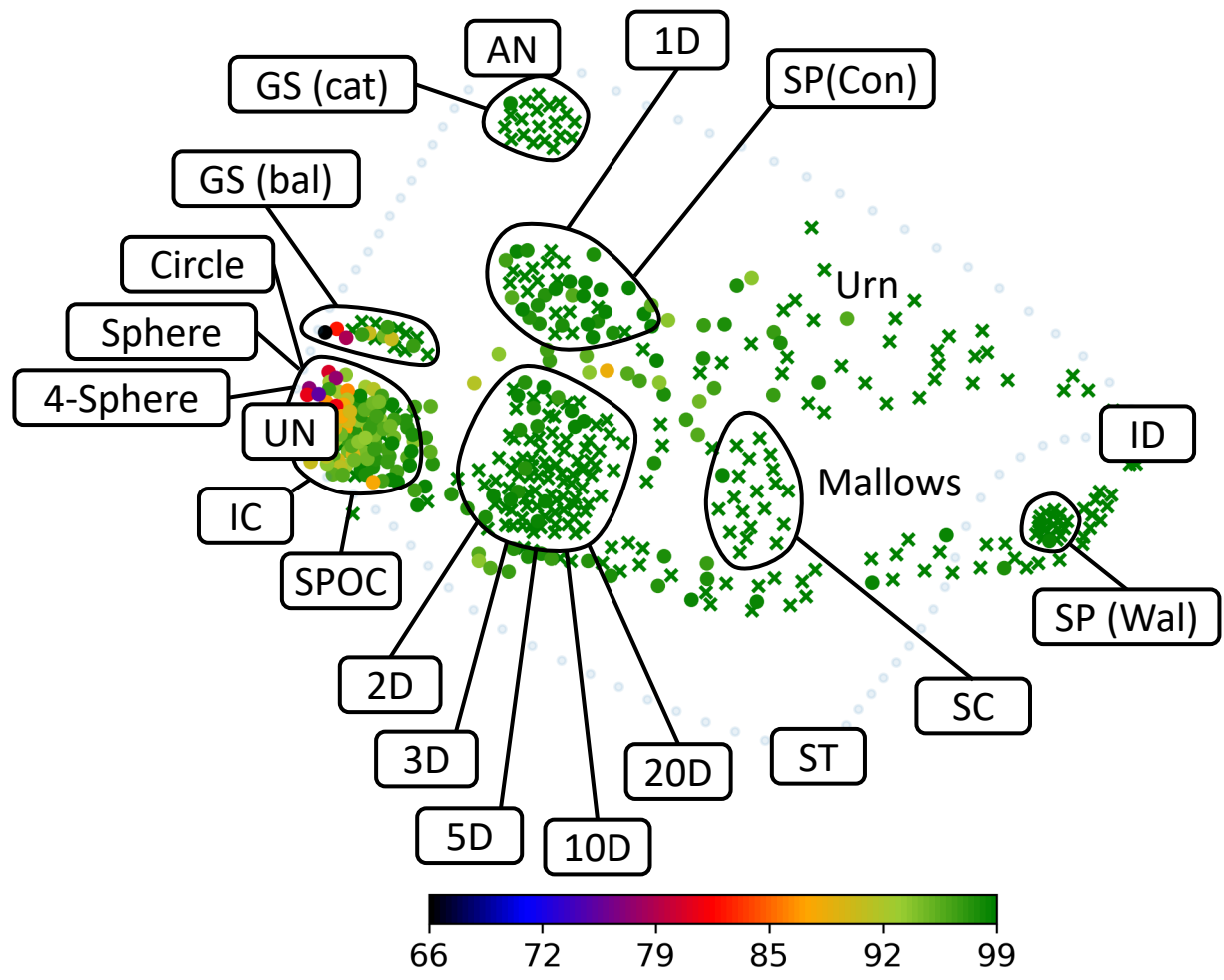
# Copeland Rule



**Condorcet winner**  
 A candidate that wins all pairwise comparisons

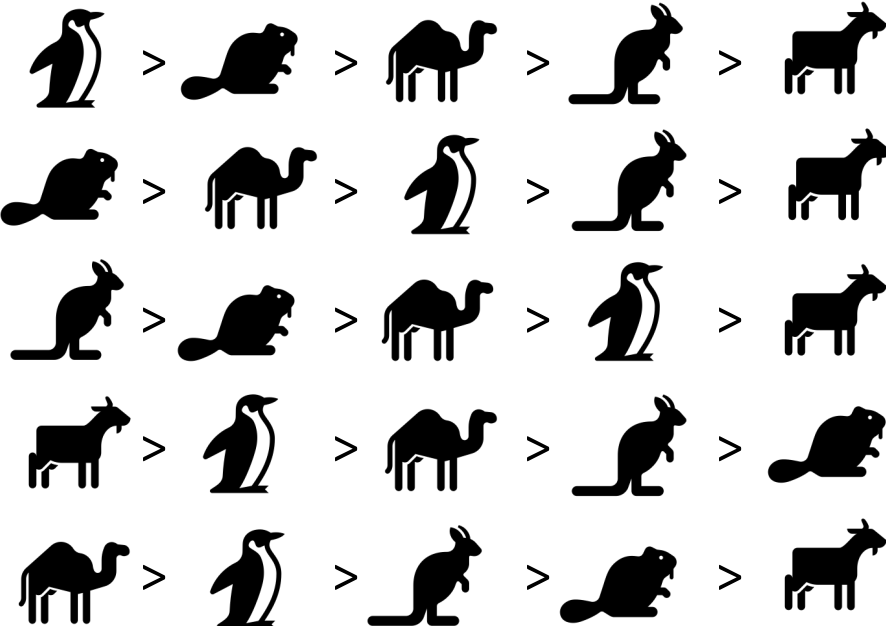


The candidate with the highest score wins



Highest Copeland Score

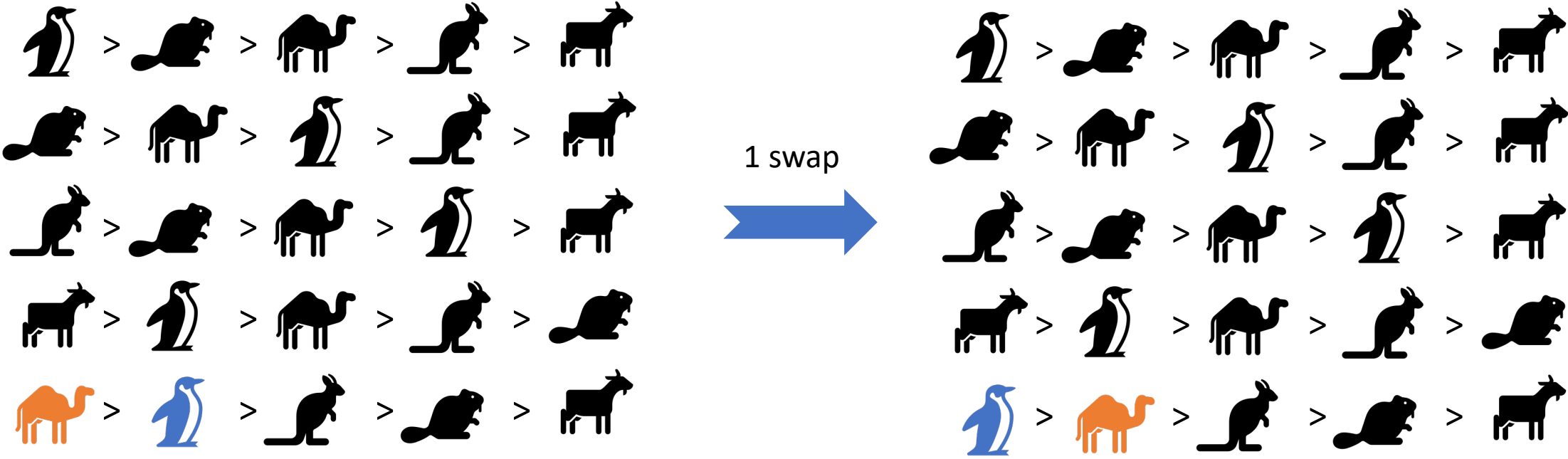
# Dodgson Rule



Score of a candidate is the minimal number of swaps needed to make him or her a Condorcet winner

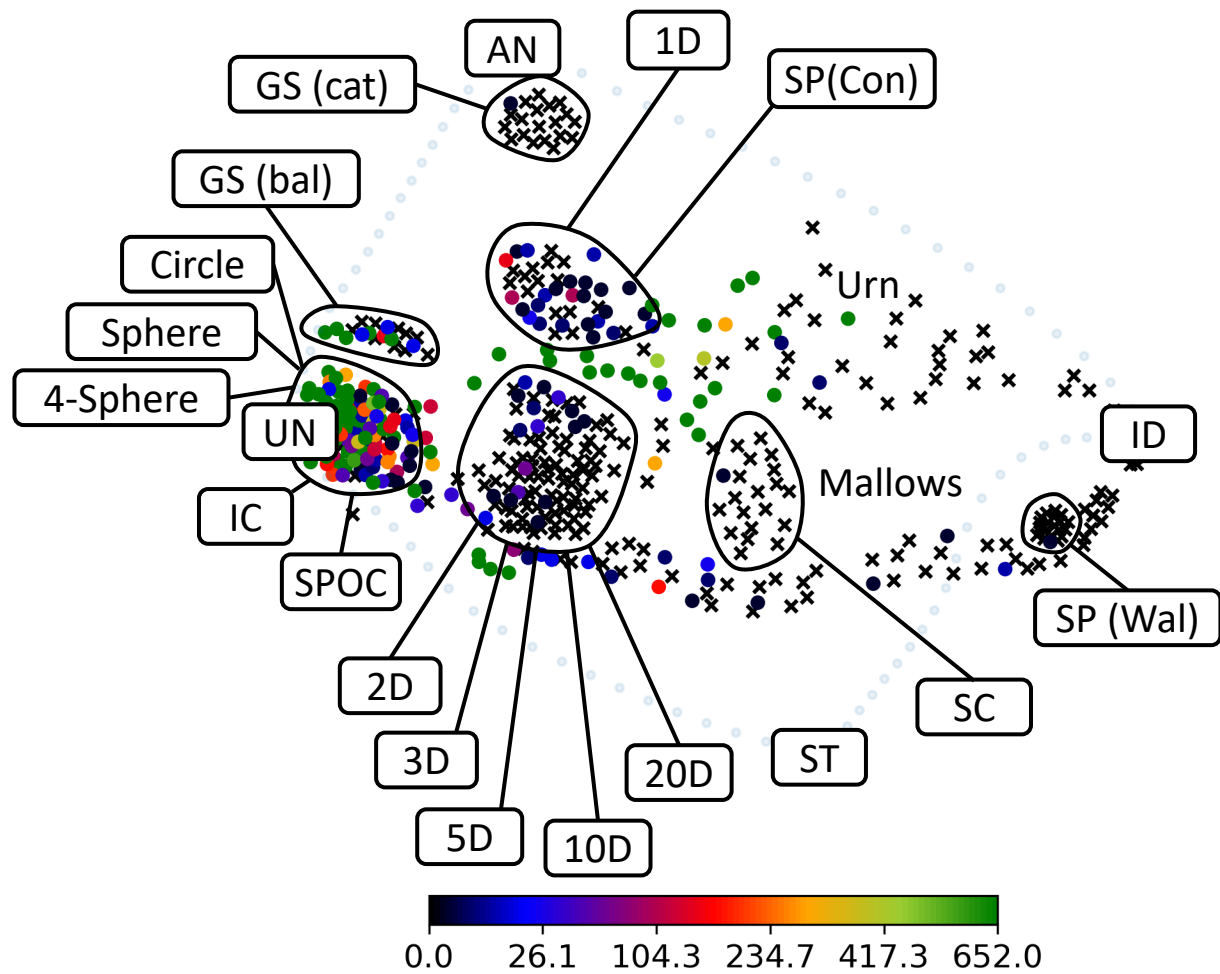
The candidate with the lowest score wins

# Dodgson Rule



Score of a candidate is the minimal number of swaps needed to make him or her a Condorcet winner

The candidate with the lowest score wins



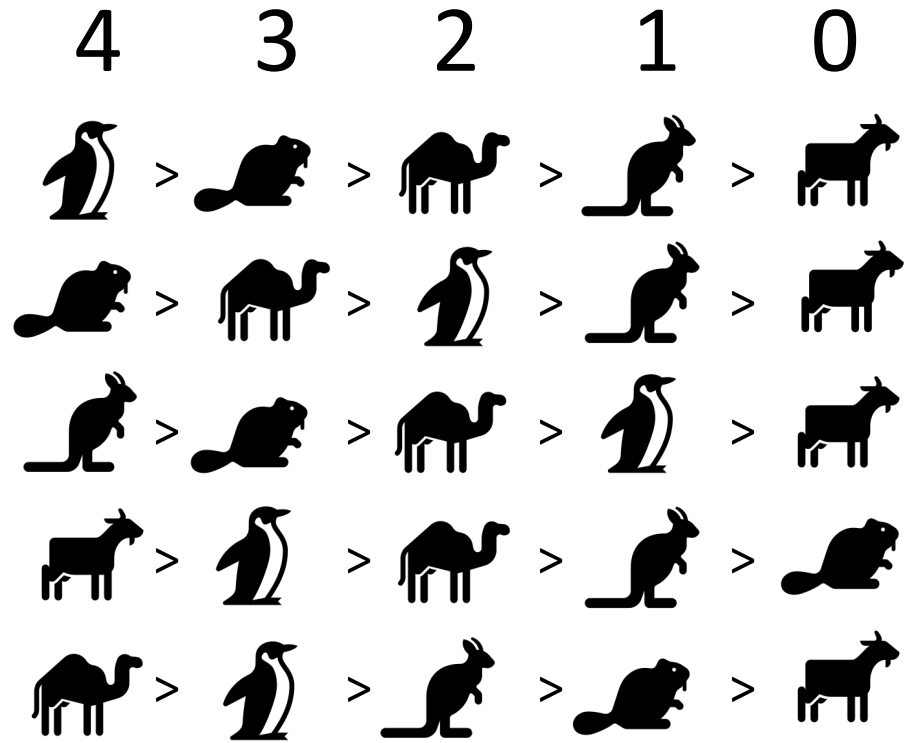
Lowest Dodgson Score

# Winning Committee Score

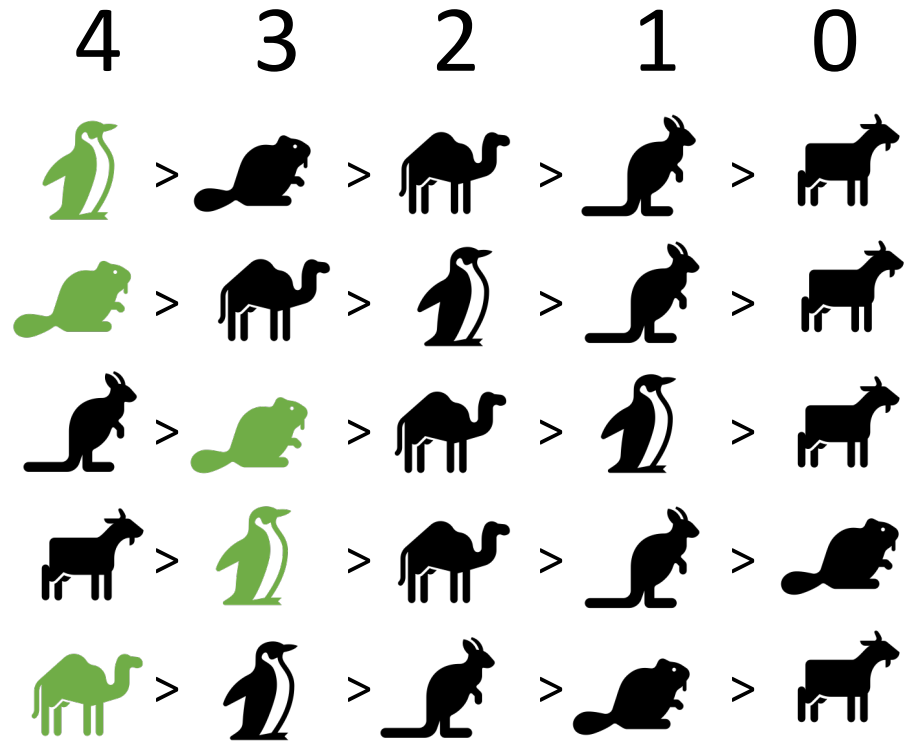
Visualizing Experiment Results



# Chamberlin—Courant (CC) Rule

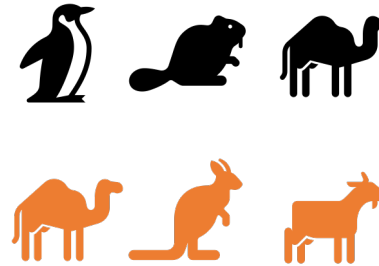
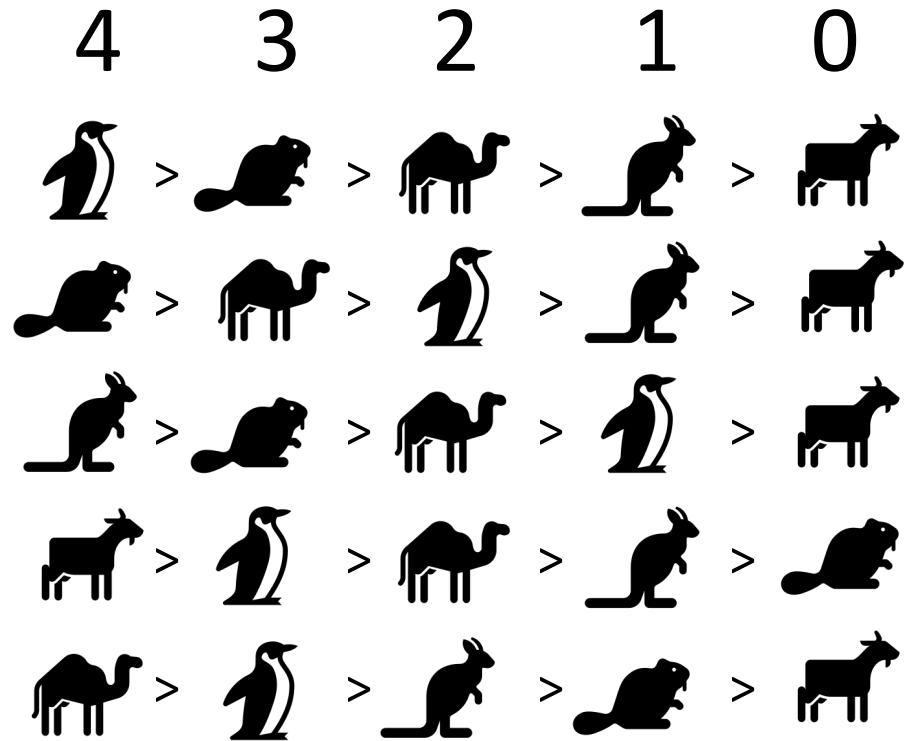


