Single-Peaked Preferences: Learning Axes from Samples

Sonja Kraiczy, Edith Elkind
University of Oxford

In a country far, far away…

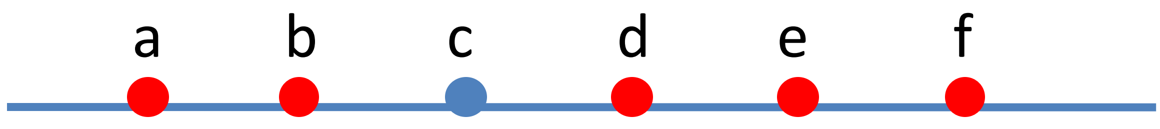
- Estonian Parliament 2021:
  - *Sotsiaaldemokraatlik Erakond*
  - *Eesti Reformierakond*
  - *Isamaa*
  - *Eesti Konservatiivne Rahvaerakond*
  - *Eesti Keskerakond*
- Can you order these parties from left to right on the political spectrum?

How many voters do we need to ask?

- Estonian Parliament 2021:
  - *Social Democratic Party (SDE)*
  - *Estonian Reform Party (Reform)*
  - *Pro Patria (PP)*
  - *Conservative People's Party of Estonia (Con)*
  - *Estonian Centre Party (Centre)*
- v_1: PP > Con > Reform > Centre > SDE
- v_2: Reform > Centre > PP > SDE > Con
- v_3: Centre > SDE > Reform > PP > Con

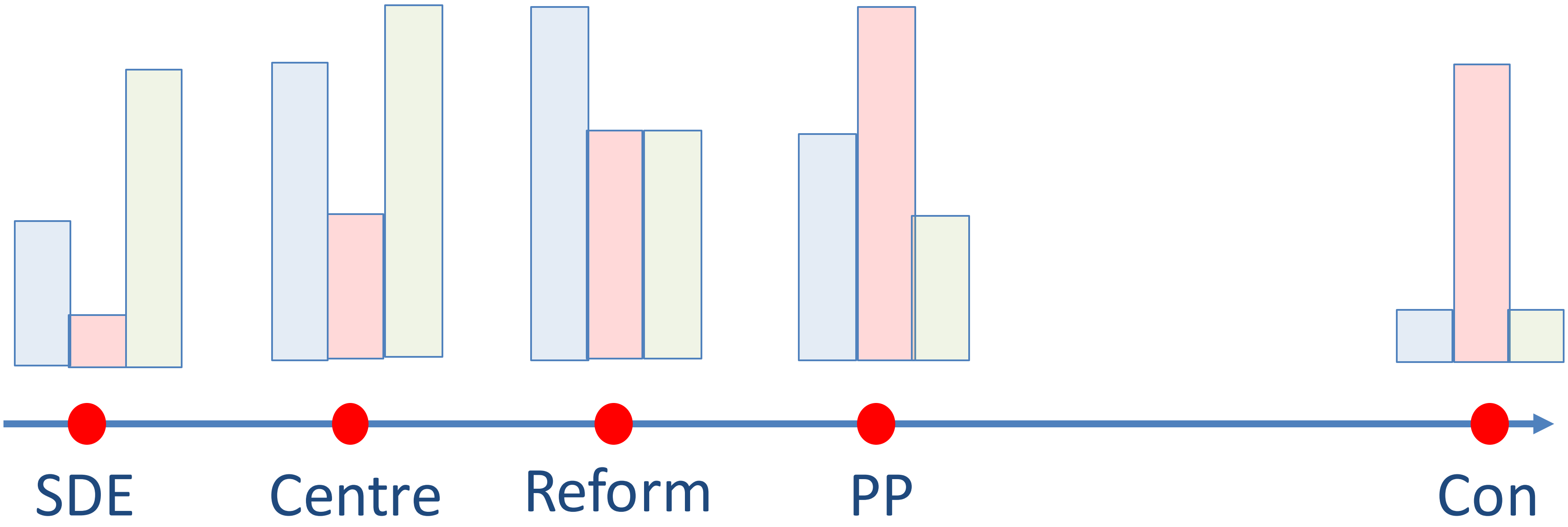
Single-peaked preferences

- <u>Setup</u>:
  - a set of m candidates C
  - each voter ranks candidates from best to worst
  - $top(v)$: most preferred candidate of voter v
- <u>Definition</u>: a vote v is single-peaked (SP) wrt an ordering $<$ of candidates (axis) if it holds that:
  - if $top(v) < d < e$, v prefers d to e
  - if $a < b < top(v)$, v prefers b to a



Example: Estonian Parliament

- v_1: PP > Con > Reform > Centre > SDE
- v_2: Reform > Centre > PP > SDE > Con
- v_3: Centre > SDE > Reform > PP > Con



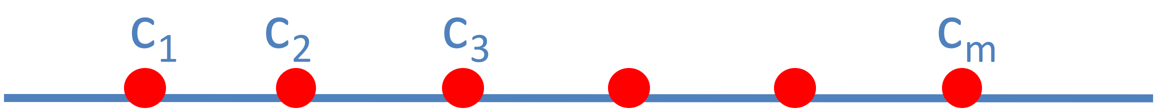
Example: Temperature

- Perfect water temperature?