# Chamberlin—Courant (CC) Rule



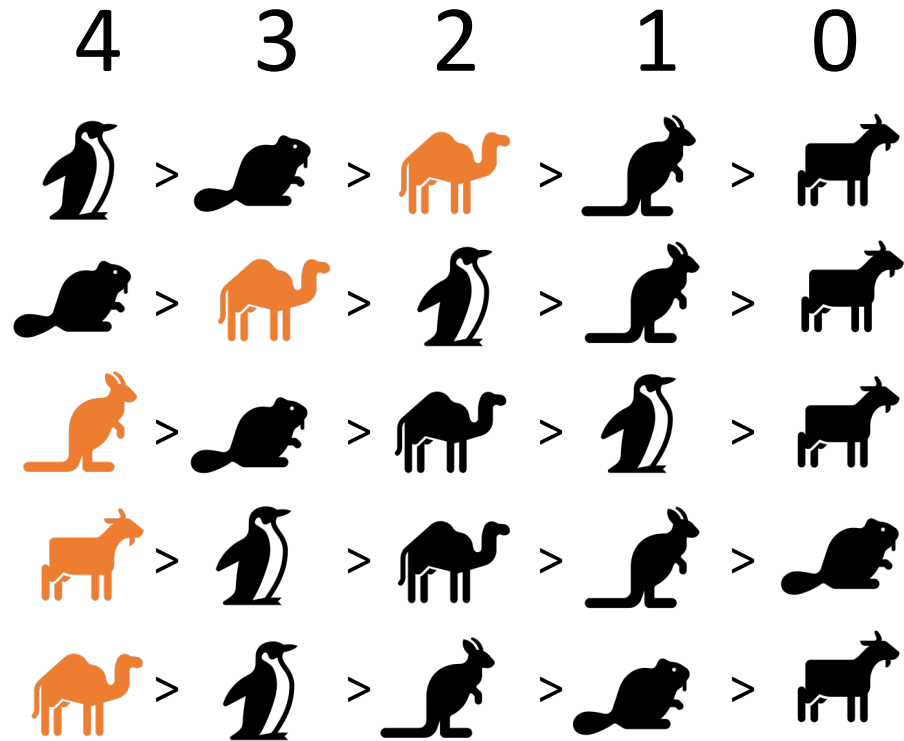
$$18 = 4+4+3+3+4$$

# Chamberlin—Courant (CC) Rule



$$18 = 4+4+3+3+4$$

# Chamberlin—Courant (CC) Rule

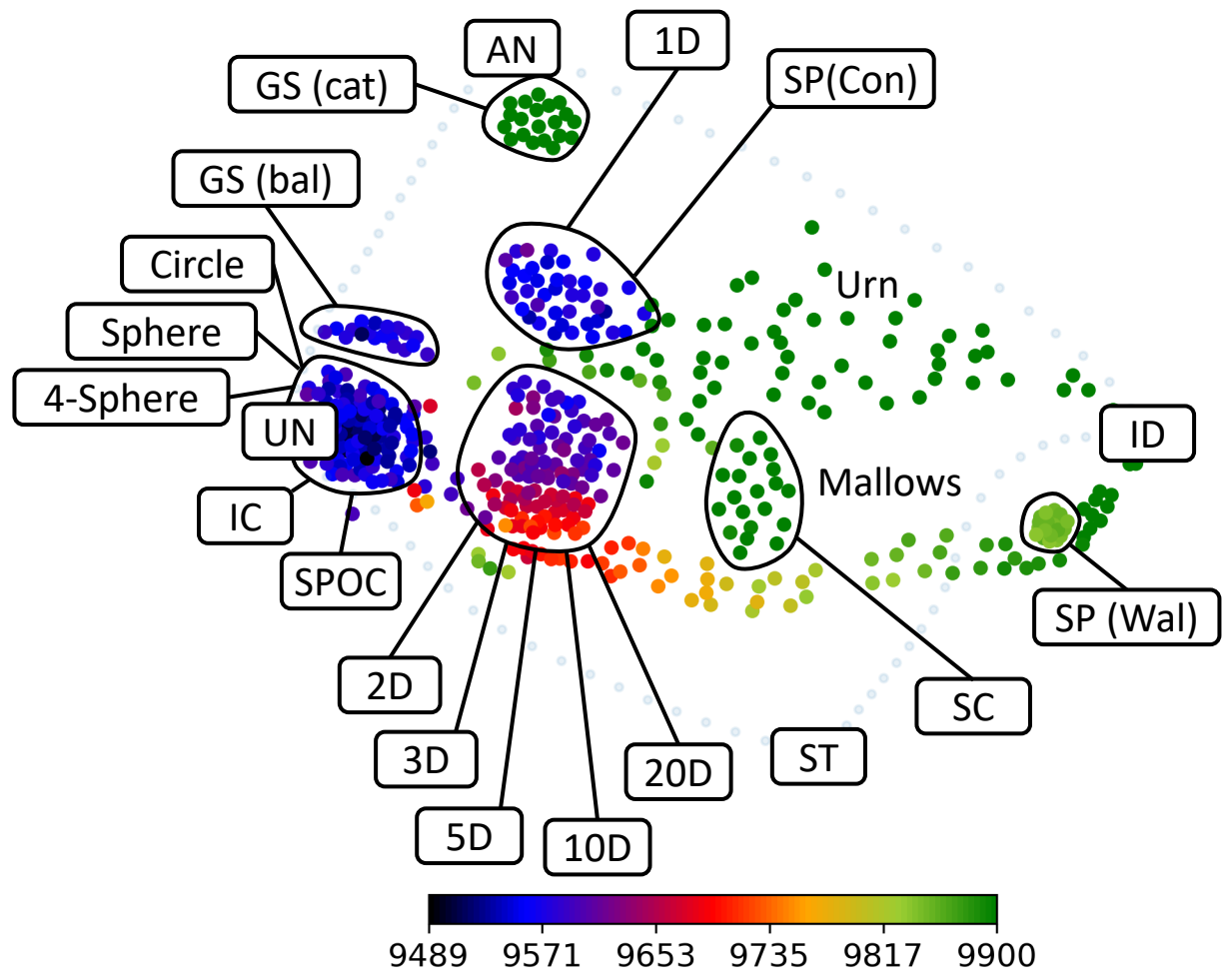


$$18 = 4+4+3+3+4$$



$$17 = 2+3+4+4+4$$

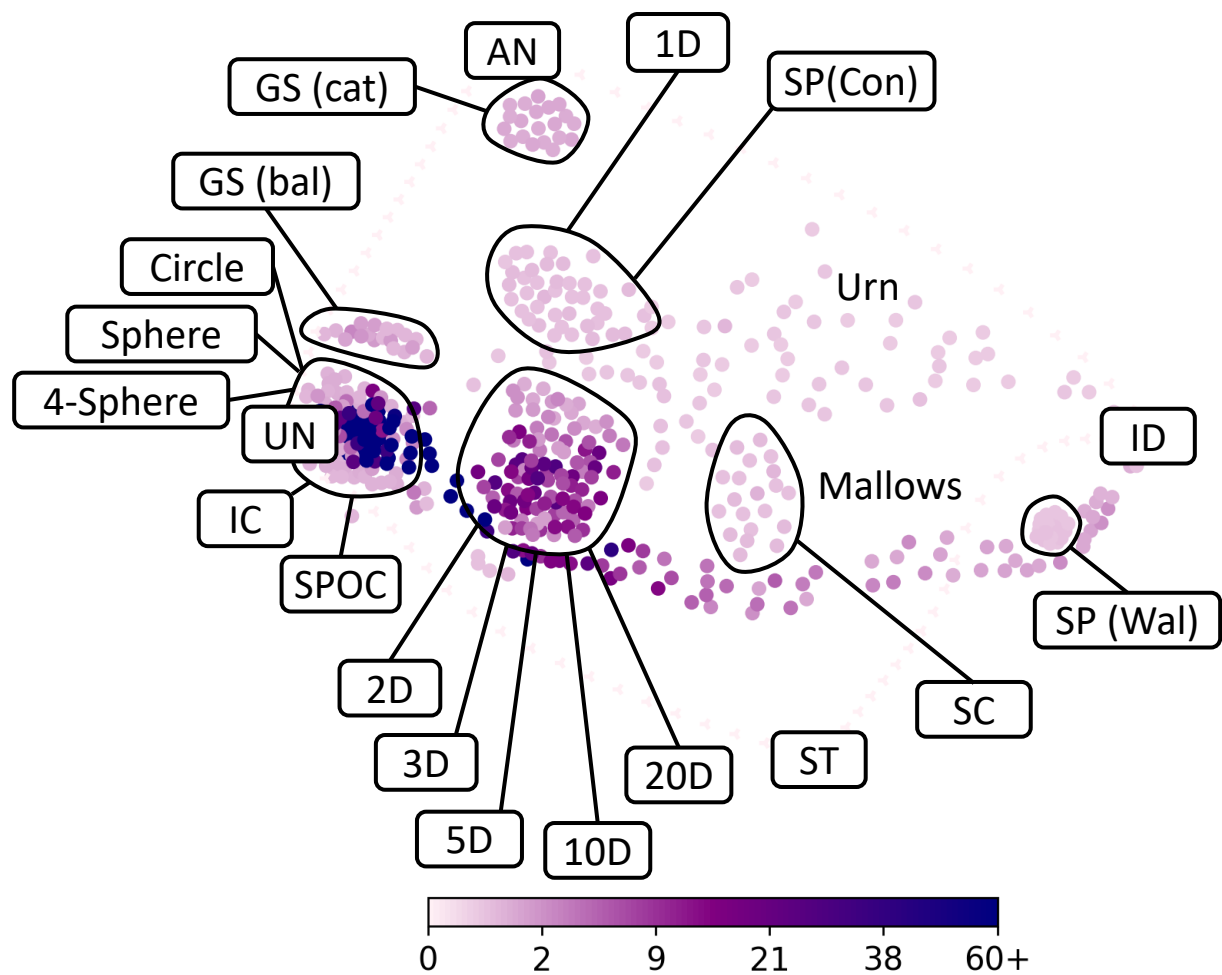
Committee with the highest score wins



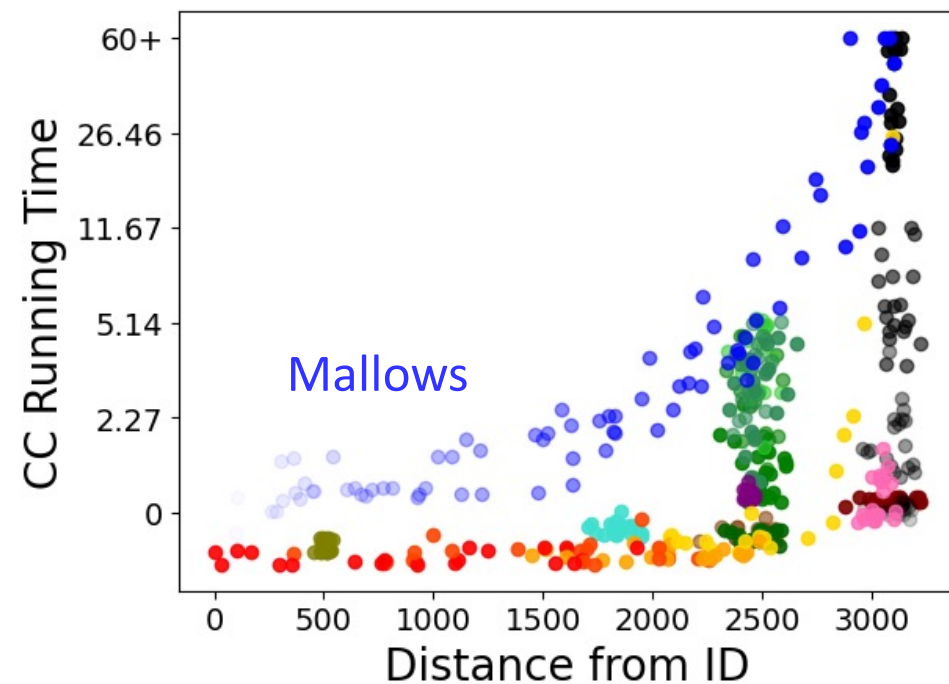
Highest CC Score

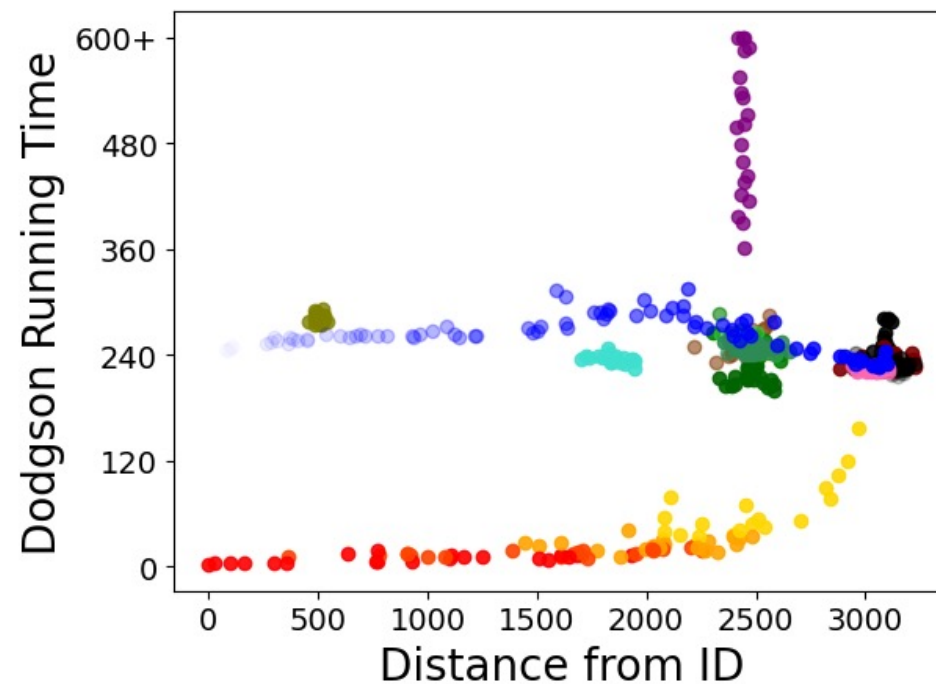
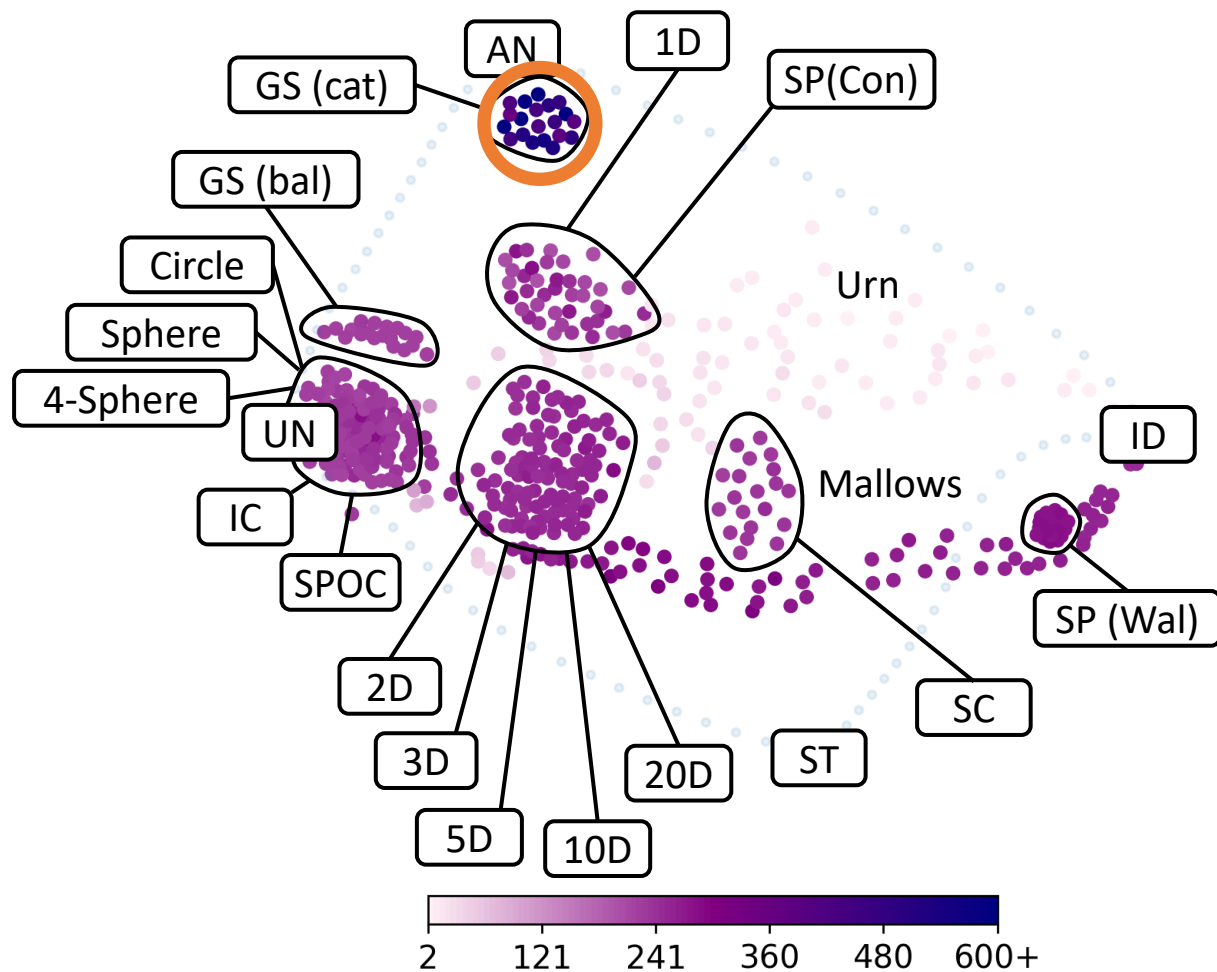
# Running Time

Visualizing Experiment Results



CC - Running Time (in seconds)





Dodgson - Running Time (in seconds)

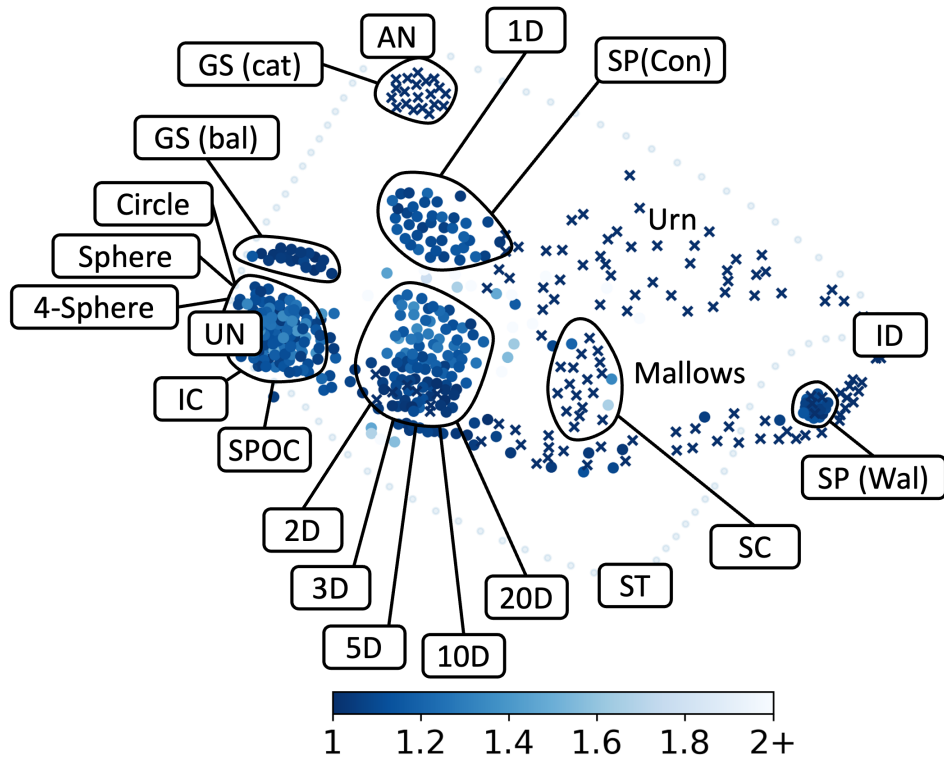


# Approximation Ratio

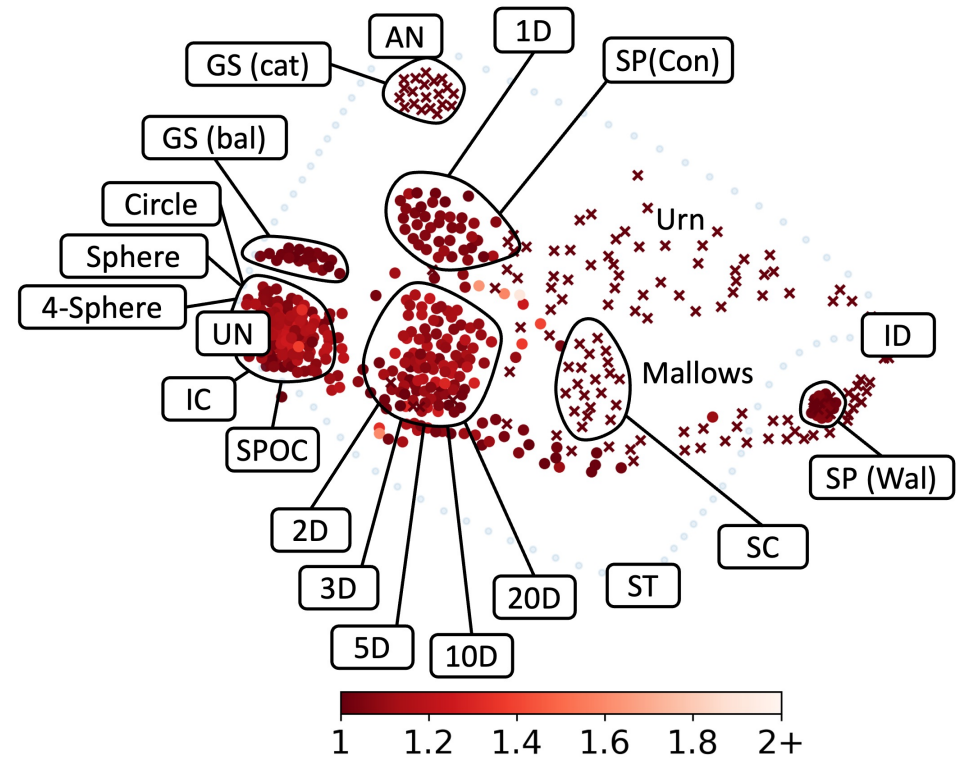
Visualizing Experiment Results

In each step, add the candidate who increases the committee's score the most

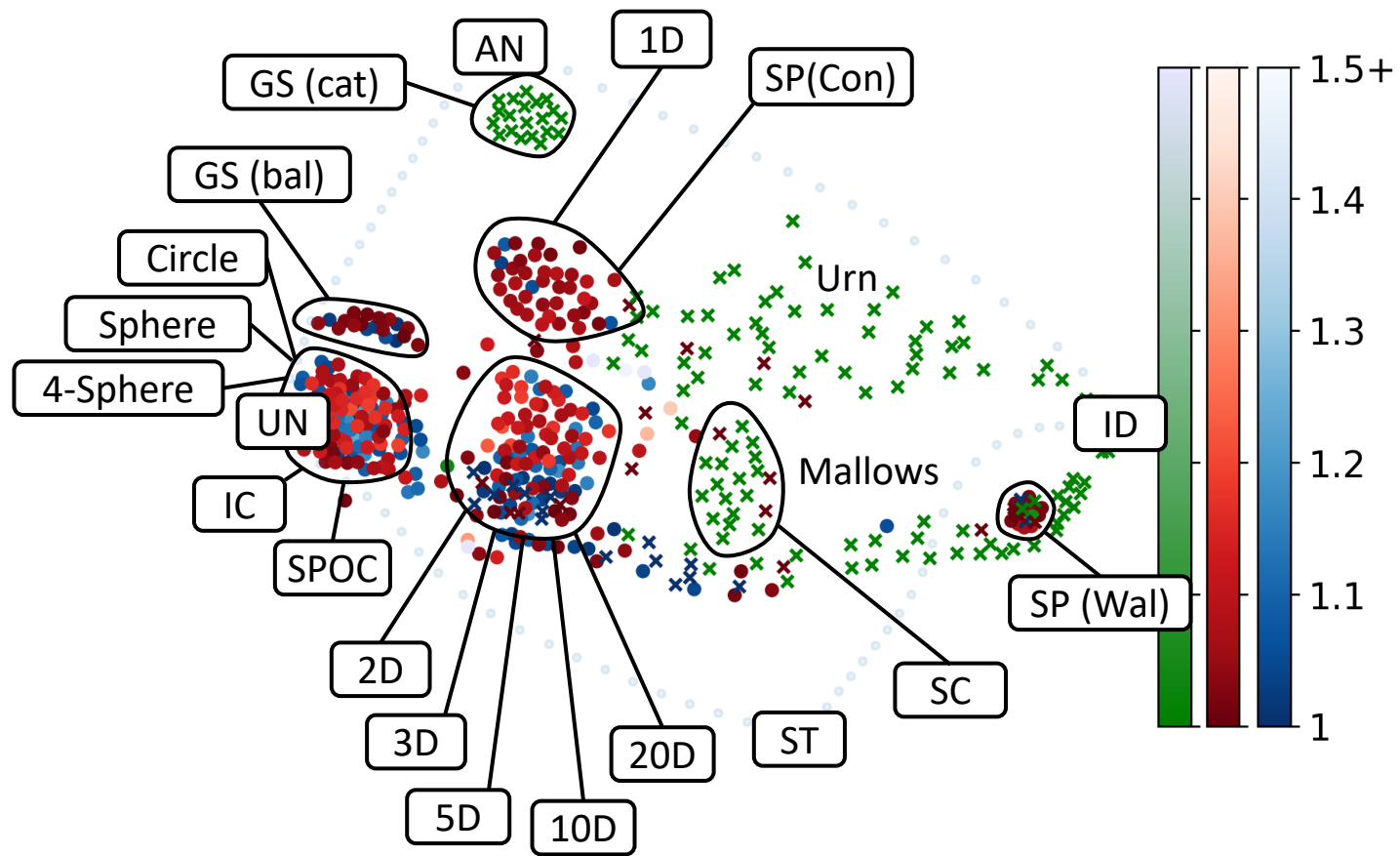
In each step, remove the candidate who decreases the committee's score the least



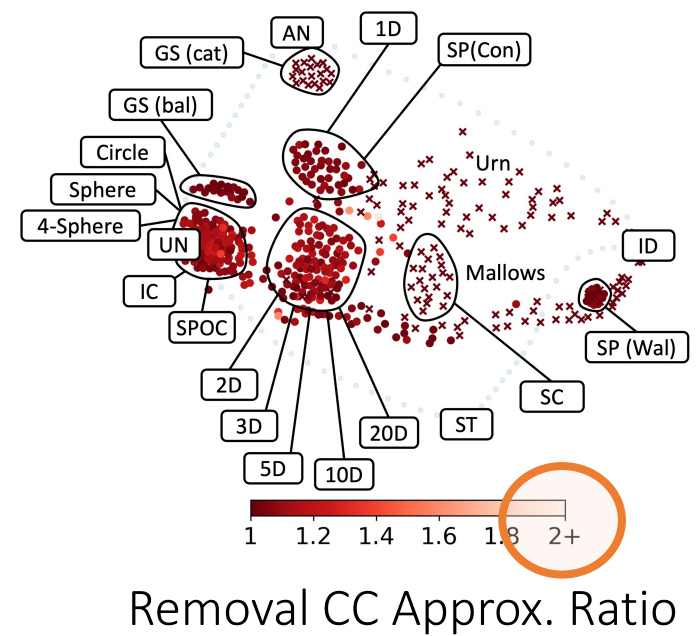
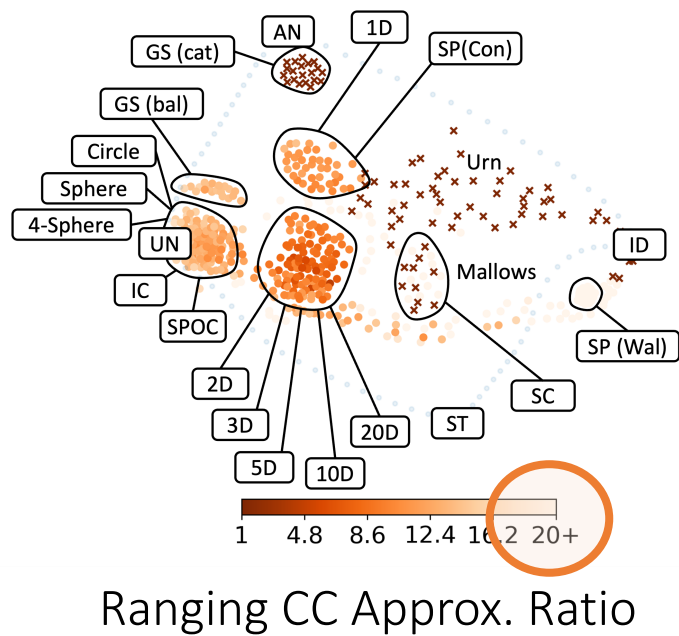
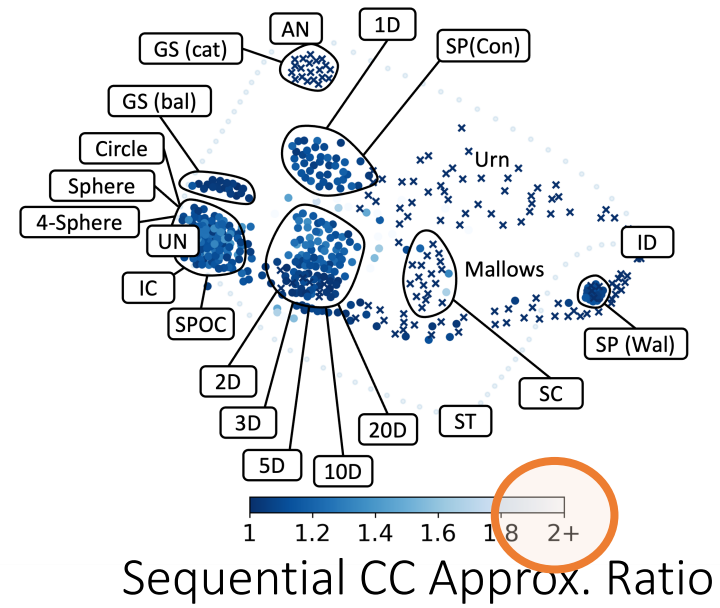
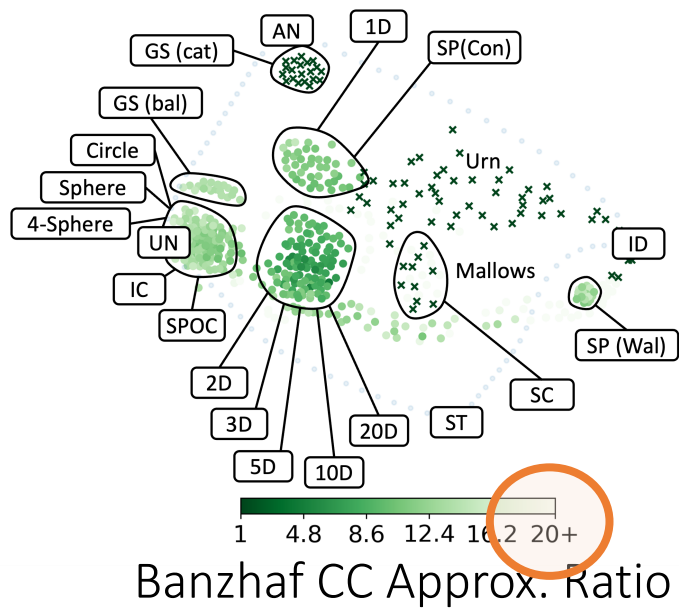
Sequential CC Approx. Ratio



Removal CC Approx. Ratio



Sequential CC vs Removal CC



 Collecting, Classifying, Analyzing, and Using Real-World Ranking Data, Boehmer and Schaar, AAMAS-23

 Putting a Compass on the Map of Elections, Boehmer et al., IJCAI-21

 PrefLib: A Library for Preferences <http://www.preflib.org>, Mattei and Walsh, ADT-13



10 minutes

# Putting Real-World Elections on the Map

# Preflib Data

PrefLib ID	Name	Type	#Elections	Avg. <i>m</i>	Avg. <i>n</i>	Avg. Inc
1	Irish	political	3	11.67	46003.67	0.39
2	Debian	survey	8	6.25	419	0.08
3	NASA	survey	1	32	10	0.1
4	Netflix	user ratings	200	3.5	818.79	0.0
5	Burlington	political	2	6	9384	0.27
6	Skate	survey	48	23.31	8.67	0.0
7	ERS	association	87	8.74	409.31	0.25
8	Glasgow	political	21	9.9	8970.29	0.5
9	AGH	survey	2	8	149.5	0.0
10	Ski	sport	2	260.5	4	0.23
11	Web	meta-search	77	1874.74	4.04	0.36
12	T-Shirt	survey	1	11	30	0.0
14	Sushi	survey	1	10	5000	0.0
15	Clean Web	meta-search	79	78.15	4.04	0.0
16	Aspen	political	2	8	2502	0.26
17	Berkley	political	1	4	4173	0.13
18	Minneapolis	political	4	218	34370.5	0.76
19	Oakland	political	7	7	52449.29	0.39

637 elections from 35 datasets:

- Humans expressing opinions concerning candidates for a position (political, association)
- Humans expressing preferences over objects (survey, user ratings)
- Humans ranking items in a test (human tests)

20	Pierce	political	4	5	188627	0.29
21	San Francisco (sf)	political	14	10.43	61635.79	0.51
22	San Leandro (sl)	political	3	5.33	23666	0.27
23	Takoma Park	political	1	4	204	0.13
24	Mechanical Turk dots	human tests	4	4	795.75	0.0
25	Mechanical Turk puzzle	human tests	4	4	795	0.0
26	French Presidential	political	6	16	430.83	0.68
27	Proto French	political	1	15	398	0.7
28	APA	association	12	5	16991.33	0.16
30	UK Labor Leadership	political	1	5	266	0.21
31	Vermont	political	15	3.93	1160.73	0.42
32	Education Survey	survey	7	13.57	21.86	0.39
33	San Sebastian Poster	survey	2	17	61.5	0.59
34	Cities	survey	2	42	392	0.73
35	Breakfast Items	survey	6	15	42	0.0
57	Austrian Parliamentary	political	9	12.22	4792773.11	0.84

Decisive features:

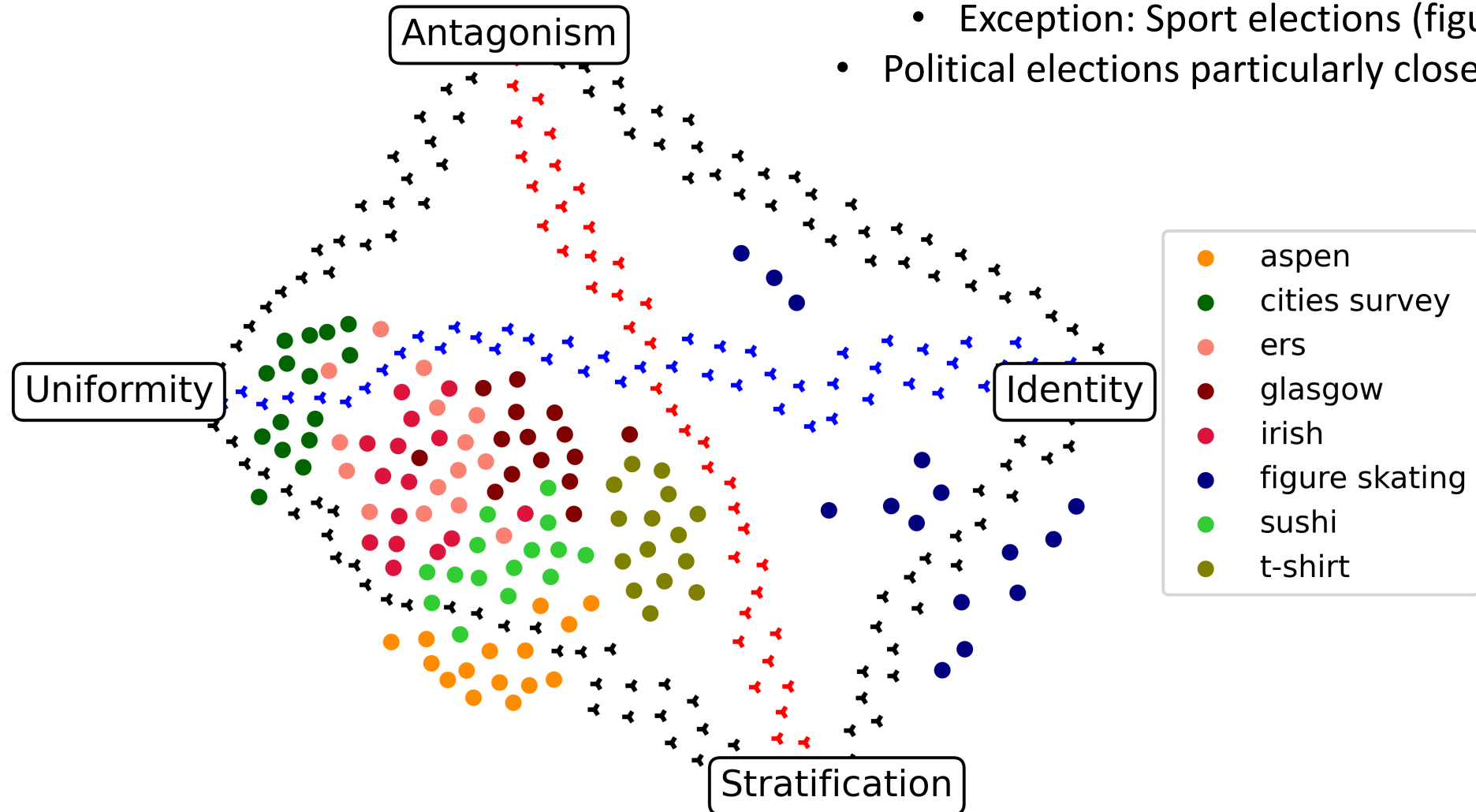
- Typically below 10 candidates or below 10 voters.
- Often highly incomplete votes, voters typically rank only small subset of candidates.

Usable for map:

Irish, Skate, ERS, Glasgow, T-Shirt, Sushi, Aspen, and Cities

# Map of Preflib Elections

- Most elections fall in bottom left
  - Exception: Sport elections (figure skating)
- Political elections particularly close to each other



# A Second Datasource

Collecting, Classifying, Analyzing, and Using Real-World Ranking Data, AAMAS 2023.

## Time-Based Elections

- Multi-race competitions (Formula 1 season/Tour de France)
- Top-x rankings at different times (Spotify, boxing, tennis top 100, american football)

...

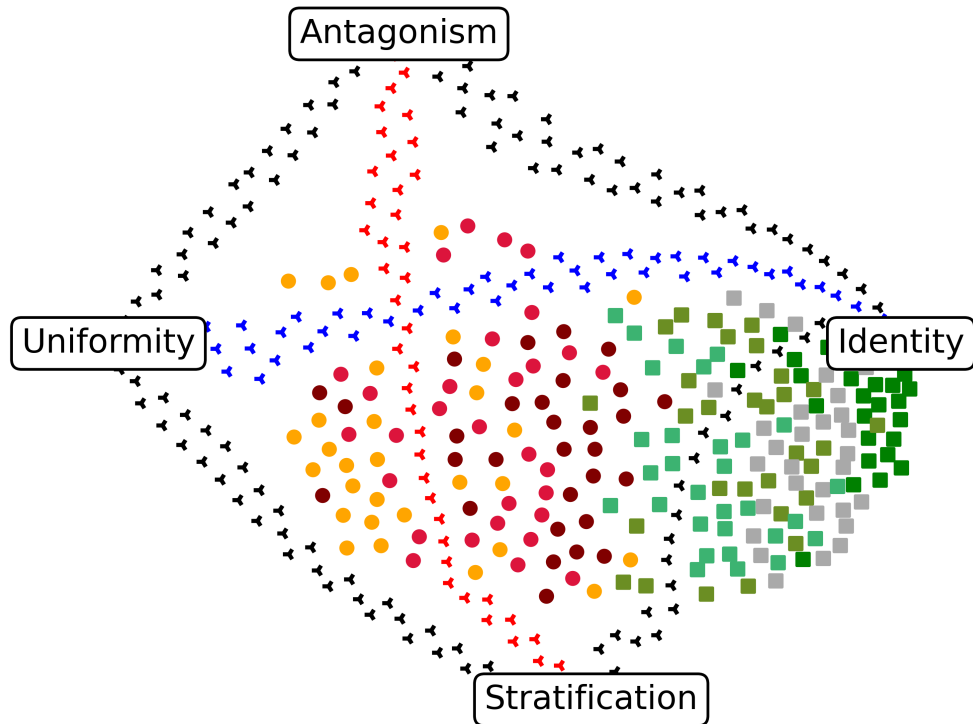
## Criterion-Based Elections

- Indicator-based rankings (cities, countries, universities)
- Top-x rankings from different sources (Spotify, american football)

...

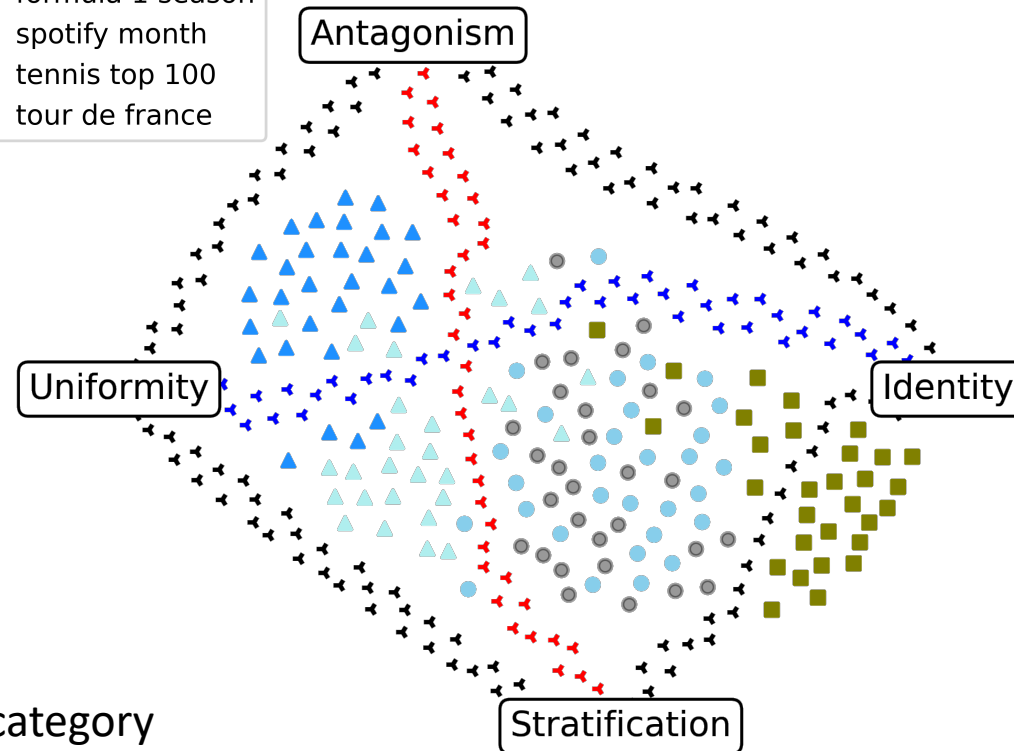


# Map of Real-World Elections

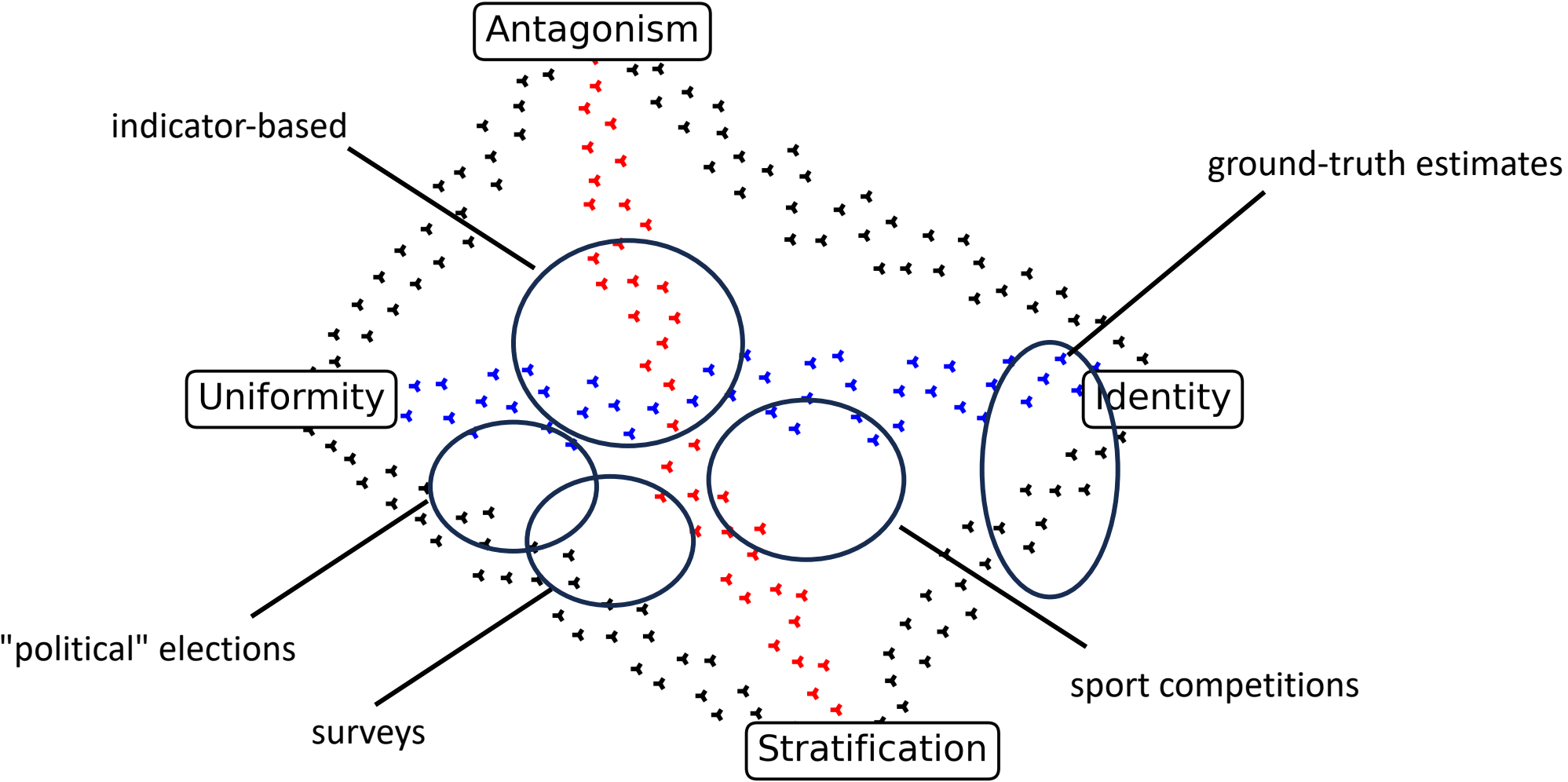


- Close to Identity elections (mostly time-based)
- Elections from "the middle"
- ▲ "Outliers" closer to Uniformity

→ very consistent behavior of data sources from one category

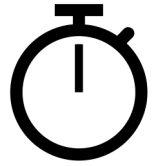


# Different Types of Real-World Elections



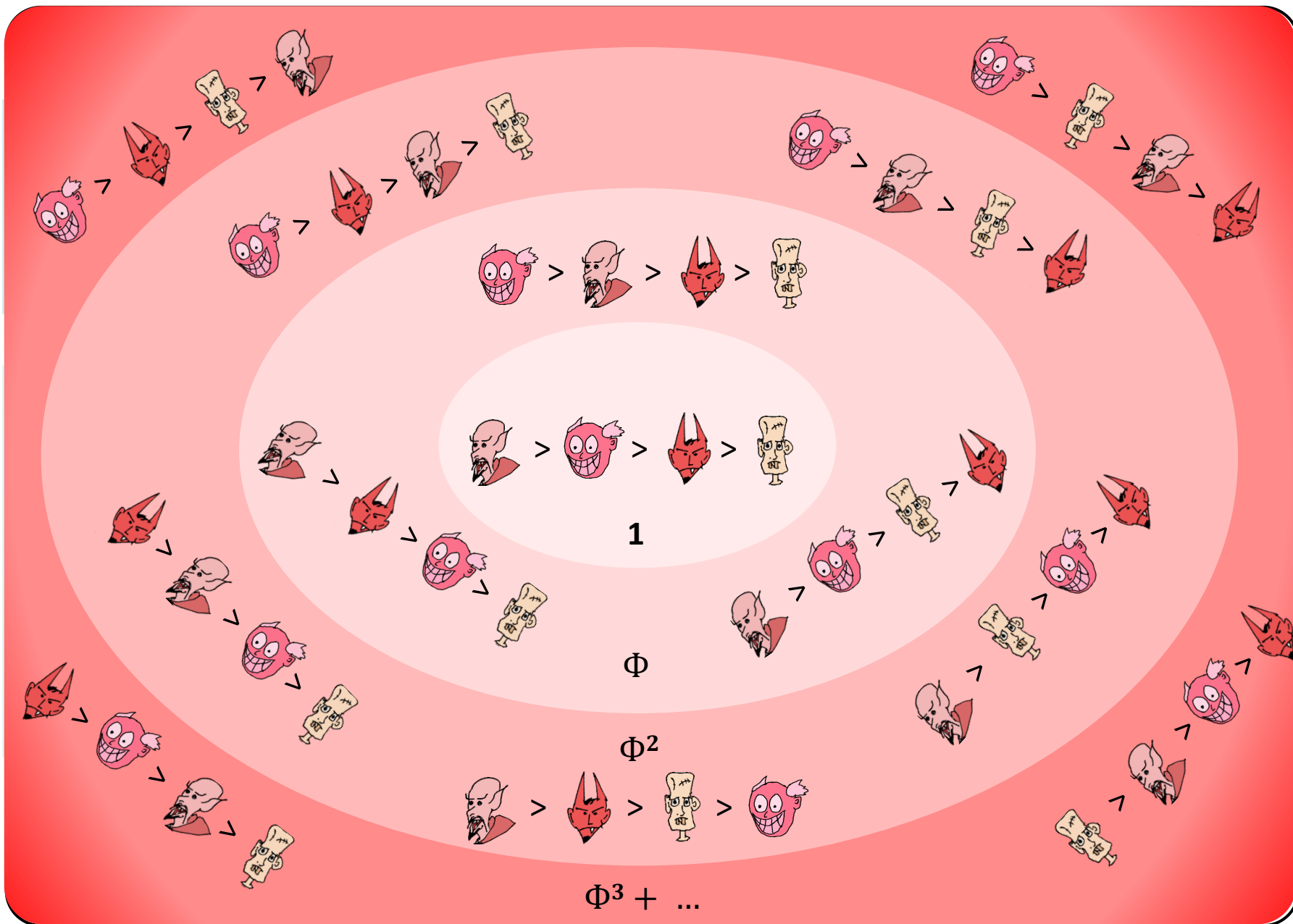
 Putting a Compass on the Map of Elections, Boehmer et al., IJCAI-21

 Properties of the Mallows Model Depending on the Number of Alternatives: A Warning for an Experimentalist, Boehmer et al., ICMI-23



15 minutes

# Using the Map to Generate Realistic Data: (Normalized) Mallows Model



## Mallows Model

*Input*

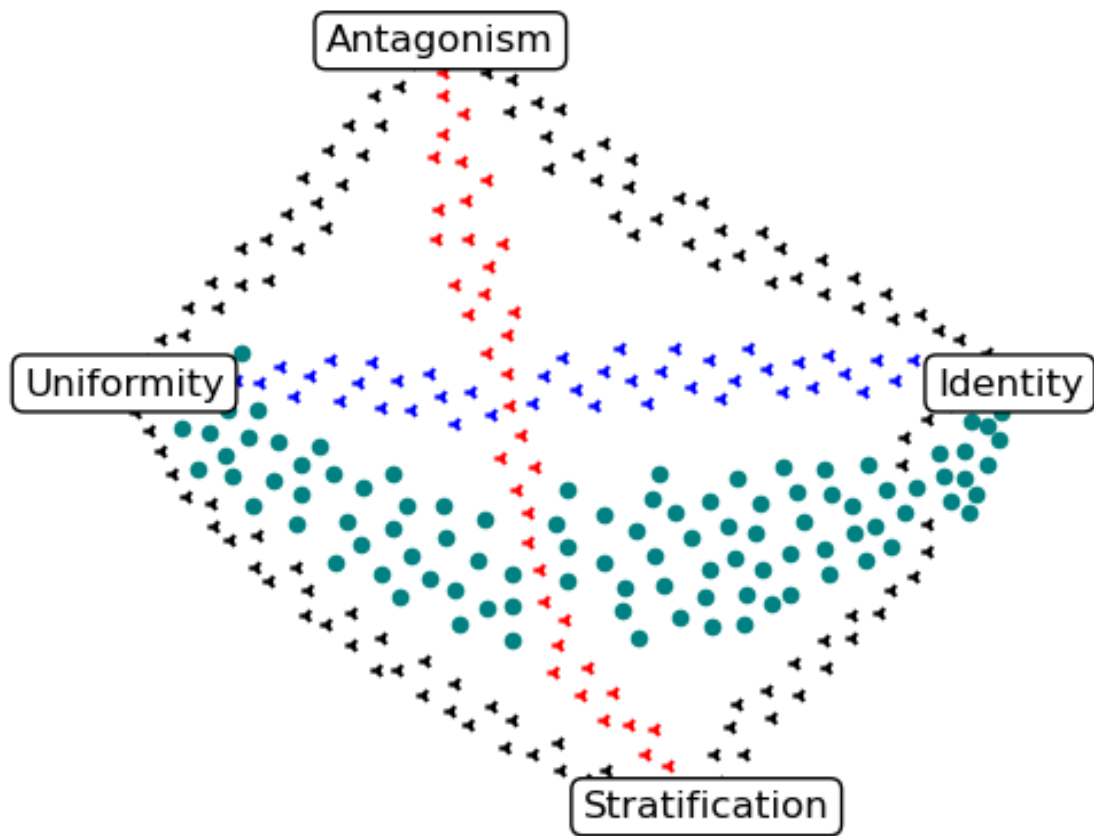
Central vote  $v^*$  + dispersion  
parameter  $\varphi$

*Sampling*

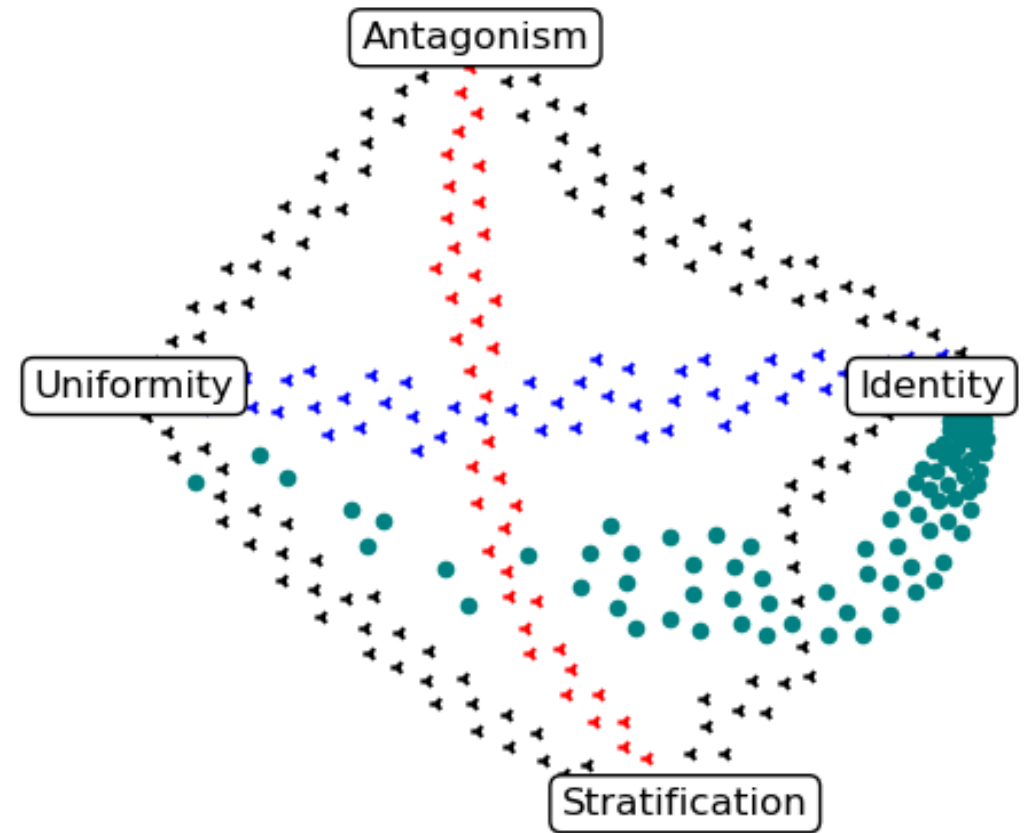
Probability of sampling vote  $v$   
proportional to:

$$\varphi^{\text{swap}(v, v^*)}$$

# Mallows Model with Uniformly Sampled $\varphi$

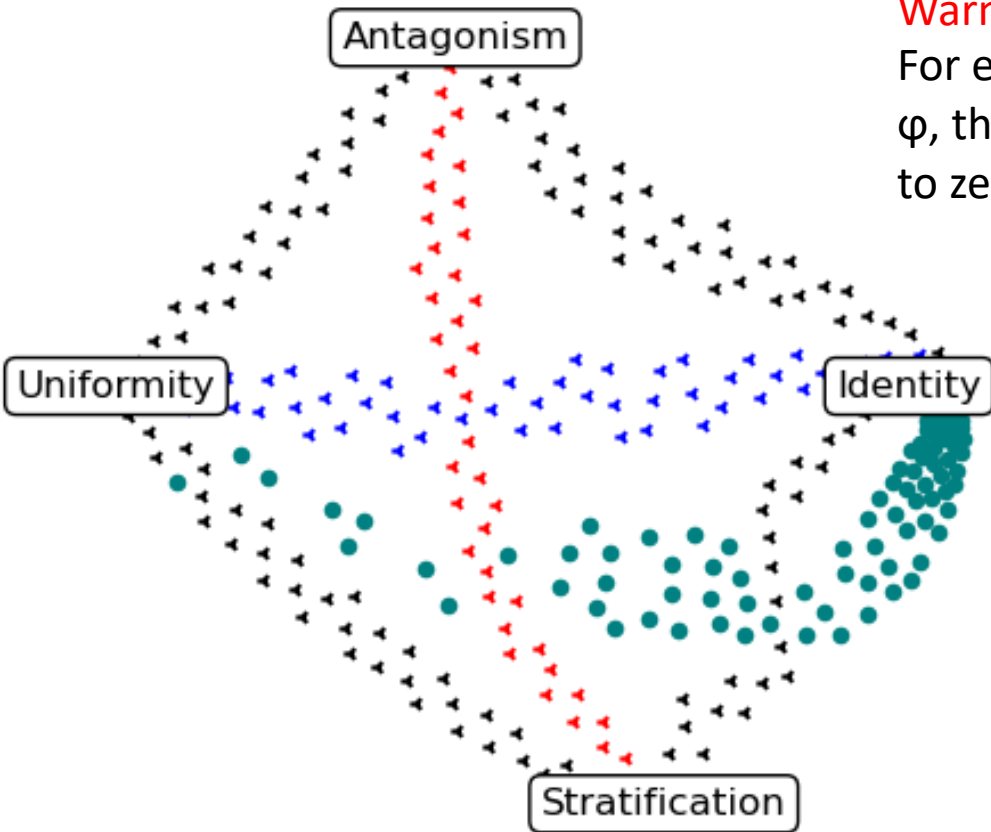


100 voters and 10 candidates



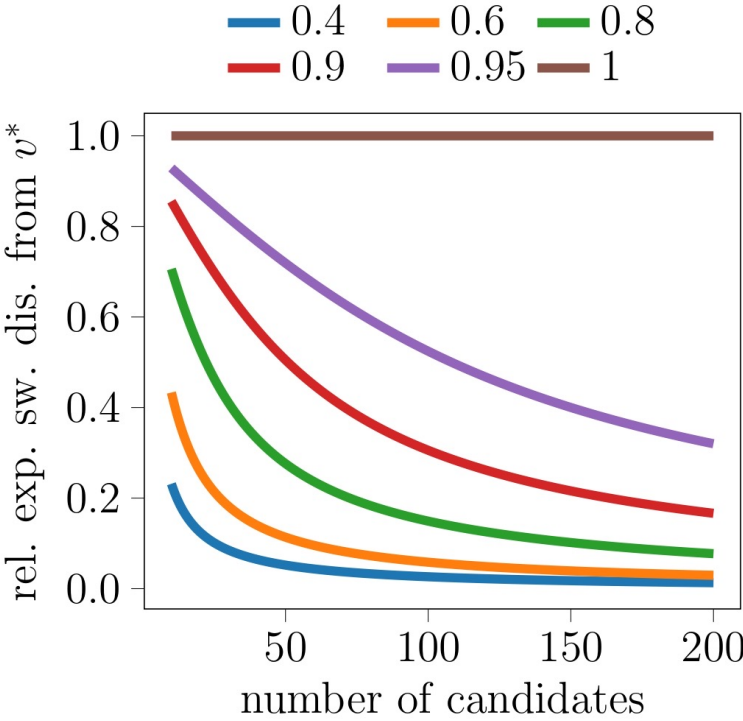
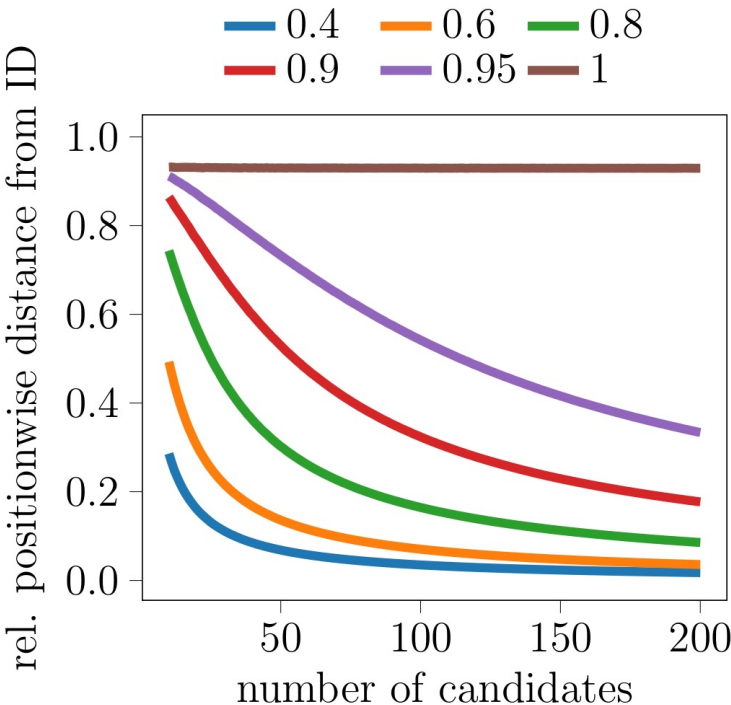
100 voters and 50 candidates

# Mallows Model with Uniformly Sampled $\varphi$



100 voters and 50 candidates

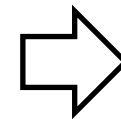
**Warning**  
For each fixed dispersion parameter  $\varphi$ , the relative distance to ID goes to zero.



# Problems with Mallows Model

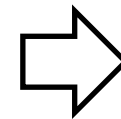
## Common Implicit Assumptions

A fixed dispersion parameter produces "structurally similar" elections for different candidate numbers.

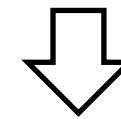


Fixed dispersion parameter for different candidate numbers in one experiment.

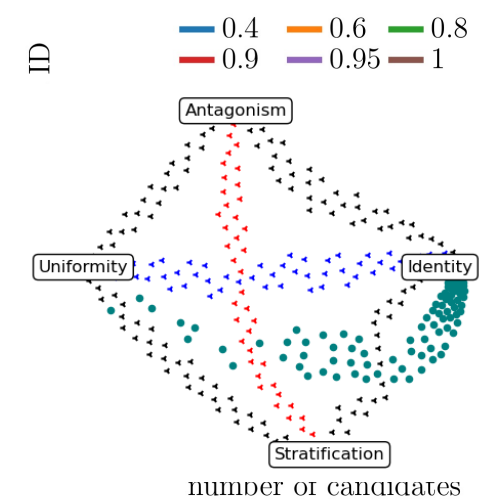
A uniformly at random chosen dispersion parameter "uniformly covers" the space between identity and uniformity.



Don't know what dispersion to use? Just choose uniformly at random, it's the *natural* agnostic choice.



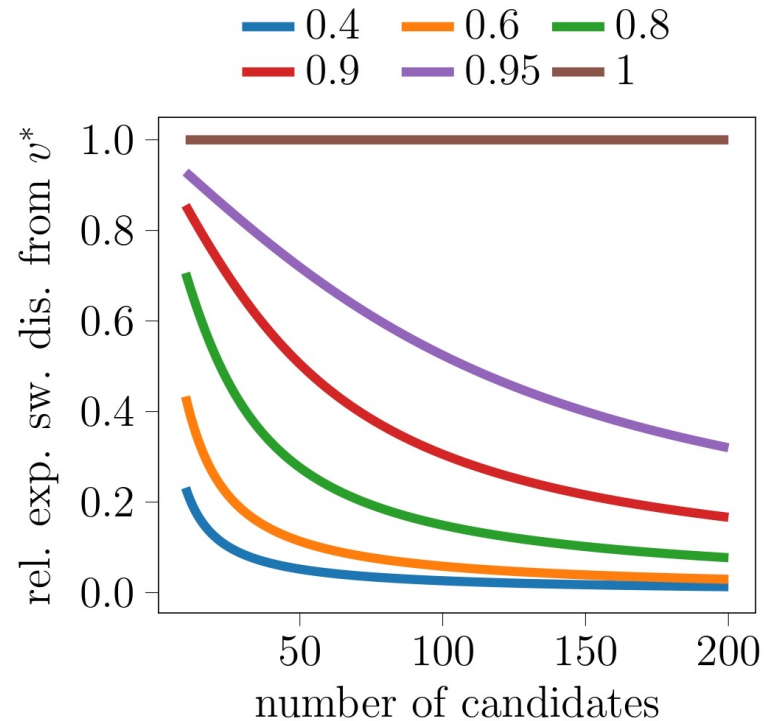
**Possibility for methodological errors!**



# What Can We Do?

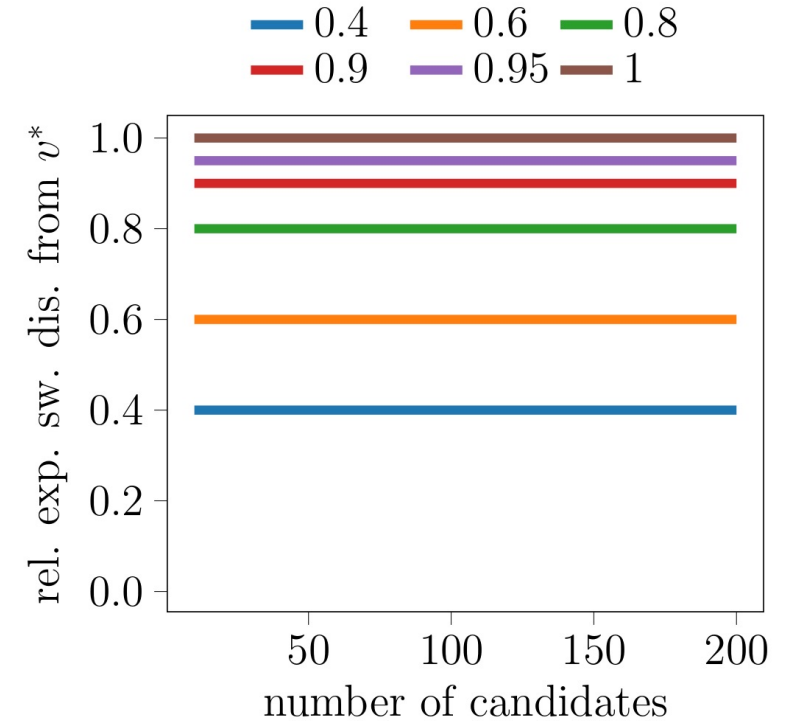
## Mallows Model

Sampled votes become more and more similar to central one



## Normalized Mallows Model

Keep expected swap distance from central order fixed





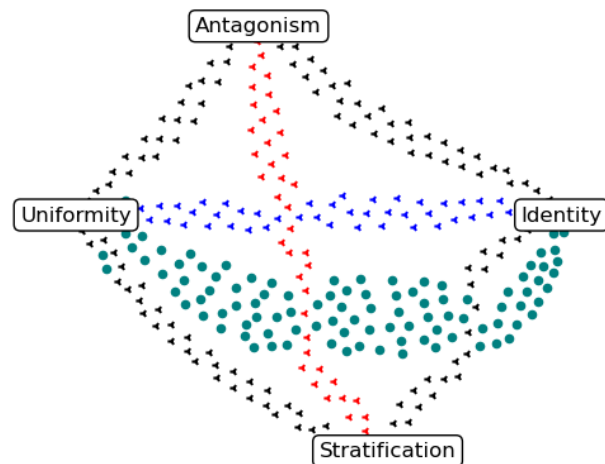
# Normalized Mallows Model

## Idea

- Keep expected swap distance from central order fixed

## Advantage

- Uniform parameter values lead to uniform coverage of election space
- "Consistent" behavior for varying number of candidates
- Easy-to-interpret parameter values



## Input

Central vote  $v^*$  with  $m$  candidates + "new" parameter  $\text{norm-}\varphi$

## Conversion

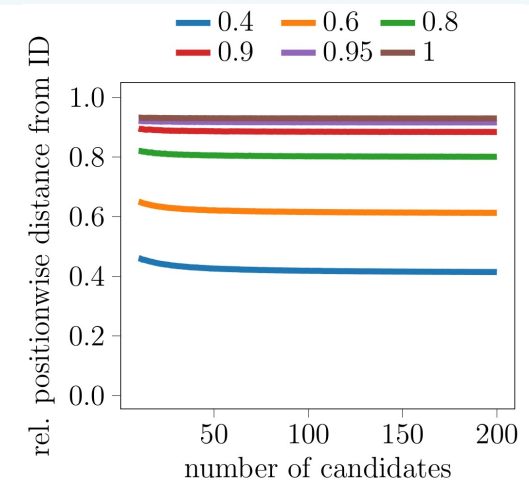
Choose a value  $\varphi$  of the dispersion parameter s.t. expected swap distance between central and sampled vote:

$$\text{norm-}\varphi \cdot \frac{1}{4} m(m-1)$$

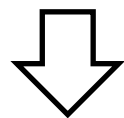
## Sampling

Probability of sampling vote  $v$  proportional to:

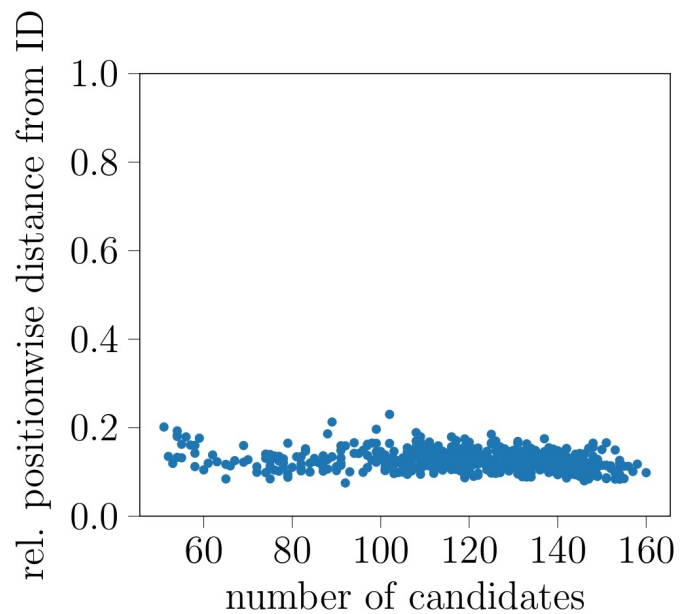
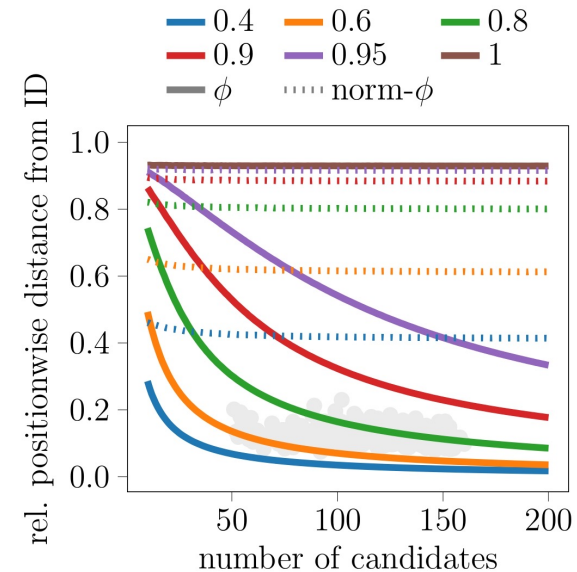
$$\varphi^{\text{swap}(v,v^*)}$$



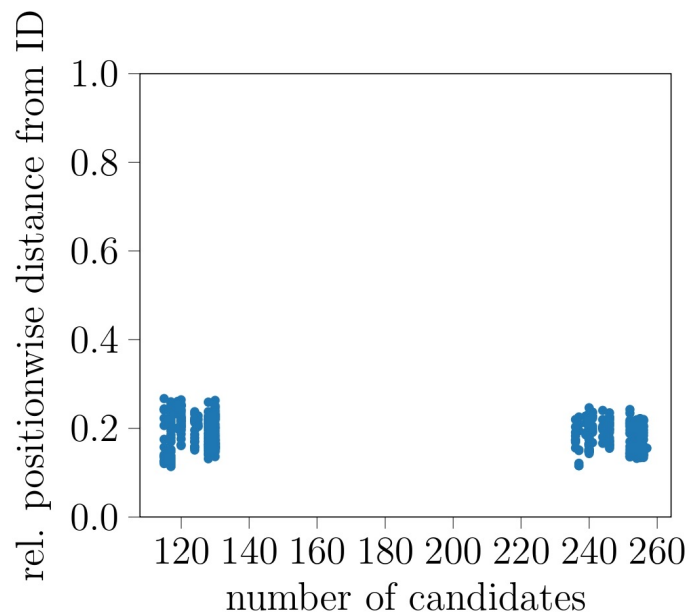
# Real-World Evidence



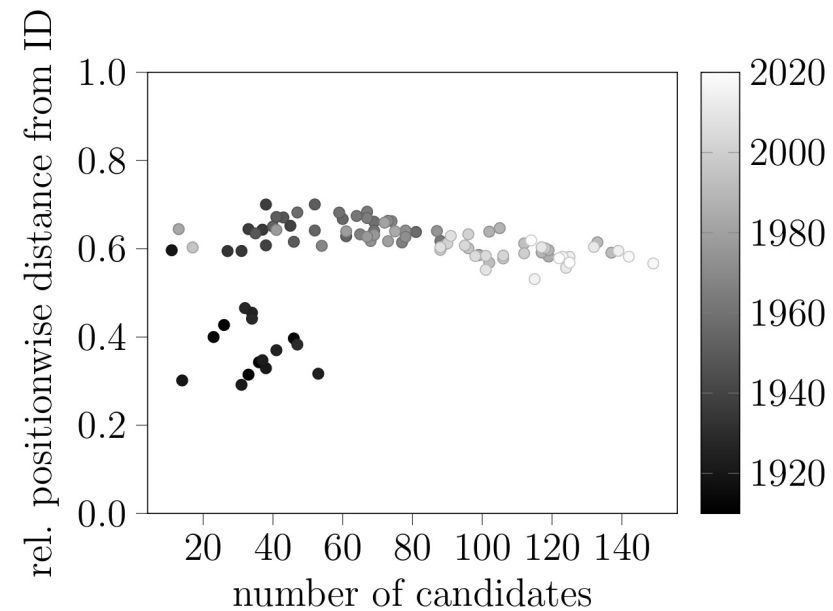
**Behaves as normalized Mallows model**



**Spotify charts**




**American football power rankings**



**Tour de France**

# Mallows Model: Warnings

- Be careful when varying the number of candidates: Trends could be artifact of Mallows model.
- Statements about certain ranges of dispersion parameter unlikely to generalize for other candidate numbers.
- Be careful how to select values of dispersion parameter in experiments to ensure meaningful coverage.
- Problems get intensified for generalizations such as Mallows mixtures.

 Expected Frequency Matrices of Elections: Computation, Geometry, and Preference Learning, Boehmer et al., NeurIPS-22

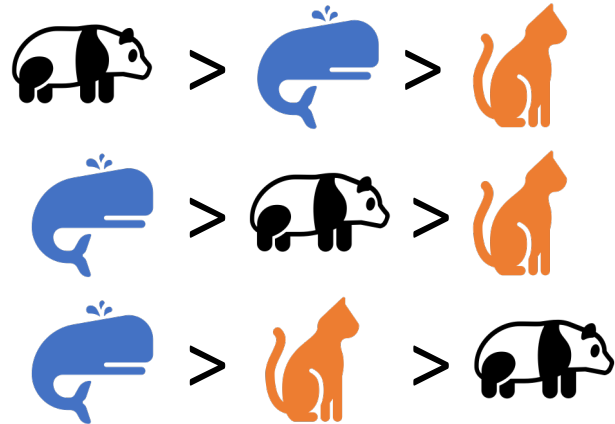
 Application-Oriented Collective Decision Making, Boehmer, PhD thesis






10 minutes

# Understanding Real-World Elections via Preference Learning




# Frequency Matrix



**Position Matrix**

			
1	2	0	1
2	1	1	1
3	0	2	1

**Frequency Matrix**

			
1	2/3	0	1/3
2	1/3	1/3	1/3
3	0	2/3	1/3

**Frequency Matrix of Vote Distribution (aka. probability distribution over votes)**

Entry (i,j): Probability that j is ranked in position i in a sampled vote.

# Learning Real-World Data

## **Idea**

*Given* parameterized vote distribution and (real-world) election  
*Compute* parameters most likely to produce election

## **Motivation**

- Quantify nature of examined elections
- Identify parameter values leading to realistic data

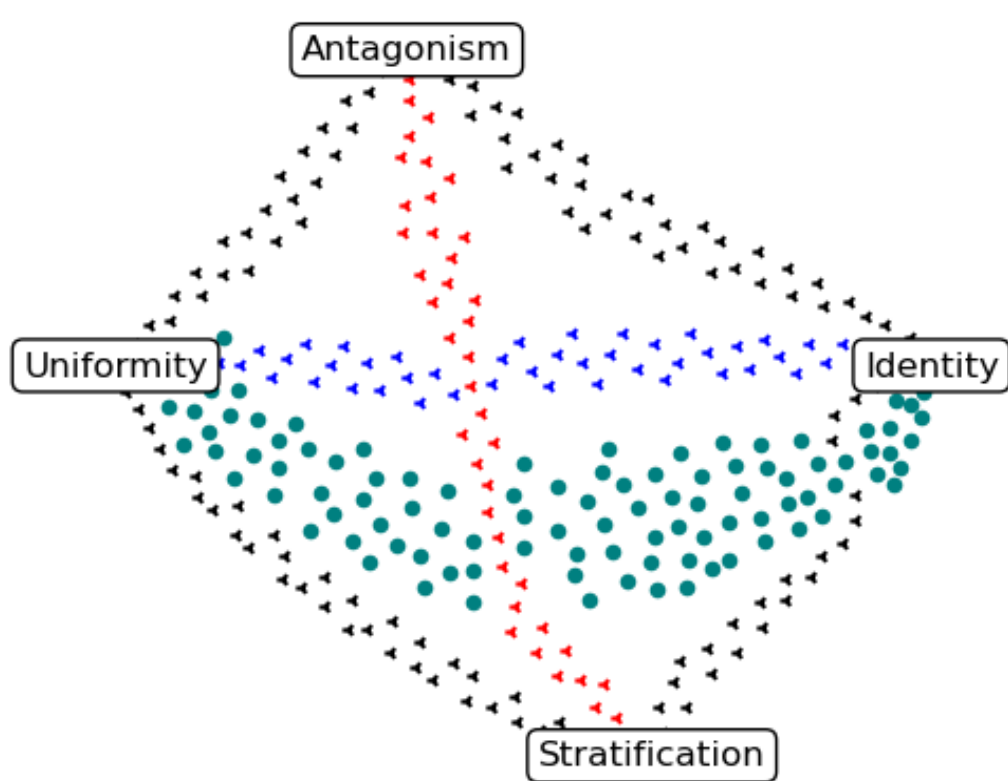
## **Approach**

For different distribution parameters:

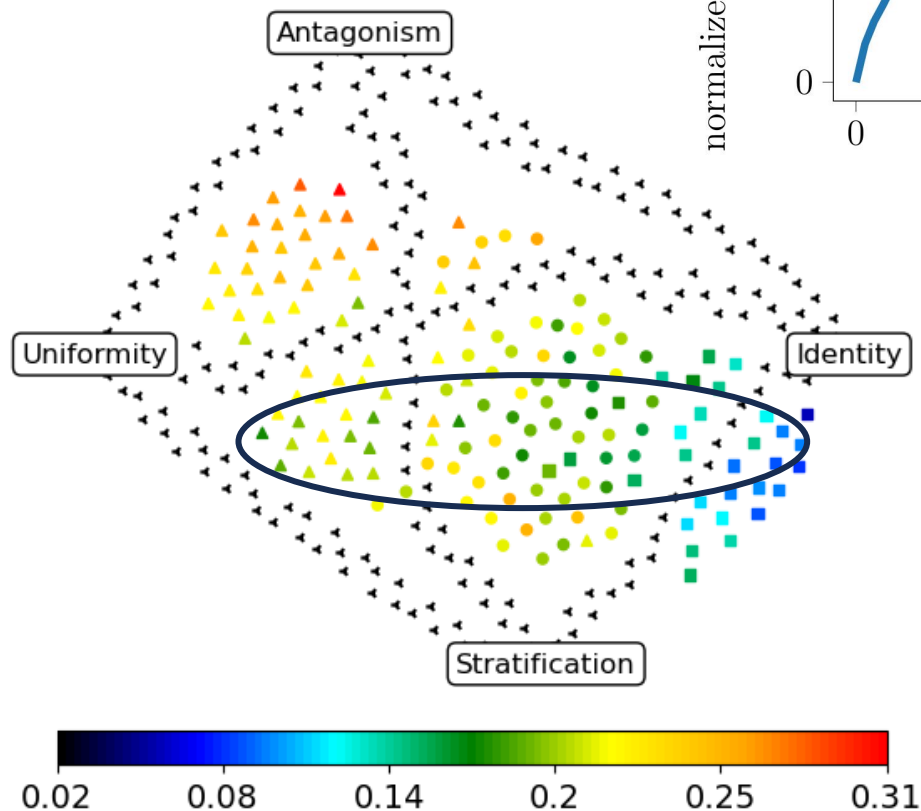
- Compute distance between frequency matrix of distribution and election

Return distribution parameters resulting in smallest distance

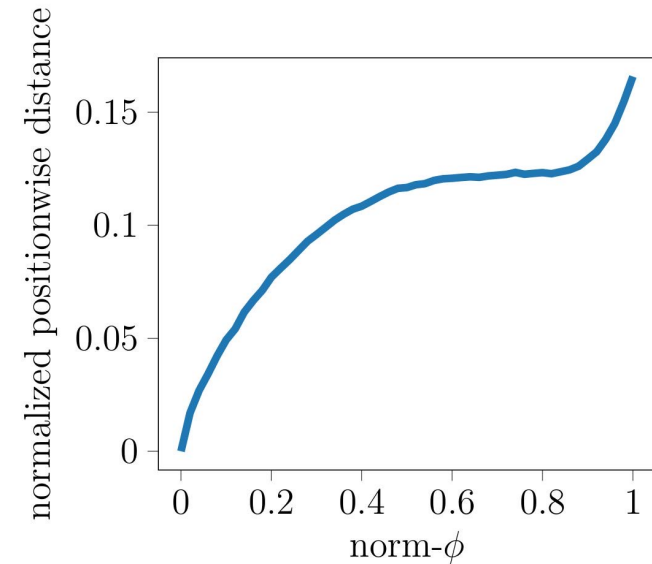
# Learning Single Mallows Models



Normalized dispersion parameter  $\text{norm-}\phi$   
of closest Mallows model  
Avg. 0.49



Normalized EMD-positionwise distance to  
closest Mallows matrix  
Avg. 0.192



# Learning Mixtures of Mallows Model

## Idea

Heterogeneous electorate with multiple central votes

## Procedure

Given two central votes  $v_1^*$  and  $v_2^*$  (over same candidate set), two dispersion parameters  $\text{norm-}\varphi > \text{norm-}\psi$ , and probability  $p$

- With probability  $p$ , sample from Mallows model with  $\text{norm-}\varphi$  and  $v_1^*$
- With probability  $1-p$ , sample from Mallows model with  $\text{norm-}\psi$  and  $v_2^*$

## Frequency matrix

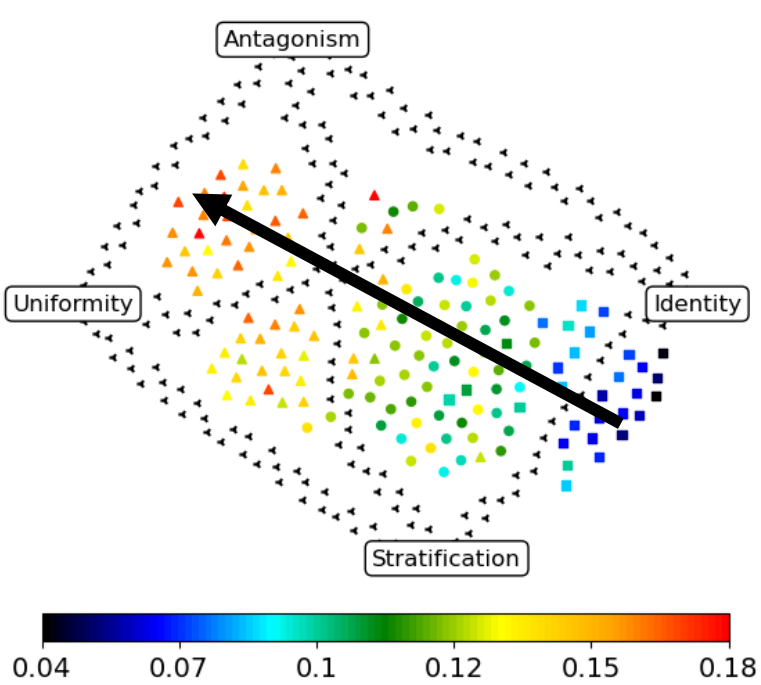
Weighted sum of matrices of individual models



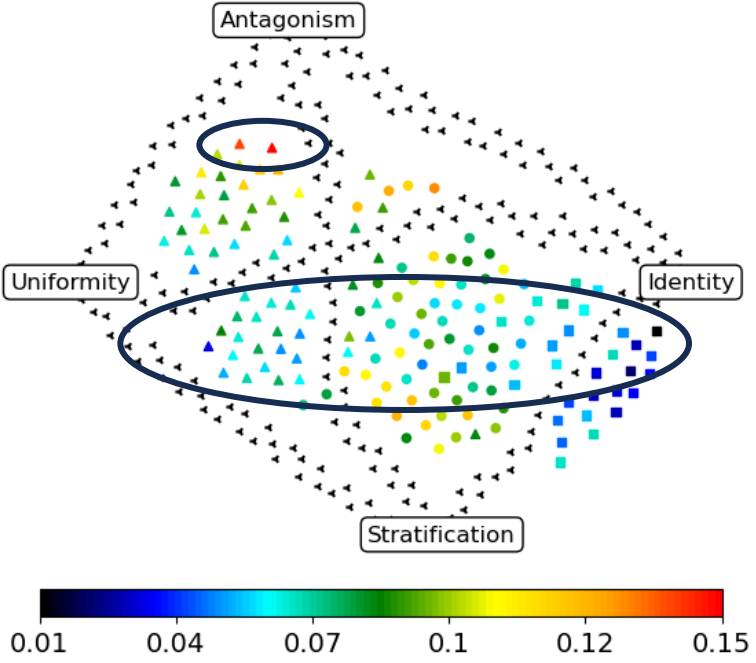
$$p * M(\text{norm-}\varphi, v_1^*) + (1-p) * M(\text{norm-}\psi, v_2^*)$$

with  $\text{norm-}\varphi > \text{norm-}\psi$

# Learning Mixtures of Mallows Model



Distance to frequency matrix of closest Mallows mixture  
Avg. 0.12 (-0.07)

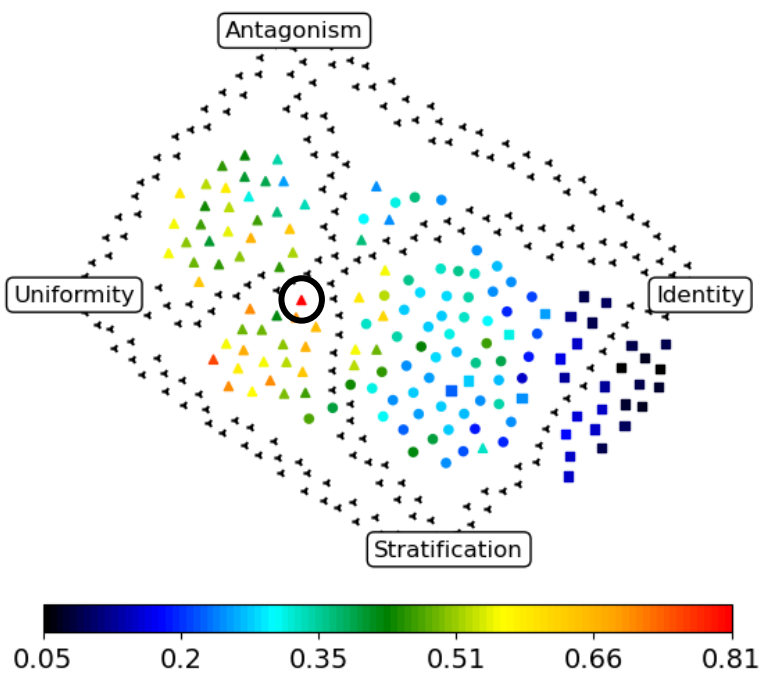


Distance "gain" by using mixture instead of  
single Mallows model

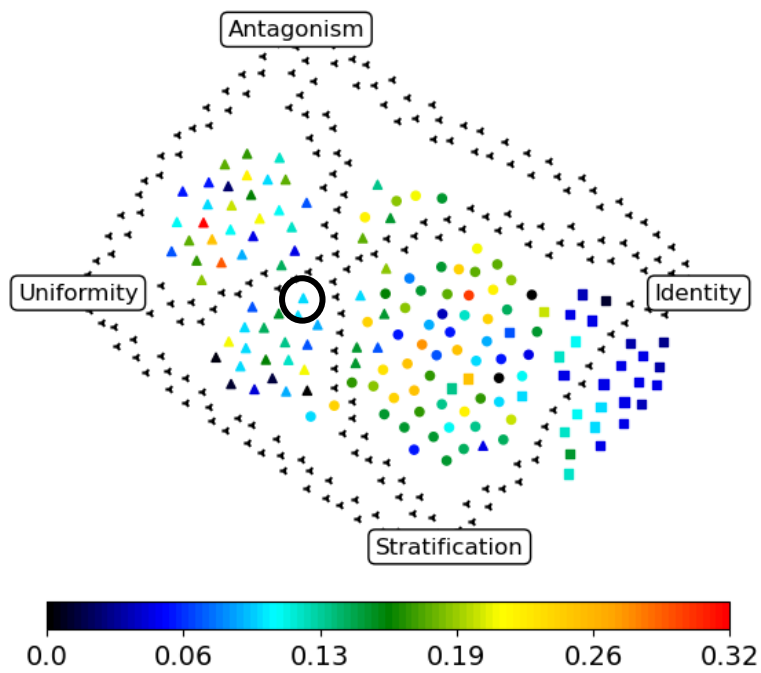
# Learning Mixtures of Mallows Model

$$p * M(\text{norm-}\varphi, v_1^*) + (1-p) * M(\text{norm-}\psi, v_2^*)$$

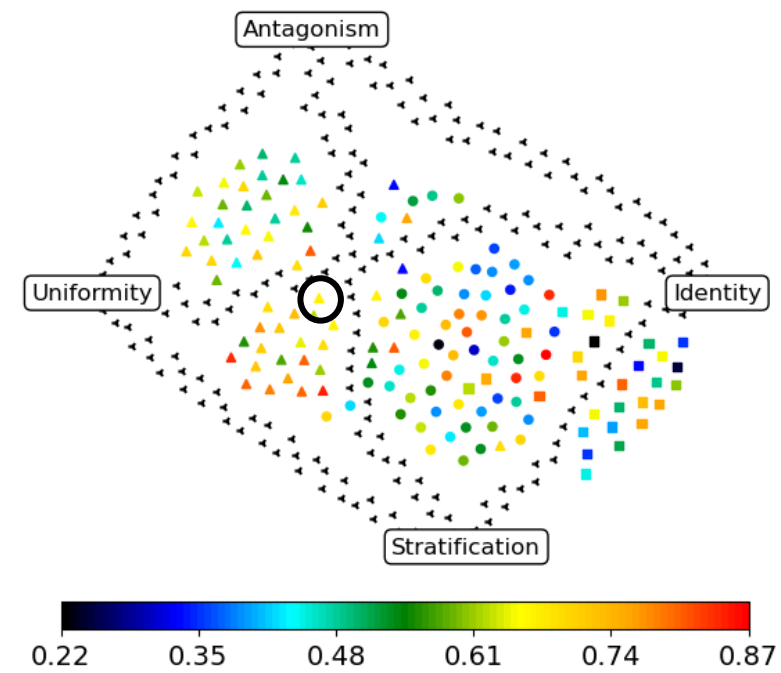
with  $\text{norm-}\varphi > \text{norm-}\psi$



Normalized dispersion parameter  $\text{norm-}\varphi$   
of closest mixture  
Avg. 0.353



Normalized dispersion parameter  $\text{norm-}\psi$   
of closest mixture  
Avg. 0.128



Sampling probability of closest mixture  
Avg. 0.6

# How to Sample Realistic Data Using the Mallows model?

## Observations

- Mallows elections capture relevant part of map of elections well
- Mixtures of Mallows models even more powerful/general

## Procedure

- Normalized Mallows model with uniformly at random chosen norm- $\varphi$  between 0 and 0.92
- Mixtures of Mallows models:  $p \in [0.35, 0.8]$ , norm- $\varphi \in [0.05, 0.6]$ , norm- $\psi \in [0, 0.25]$ , and swap distance between  $v_1^*$  and  $v_2^* \in [0.35, 0.6]$

**Mapel**

Matchings

**Further Applications**

Approval Elections

Map of Rules

Data!

Introduction to voting

**Experiments in Computational Social Choice**

Preference Learning

Mallows

Real-Life Data

**Use Cases (Elections)**

Swap Distance

Map of Elections

Winners

Approximations

Distances

Force-Directed

Compass Elections

Election Results

Positionwise

Embedding Algorithms

UN

ST

ID

Committees

Running Time

Verification

Mapel

Matchings

Further Applications

Approval Elections

Map of Rules

Introduction to voting

Data!

Experiments in Computational Social Choice

Preference Learning

Mallows

Real-Life Data

Use Cases (Elections)

Swap Distance

Map of Elections

Approximations

Distances

Positionwise

Embedding Algorithms

Force-Directed

UN

Compass Elections

ID

Winners

Election Results

Running Time

Verification

ST

Committees



25 minutes

# Approval Elections

Further Applications

# Instance of Approval Election

Candidates: {  ,  ,  ,  ,  }

Voters:

$v_1$ : {  ,  }

$v_2$ : {  ,  ,  }

$v_3$ : {  ,  ,  }

$v_4$ : {  ,  ,  ,  }

$v_5$ : {  }

Approvalwise distance



# Approvalwise distance

$v_1$ : {  ,  }

$v_2$ : {  ,  ,  }

$v_3$ : {  ,  ,  }

$v_4$ : {  ,  ,  ,  }

$v_5$ : {  }

$u_1$ : {  ,  ,  ,  }

$u_2$ : {  ,  }

$u_3$ : {  ,  }

$u_4$ : {  ,  ,  }

$u_5$ : {  ,  }



Score: 3 4 1 2 3

Sorted vector: [4, 3, 3, 2, 1]



2 5 1 3 2

[5, 3, 2, 2, 1]

$$\ell_1([4, 3, 3, 2, 1], [5, 3, 2, 2, 1]) =$$

# Approvalwise distance

Pseudodistance

$v_1$ : { ,  }

$v_2$ : { , ,  }

$v_3$ : { , ,  }

$v_4$ : { , , ,  }






$u_1$ : { , , ,  }






$u_2$ : { ,  }

$u_3$ : { ,  }

$u_4$ : { , ,  }

Can be computed in polynomial time





  
 Score: 3 4 1 2 3  
 Sorted vector: [4, 3, 3, 2, 1]





  
 2 5 1 3 2  
 [5, 3, 2, 2, 1]

$$l_1([4, 3, 3, 2, 1], [5, 3, 2, 2, 1]) = 2$$

# Hamming distance

# Hamming distance

$v_1$ : { ,  }

$v_2$ : { , ,  }

$v_3$ : { , ,  }

$v_4$ : { , , ,  }

$v_5$ : {  }

$u_1$ : { , , ,  }

$u_2$ : { ,  }

$u_3$ : { ,  }

$u_4$ : { , ,  }

$u_5$ : { ,  }

Matching



# Hamming distance

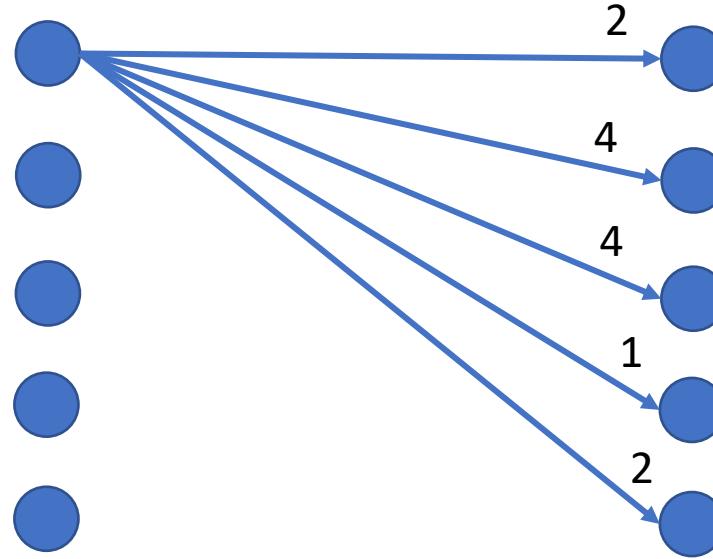
$v_1$ : { ,  }

$v_2$ : { , ,  }

$v_3$ : { , ,  }

$v_4$ : { , , ,  }

$v_5$ : {  }



$u_1$ : { , , ,  }

$u_2$ : { ,  }

$u_3$ : { ,  }

$u_4$ : { , ,  }

$u_5$ : { ,  }

# Hamming distance

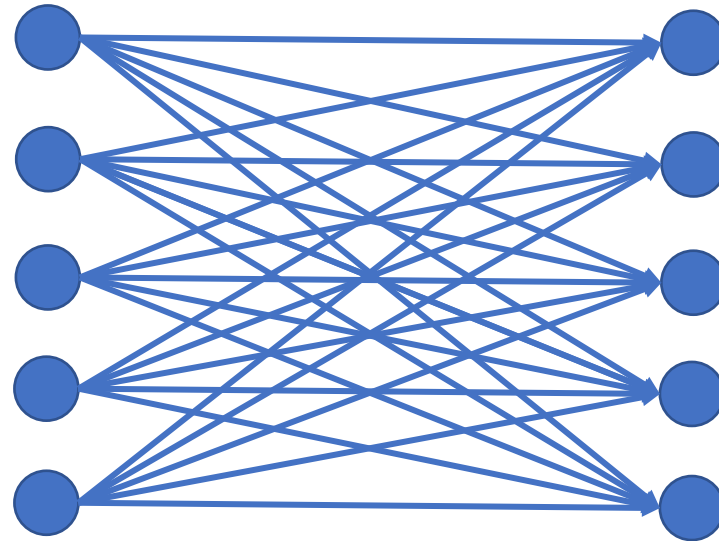
$v_1$ : { ,  }

$v_2$ : { , ,  }

$v_3$ : { , ,  }

$v_4$ : { , , ,  }

$v_5$ : {  }



$u_1$ : { , , ,  }

$u_2$ : { ,  }

$u_3$ : { ,  }

$u_4$ : { , ,  }

$u_5$ : { ,  }

# Hamming distance

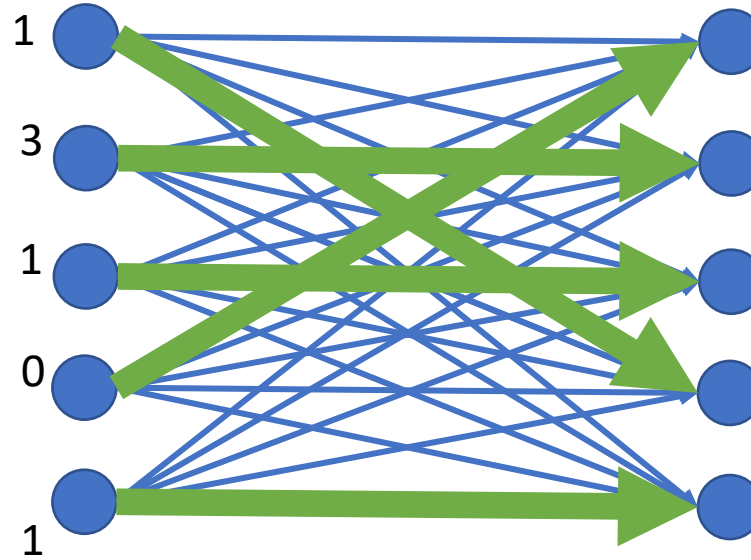
$v_1$ : { ,  }

$v_2$ : { , ,  }

$v_3$ : { , ,  }

$v_4$ : { , , ,  }

$v_5$ : {  }



$u_1$ : { , , ,  }

$u_2$ : { ,  }

$u_3$ : { ,  }

$u_4$ : { , ,  }

$u_5$ : { ,  }

$$1 + 3 + 1 + 0 + 1 = 6$$

# Hamming distance

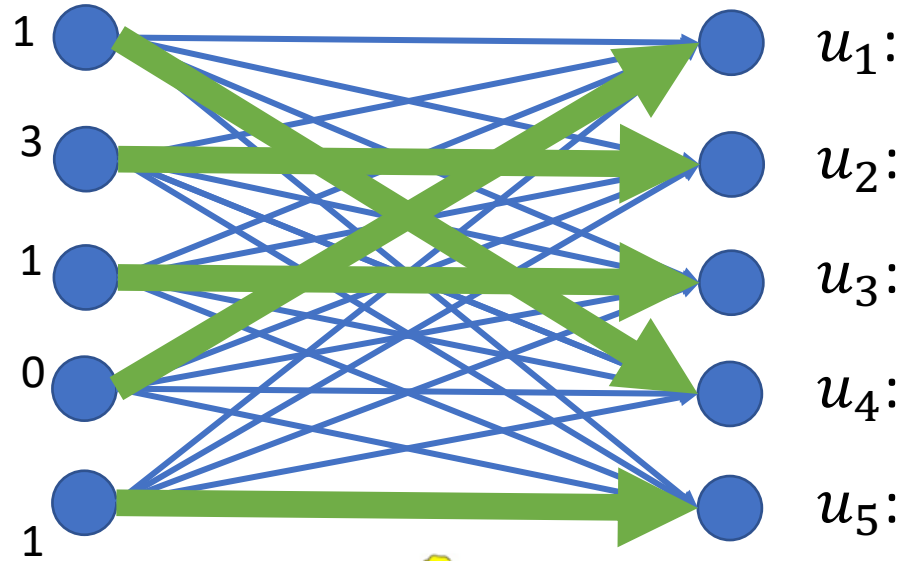
$v_1$ : { ,  }

$v_2$ : { , ,  }

$v_3$ : { , ,  }

$v_4$ : { , , ,  }

$v_5$ : {  }



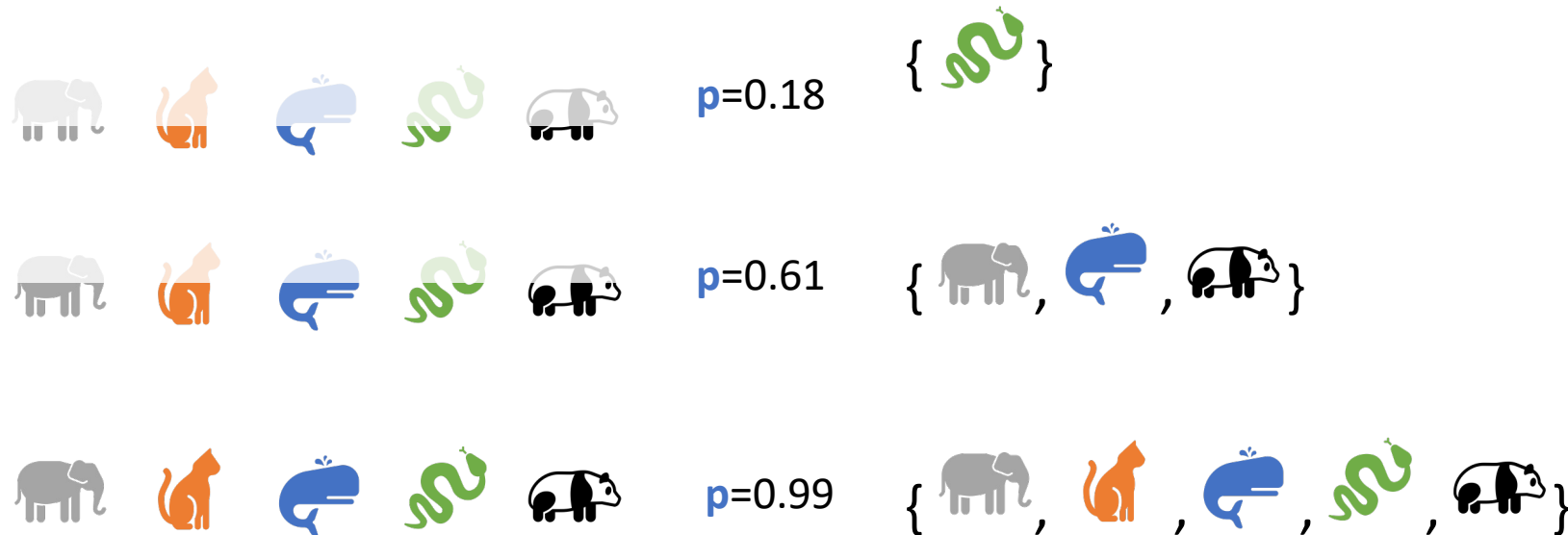
☹️ Unfortunately it is NP-hard ☹️





# p-Impartial Culture

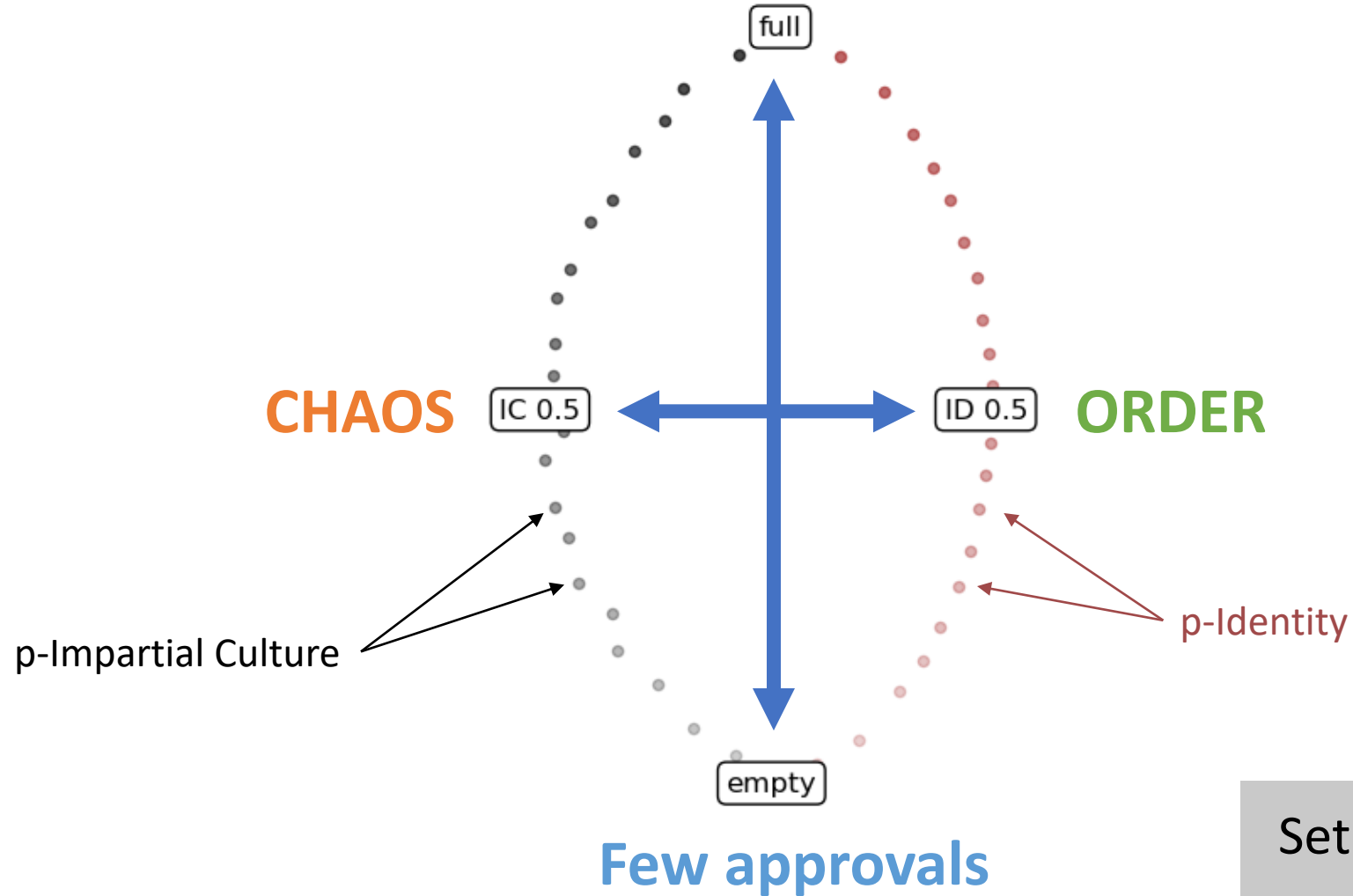
To generate a vote, for each candidate we flip an assymmetric coin, and with probability  $p$  we put that candidate in our ballot



# p-Identity

To generate first vote, we approved  $\lfloor p \cdot m \rfloor$  candidates selected uniformly at random. All other votes are its copies.

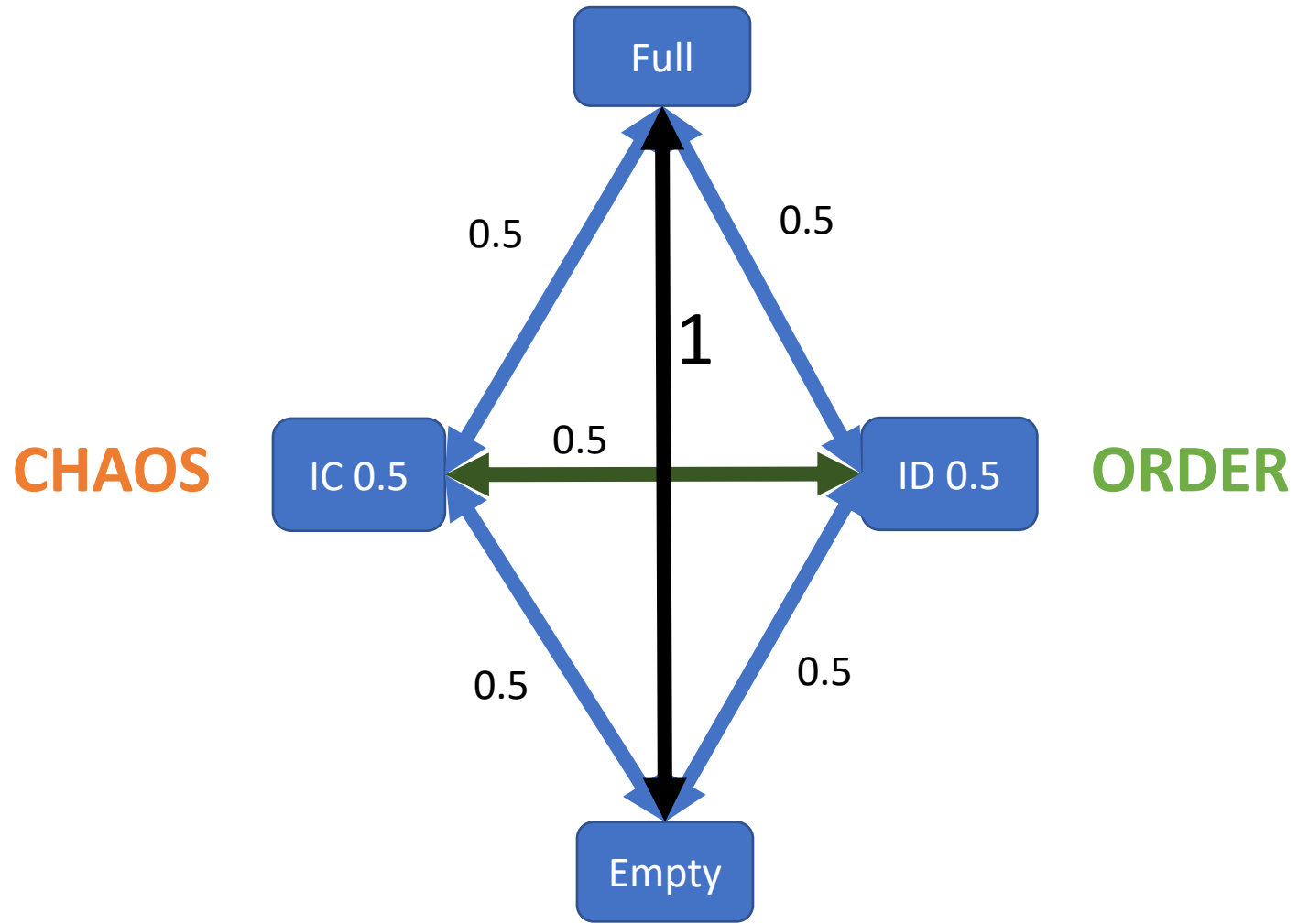
Many approvals



### Setup

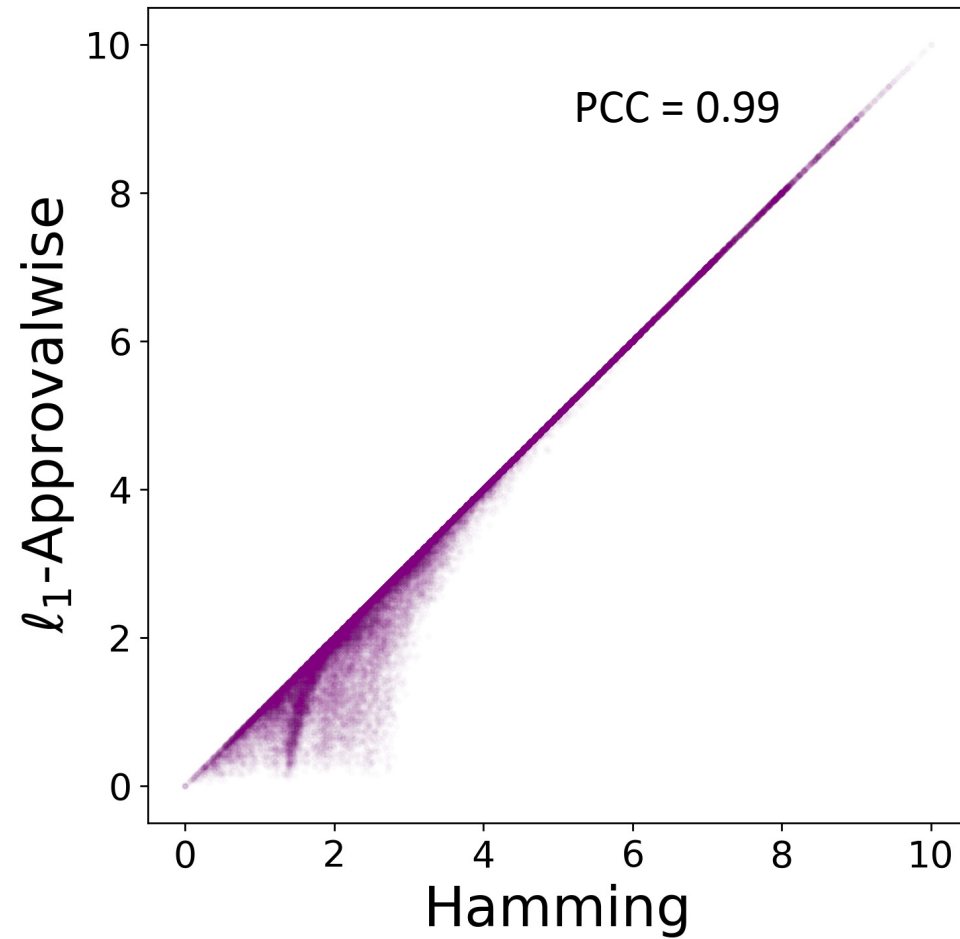
number of candidates	50
number of voters	100

Many approvals



Few approvals

# Correlation



## Setup

number of candidates	10
number of voters	50

# p-Identity with $\phi$ -Resampling

Initial ballot (from p-ID)



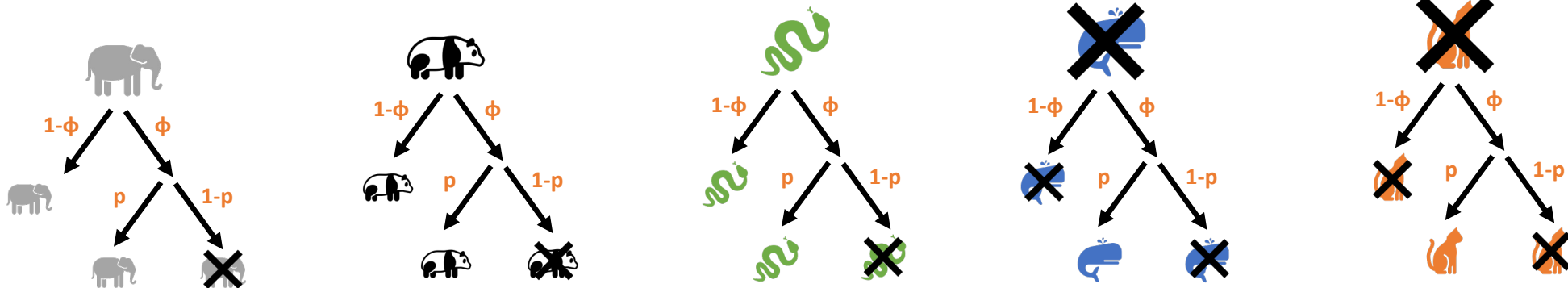
To generate a vote:

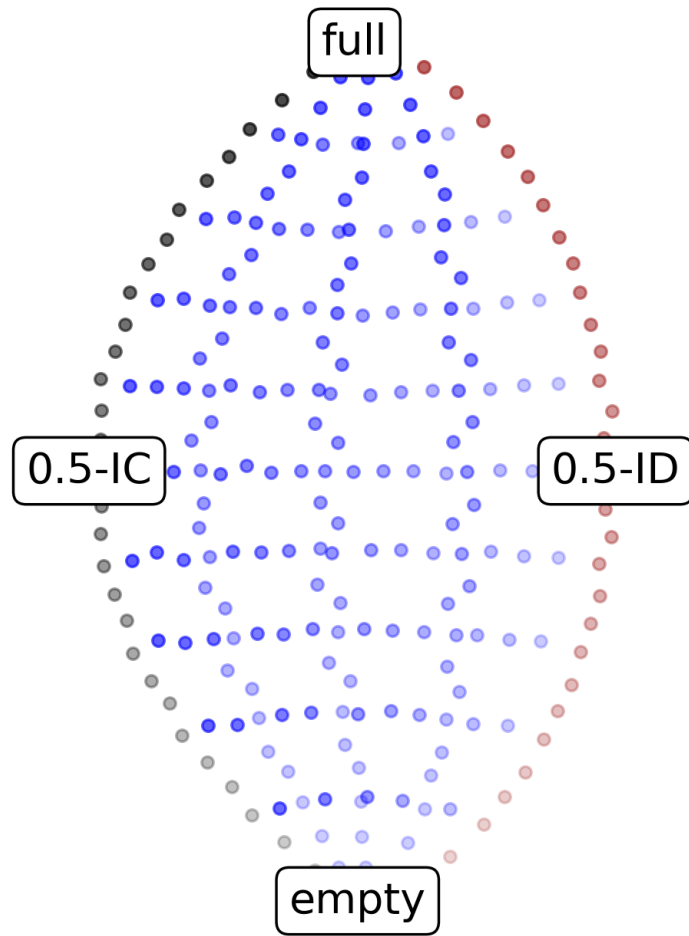
Step 0: copy initial ballot

Step 1: for each candidate, **resample** that candidate with probability  $\phi$

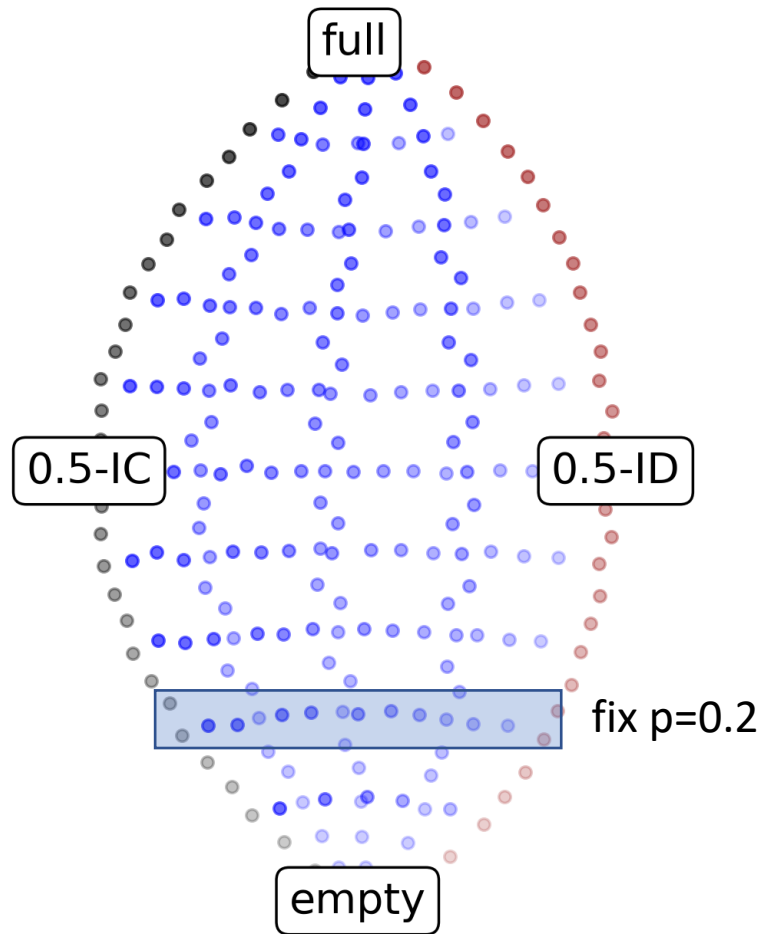
**not reverse**

(resample = toss an assymmetric coin; approve with probability  $p$ )

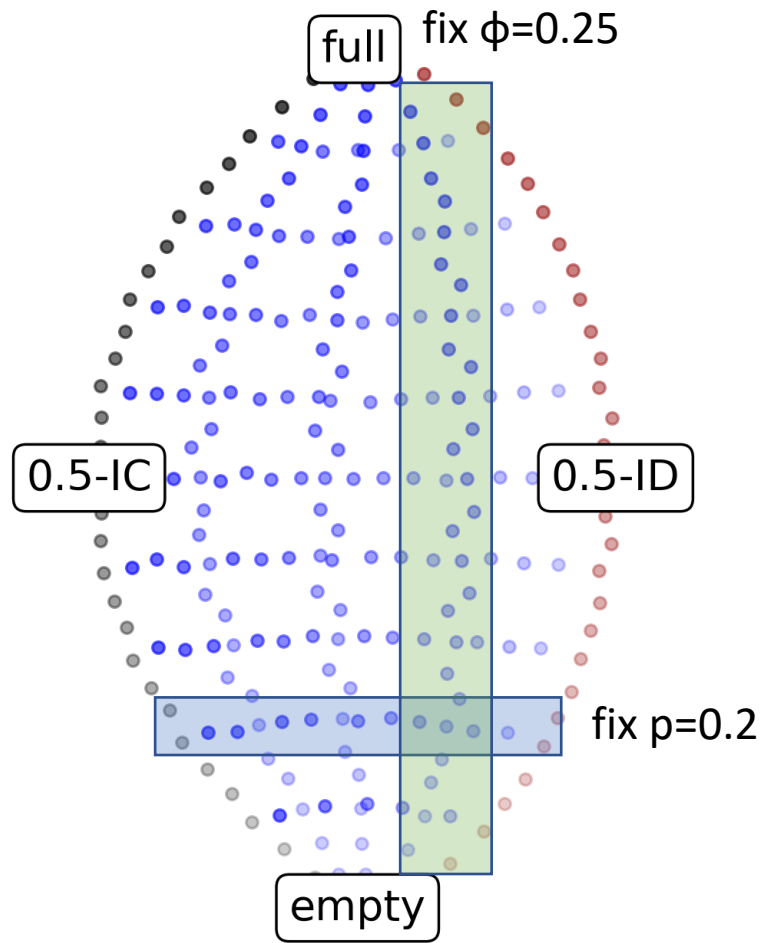




$p$ -Identity with  $\phi$ -Resampling

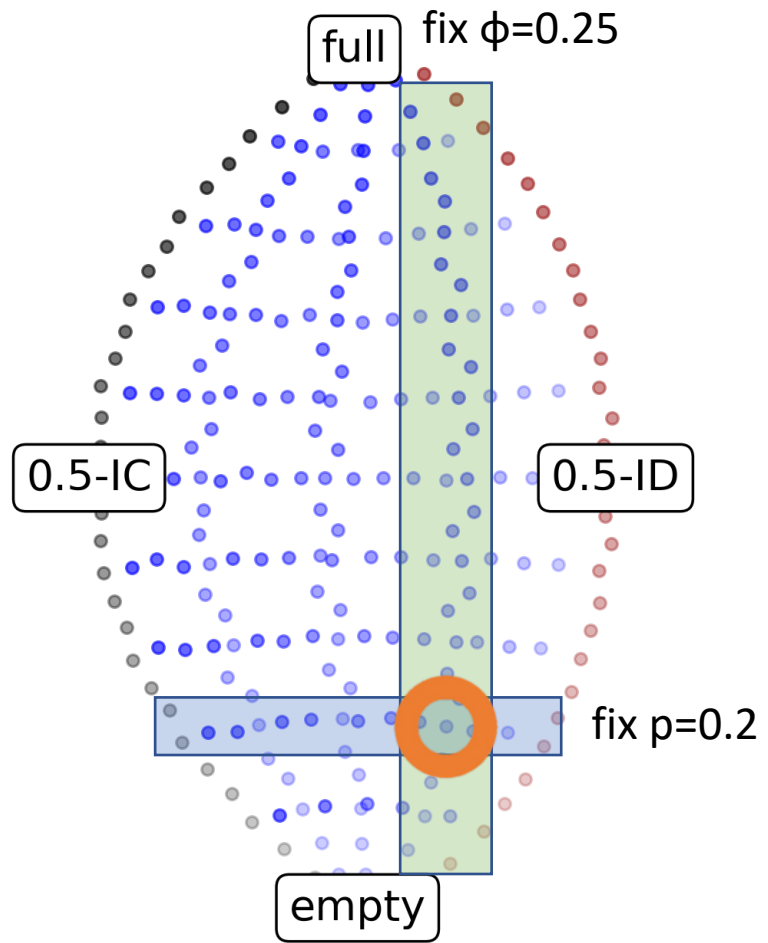


$p$ -Identity with  $\phi$ -Resampling



$p$ -Identity with  $\phi$ -Resampling





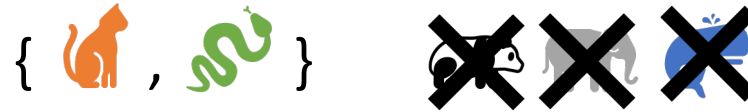
$p$ -Identity with  $\phi$ -Resampling

# Disjoint $p$ -Identity with $\phi$ -Resampling

First initial ballot



Second initial ballot



To generate a vote:

Step 0: copy one of the initial votes

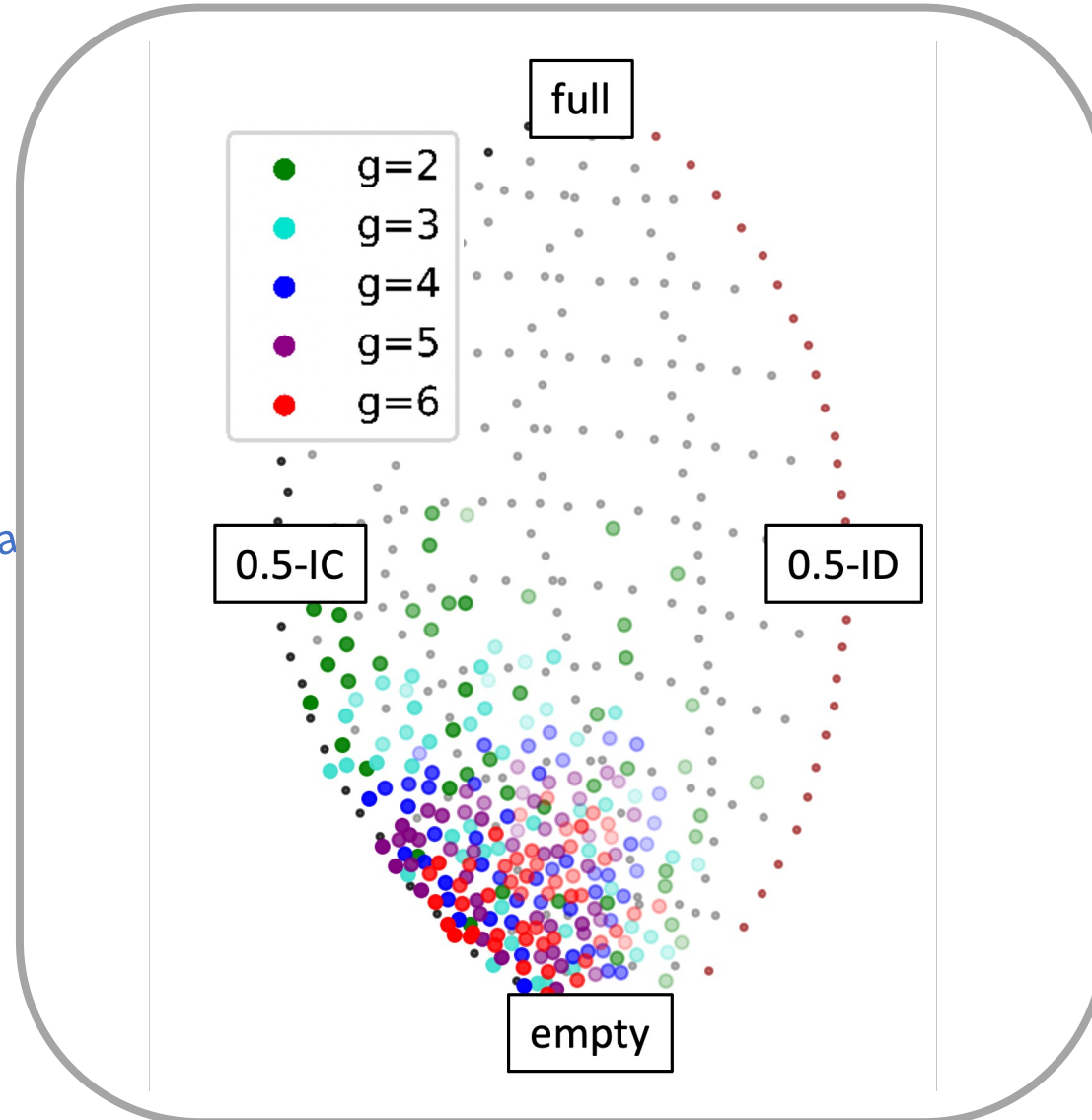
Step 1: for each candidate, **resample** that candidate with probability  $\phi$

# Disjoint $p$ -Identity with $\phi$ -Resampling

First initial ballot

Second initial ballot

To genera



# $(p, \phi)$ Noise Model

Initial ballot (from  $p$ -IC)    { , ,  }    ~~~~    ~~~~

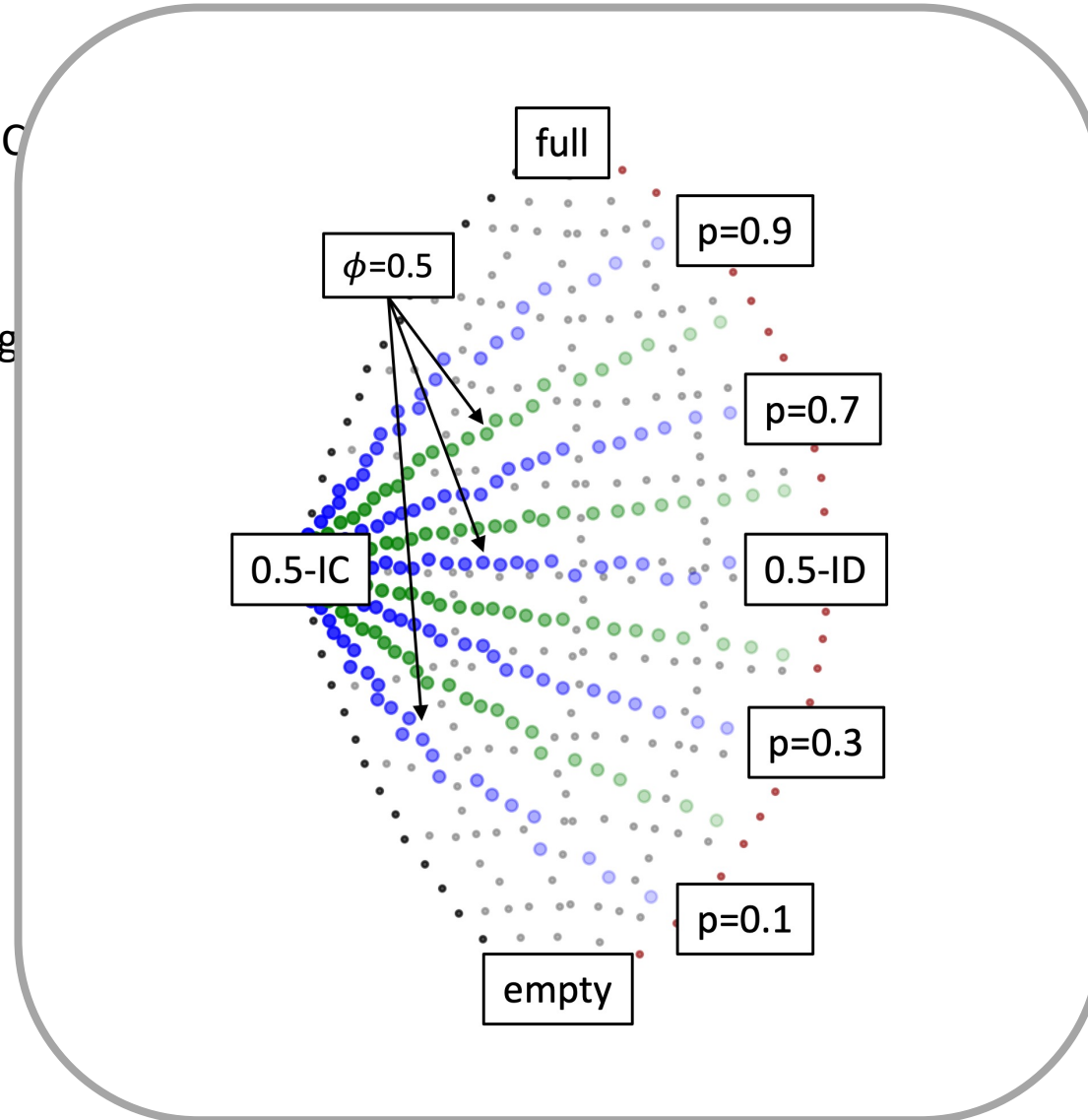
The probability of a given vote is proportional to its **Hamming** distance from the initial ballot

# $(p, \phi)$ Noise Model

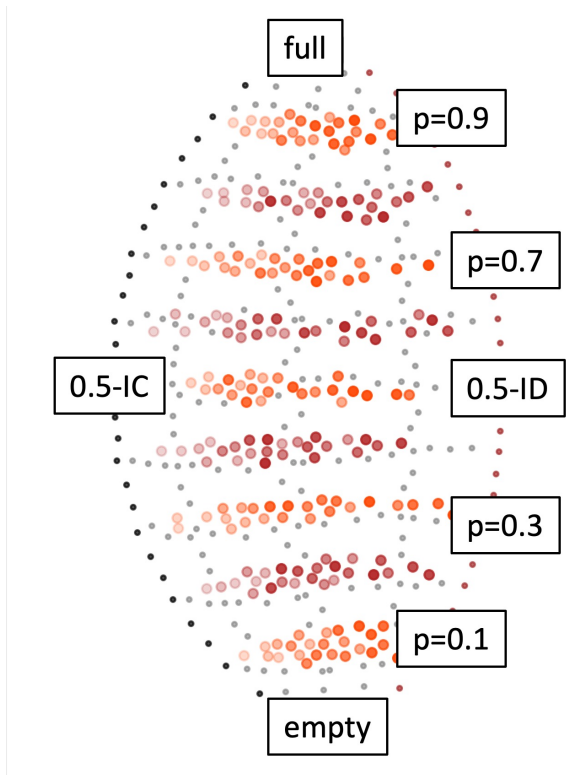
Initial ballot (from  $p$ -IC)

The probability of a g

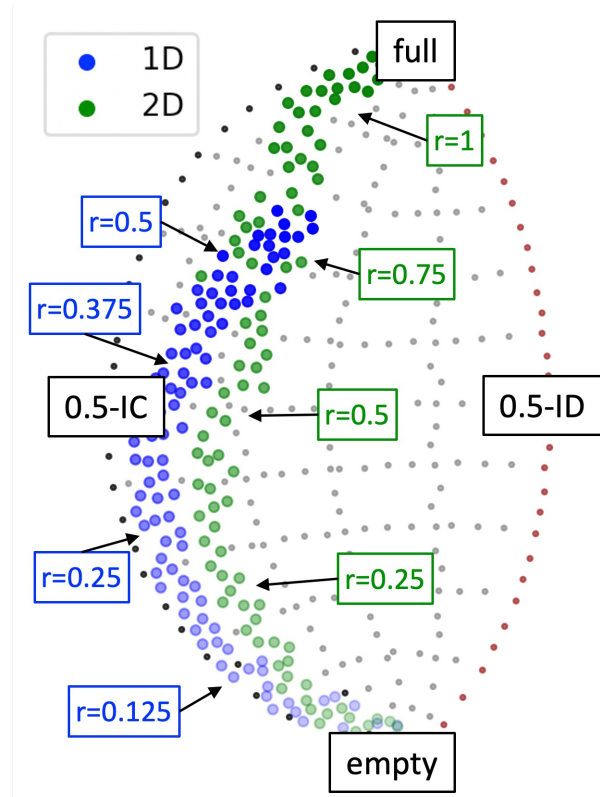
the initial ballot



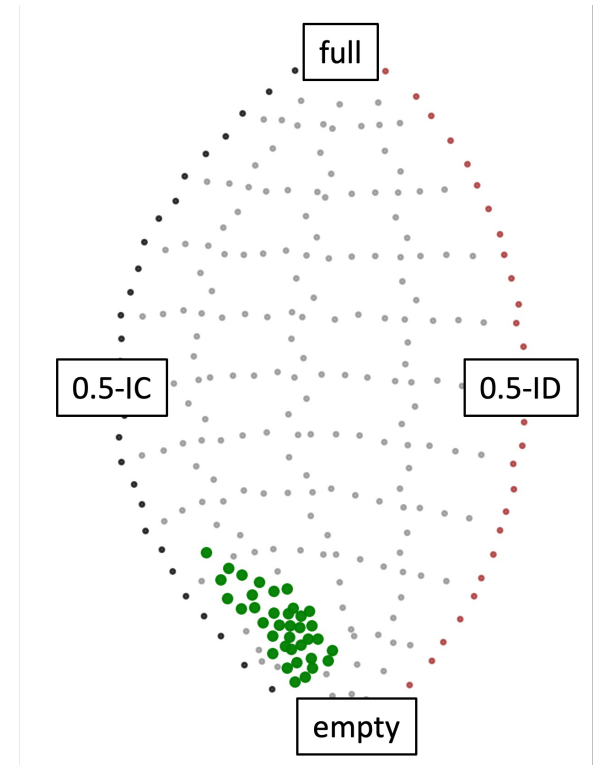
# Other Cultures



$(p, \alpha)$  Urn Model

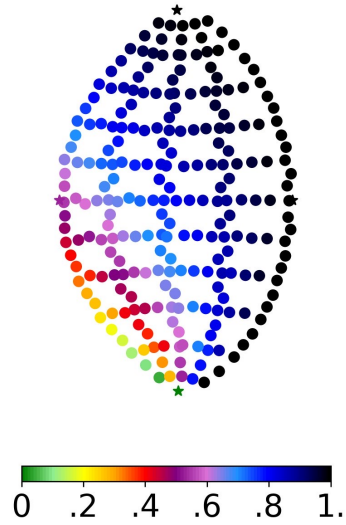


Euclidean

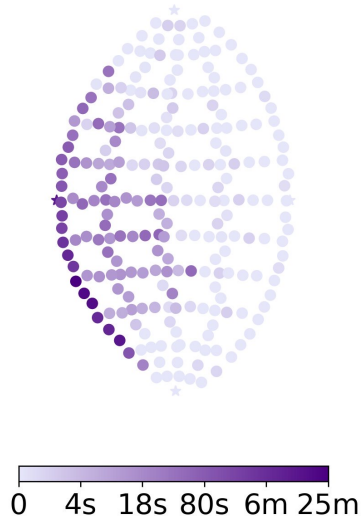


Real life data

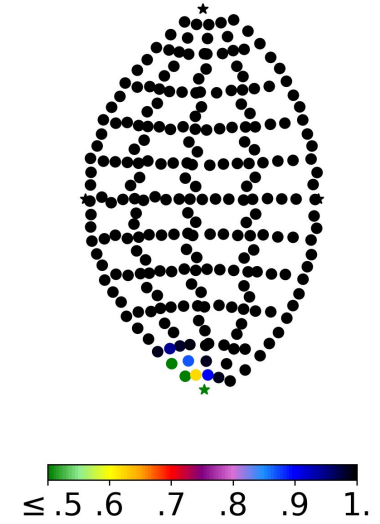
max. approval score



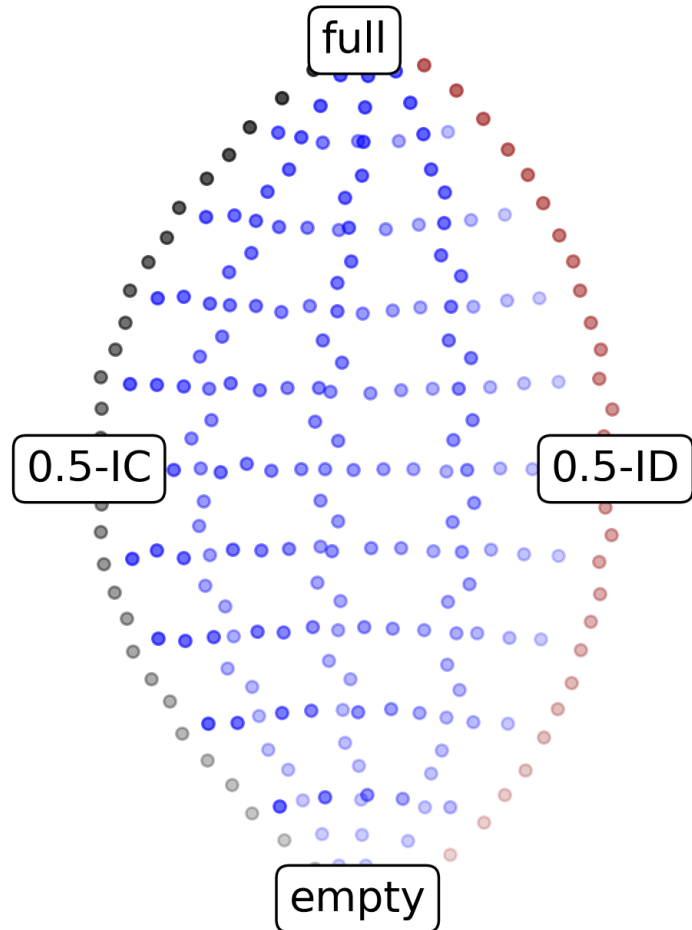
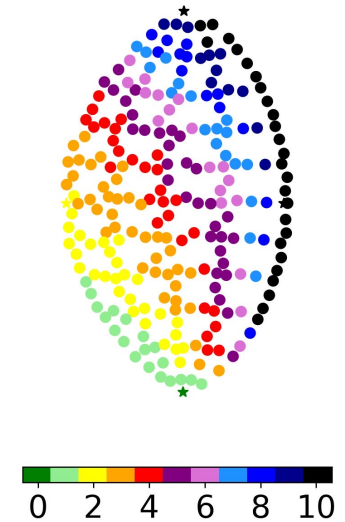
PAV runtime



voters in coh. groups



coh. level



p-Identity with  $\phi$ -Resampling

**Mapel**

Matchings

Further Applications

Approval Elections

Map of Rules

Data!

Introduction to voting

**Experiments in Computational Social Choice**

Preference Learning

Mallows

Real-Life Data

Map of Elections

Use Cases (Elections)

Swap Distance

Distances

Approximations

Positionwise

Embedding Algorithms

Force-Directed

Verification

UN

Compass Elections

ST

ID

AN

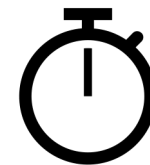
Winners

Election Results

Committees

Running Time





10 minutes

# Create your own map of elections!

Introduction to Mapel Software Package 2/2

Mapel

Matchings

Further Applications

Approval Elections

Map of Rules

Introduction to voting

Data!

Experiments in Computational Social Choice

Preference Learning

Mallows

Real-Life Data

Use Cases (Elections)

Swap Distance

Map of Elections

Approximations

Distances

AN

Winners

Election Results

Force-Directed

Compass Elections

UN

ID

Running Time

Positionwise

Embedding Algorithms

ST

Committees

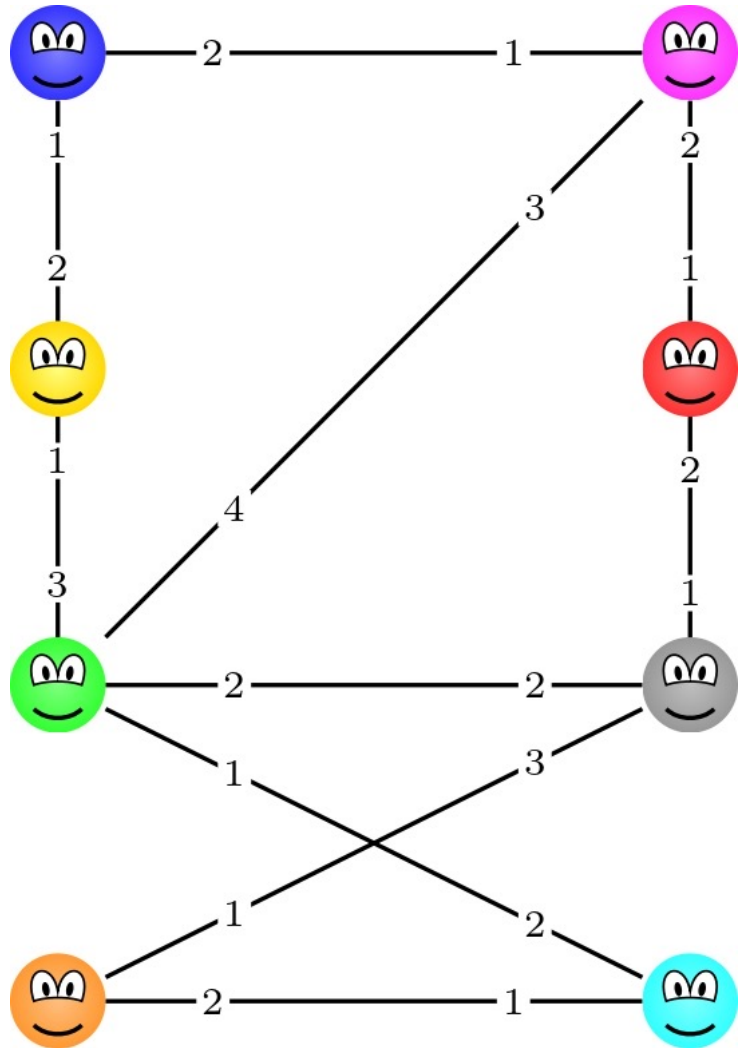
Verification



15 minutes

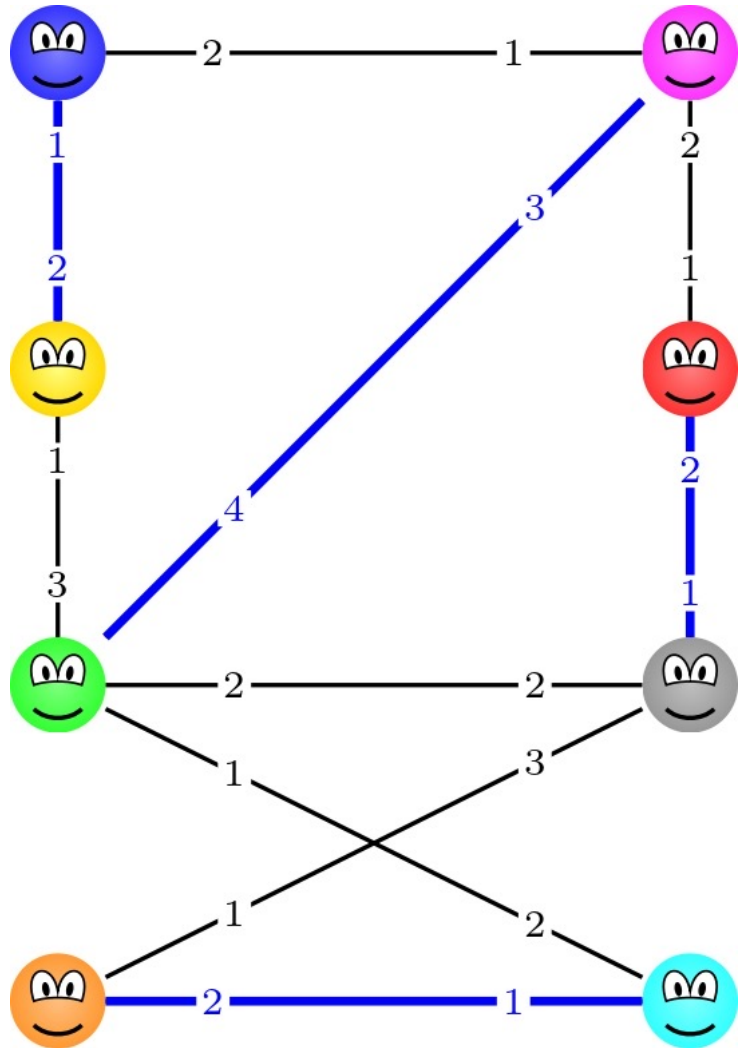
# Maps for Matchings under Preferences

# Stable Roommates



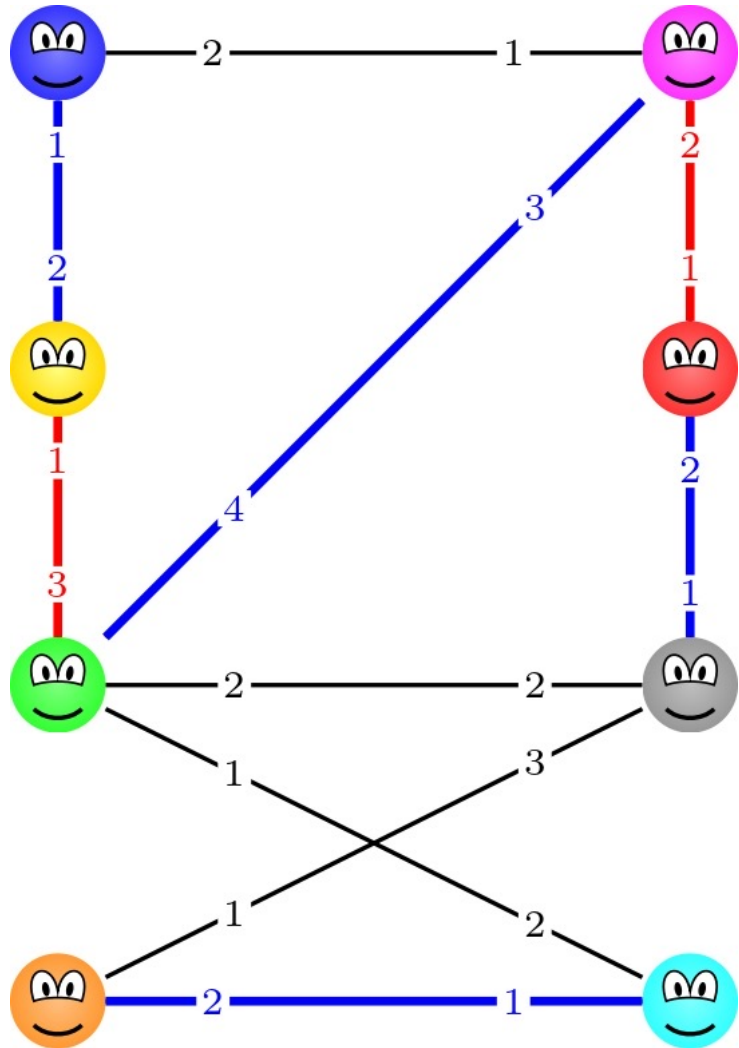
Input: Agents with strict preferences over each other.

# Stable Roommates



Input: Agents with strict preferences over each other.

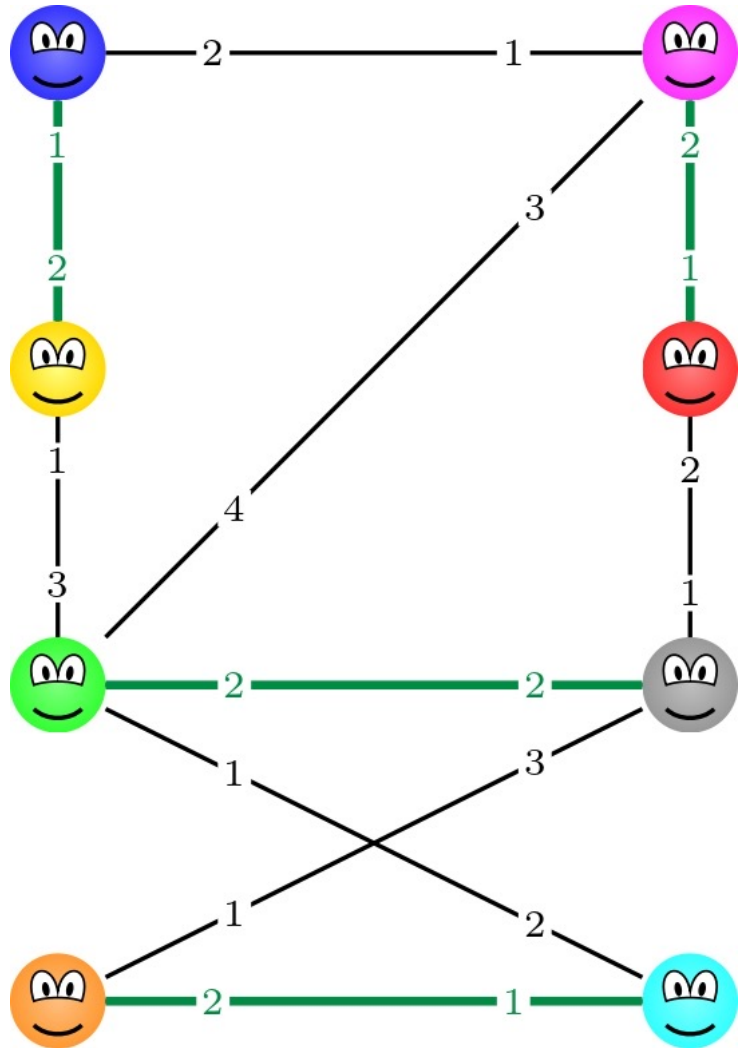
# Stable Roommates



Input: Agents with strict preferences over each other.

An agent pair **blocks** matching M if both agents prefer each other to current partner.

# Stable Roommates

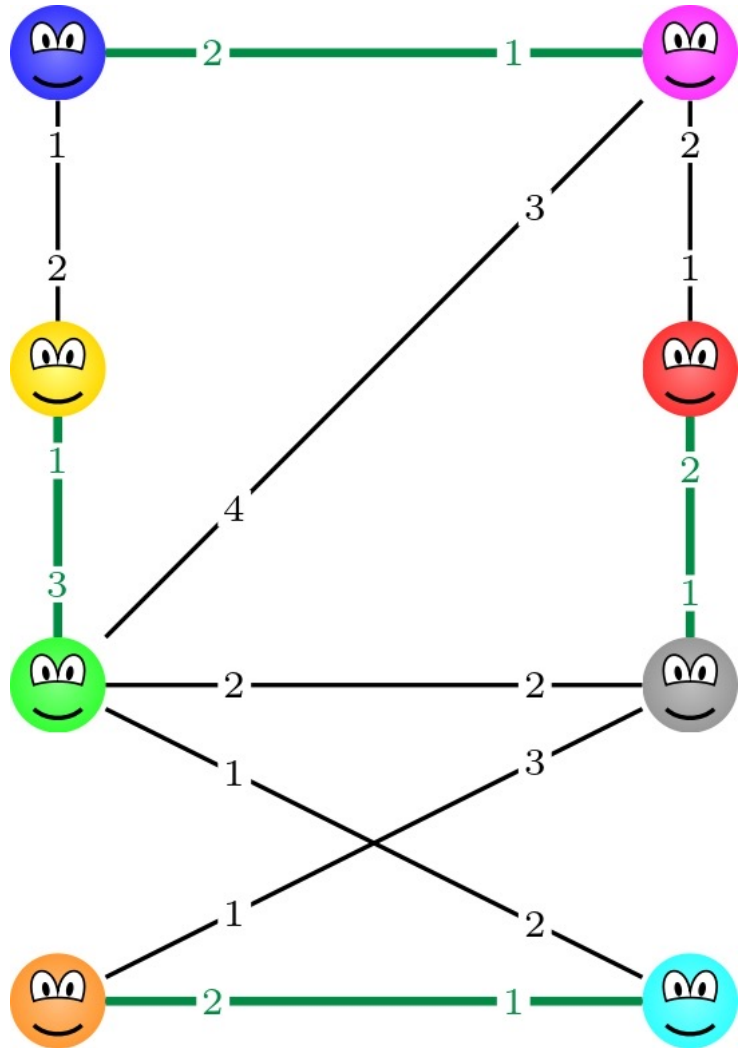


Input: Agents with strict preferences over each other.

An agent pair **blocks** matching M if both agents prefer each other to current partner.

Goal: Find a stable matching, i.e., a matching without a blocking pair.

# Stable Roommates



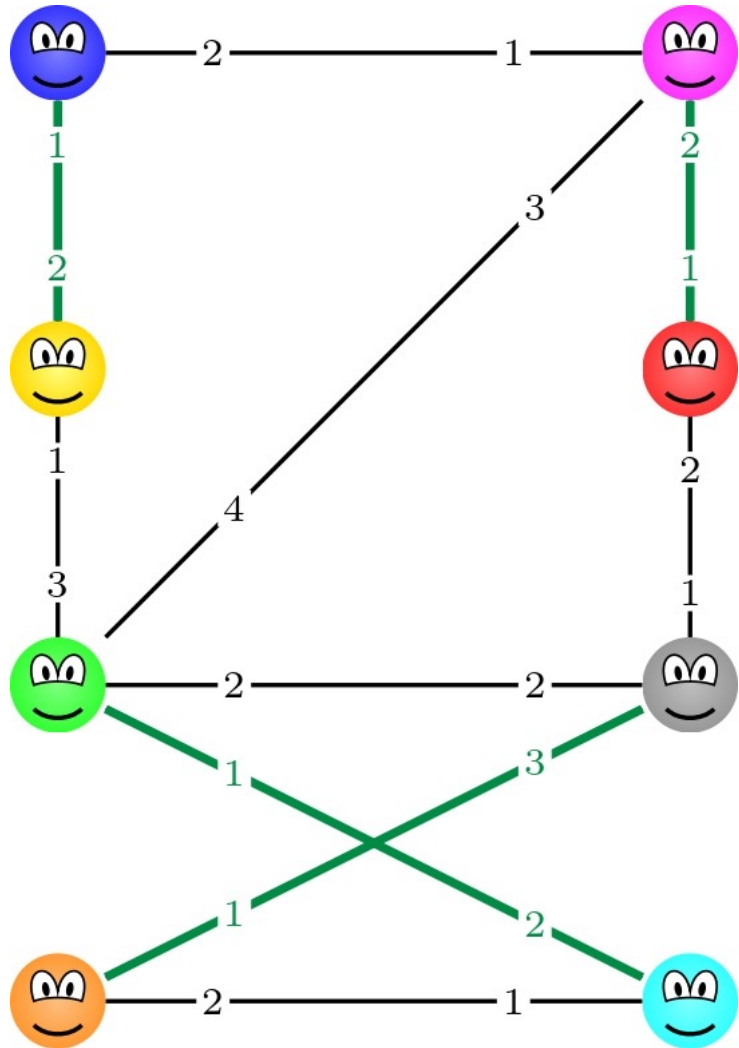
Input: Agents with strict preferences over each other.

An agent pair **blocks** matching  $M$  if both agents prefer each other to current partner.

Goal: Find a stable matching, i.e., a matching without a blocking pair.



# Stable Roommates

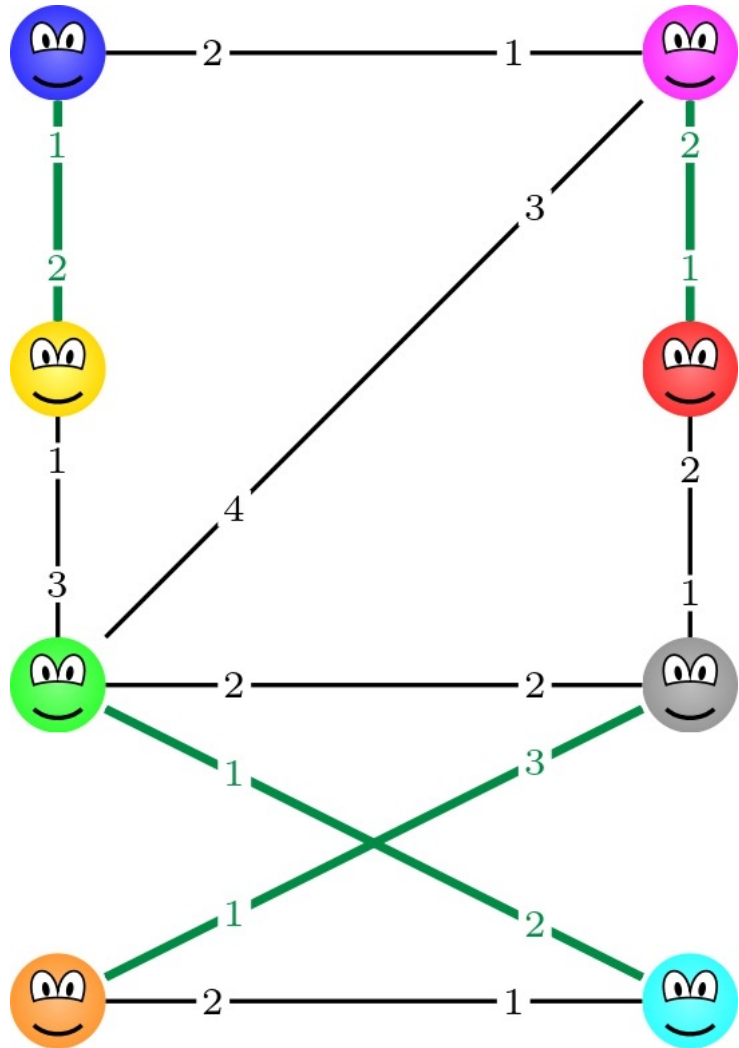


Input: Agents with strict preferences over each other.

An agent pair **blocks** matching  $M$  if both agents prefer each other to current partner.

Goal: Find a stable matching, i.e., a matching without a blocking pair.

# Stable Roommates



Input: Agents with strict preferences over each other.

An agent pair **blocks** matching  $M$  if both agents prefer each other to current partner.

Goal: Find a stable matching, i.e., a matching without a blocking pair.

# Status Quo

- Numerous works on theoretical aspects of stable matching problems with real-world impact.
- Some works contain empirical investigations **but** far away from standard with most of them only using uniformly at random sampled preferences.

# Step 1: Distance Measure

## Central Question

How to measure the similarity of two Stable Roommates instances?

(Assumption: Both instances have same number of agents)

## Positionwise Distance



General popularity/quality of agents in the instance.



Position matrix completely ignores mutual opinions, i.e., what agents think of each other (agents are "voters" and "candidates")

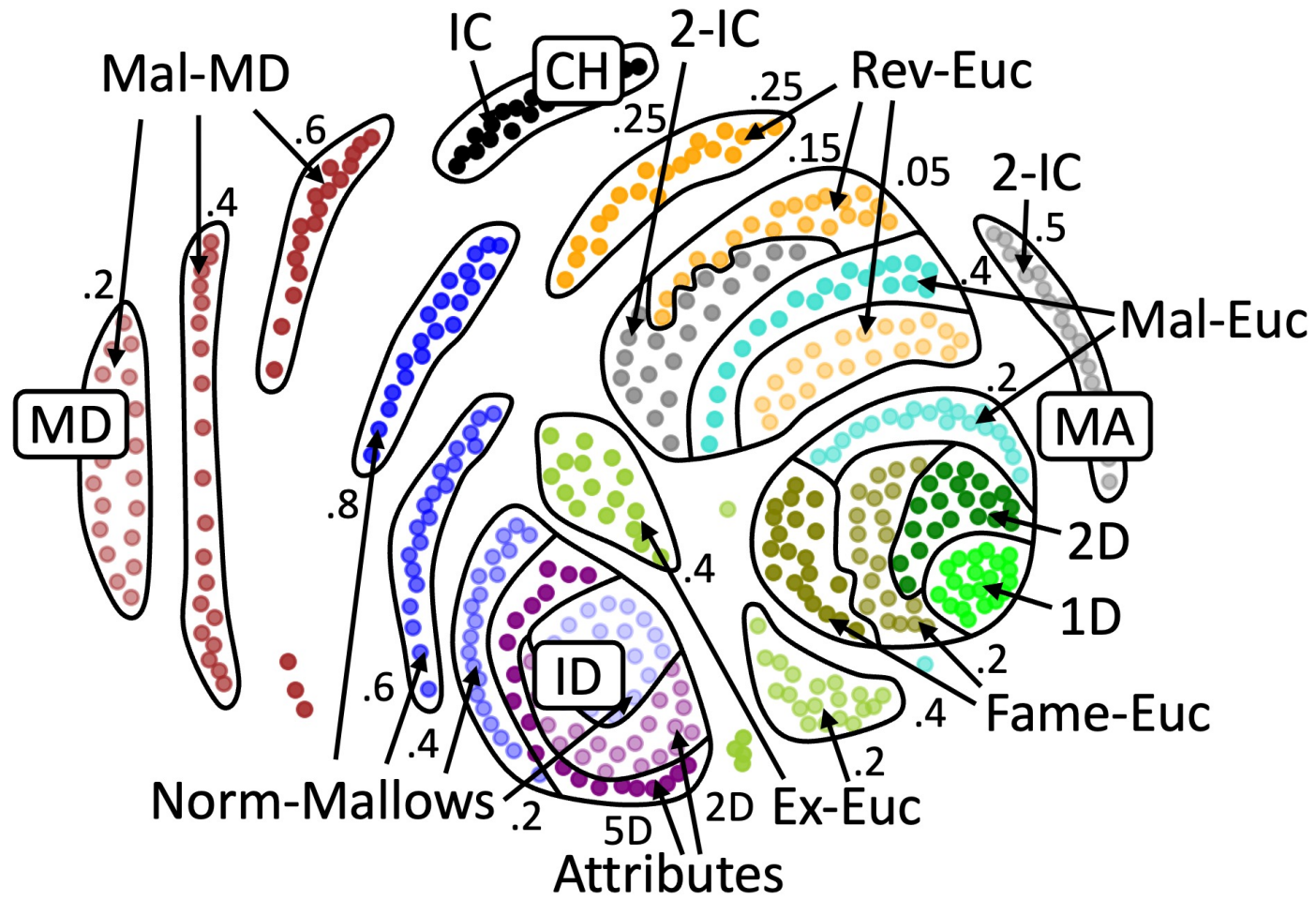


# Step 2: Generating Instances

460 instances generated from 10 statistical cultures (4 known) from:

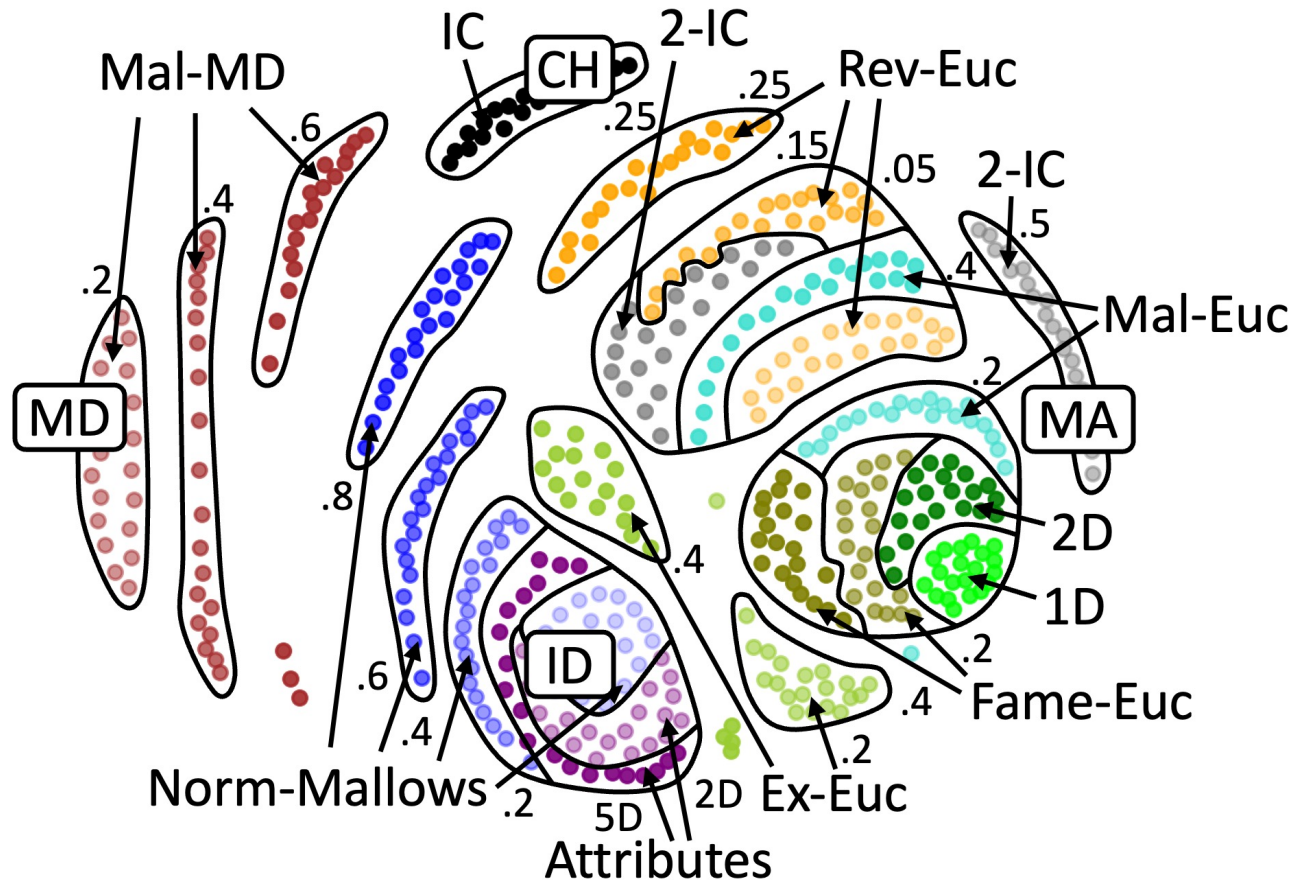
- **Impartial Culture** Agents draw preferences uniformly at random from set of all possible preferences.
- **Mallows** There is a central order  $v^*$ . Probability of sampling preference order  $v$  is proportional to  $\varphi^{\text{swap}(v,v^*)}$ .
- **Attributes** Different objective evaluation criteria but agents assign different importance to them.
- **Euclidean** Agents are points on line / in square and rank other agents by increasing distance.
- **Reverse-Euclidean** Like Euclidean but some fraction of agents rank by decreasing distance.
- **Fame-Euclidean** Like Euclidean but some agents are generally more attractive.

# Step 3: Drawing the Map



Computed using variant of forced-directed Kamada-Kawai algorithm

# Step 4: Understanding the Map



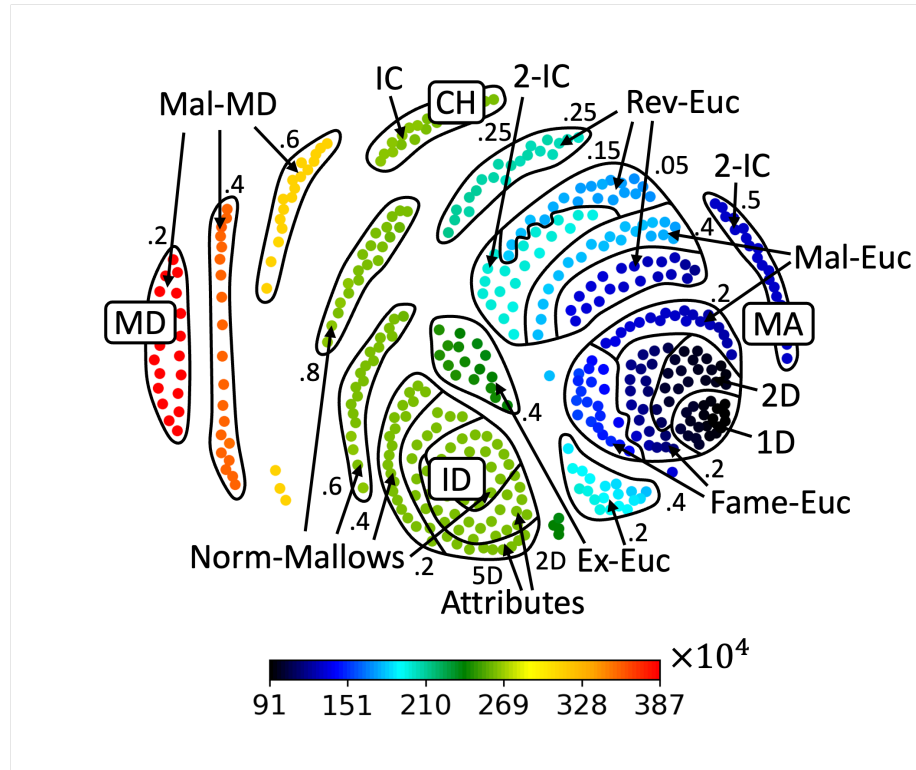
## Extreme Matrices

1. **Identity (ID)** All agents have the same preferences (master list).
2. **Mutual Agreement (MA)** Agents rank each other in same position.
3. **Mutual Disagreement (MD)** Evaluations are diametric: a ranks b in position  $i-b$  ranks a in position  $n-i+1$ .
4. **Chaos (CH)** "Chaotic" matrix.



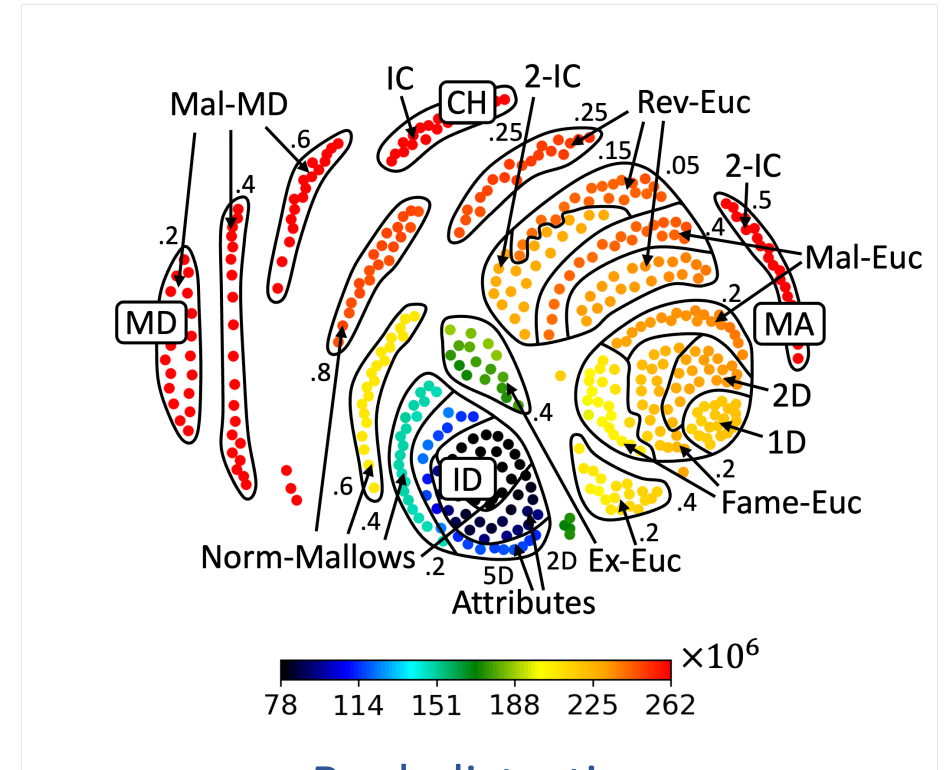
# Step 4: Understanding the Map II

## Meaning of Axes



### Mutuality value

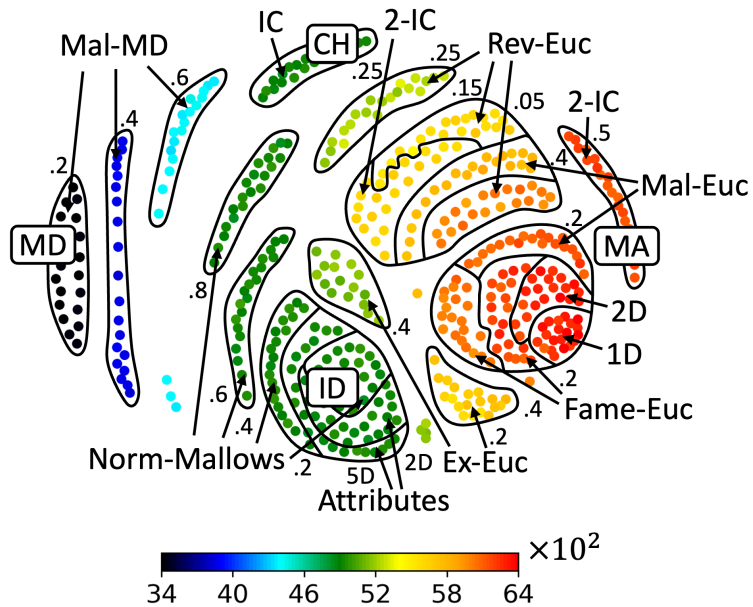
total difference between mutual evaluations of agent pairs



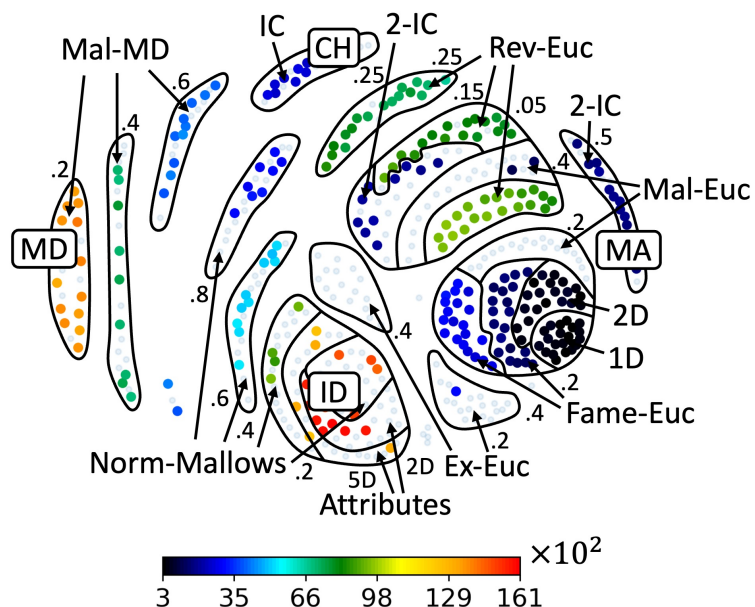
### Rank distortion

for each agent we sum up the absolute difference between all pairs of entries in MA vector

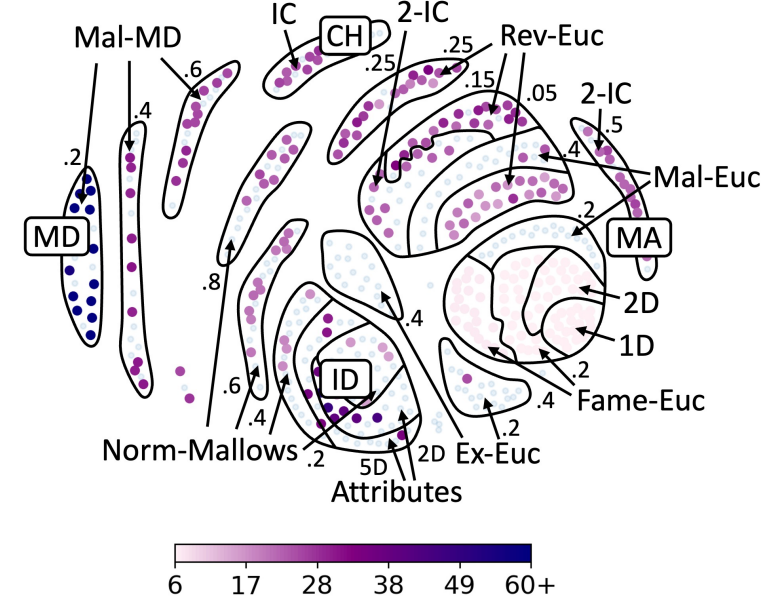
# Step 5: Using the Map



Average number of blocking pairs for perfect matching



Minimum summed rank of a stable matching



Running time of ILP for summed rank minimal matching

# Conclusion

## Take-aways

- General approach for maps applicable beyond voting including "tricks":
    - Aggregate representation
    - Force-directed algorithms
    - Give meaning to axes and regions on the map (plus compass points)
  - Instances from one statistical culture placed close to each other and exhibit similar performance in experiments.
- Usage of multi-source data crucial.

Do  
More  
Experiments!

Please...