+16 +20 +23 +25 +27 +30

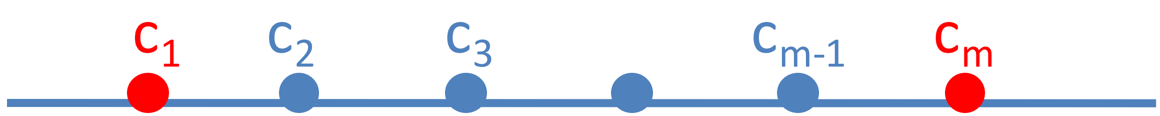
Challenge

- A population of voters, with preferences single-peaked on an axis < over C
  - assume $c_1 < c_2 < c_3 < \dots < c_m$
- We sample a random voter and ask her to report her ranking
- How many samples do we need to uniquely identify the axis (up to a swap) w.p. $1 - \delta$?
  - the answer may depend on the distribution



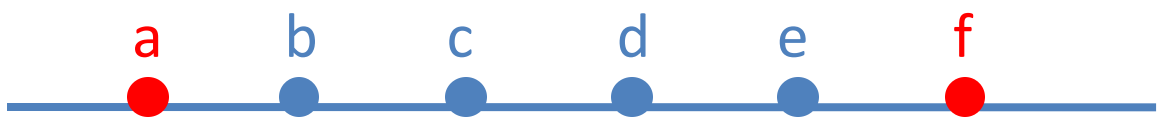
Warm-up

- Observation 1: each voter ranks c_1 or c_m last
- Observation 2: there are 2^{m-1} votes s.p. on $<$
- Best case: two votes may be enough
  - $c_1 c_2 .... c_{m-1}\, c_m$, $c_m c_{m-1} ... c_2\, c_1$
- Worst case: 2^{m-2} votes may not be enough
  - there are 2^{m-3} votes over $C \setminus \{c_1, c_m\}$ that are s.p. on $c_2 < .... < c_{m-1}$
  - to each such vote, we can append $c_1 c_m$ or $c_m c_1$
  - from these votes, we cannot decide if the axis is $c_1 < c_2 < ... < c_{m-1} < c_m$ or $c_m < c_2 < ... < c_{m-1} < c_1$



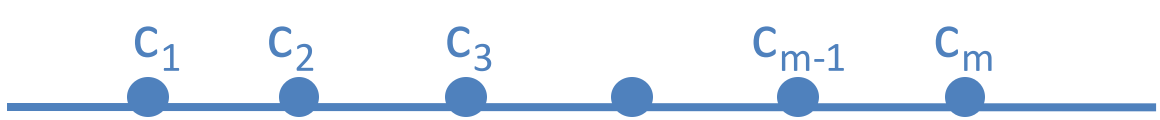
Understanding the worst case

- How do we distinguish between
  a < b < c < d < e < f and f < b < c < d < e < a?
- If all votes rank {a, f} in the last two positions, we cannot
- But suppose we have votes …a, …f, and …fba
  - i.e., in our set of votes there is no “cut” between positions 1, 2, 3, 4 and 5, 6
  - then b is “glued” to a
- <u>Lemma [DF’94]</u>: we can identify the axis iff there no cut between positions 1, …, j and j+1, …, m for any j



Average case: uniform distribution

- $\mathcal{P}(<)$: the set of all votes s.p. on $<$ (size of $\mathcal{P}(<)$: 2^{m-1})
- $\mathcal{U}(<)$: uniform distribution over $\mathcal{P}(<)$
- How do we sample votes from $\mathcal{U}(<)$?
- Bottom up:
  - last candidate is c_1 or c_m w.p. 1/2
  - if c_1 is last, then 2nd last candidate is c_2 or c_m, w.p. 1/2
  - if c_m is last, then 2nd last candidate is c_1 or c_{m-1}, w.p. 1/2
  - etc.
  - m-1 binary choices (left or right)
- Note: both c_1… and c_m… are exponentially unlikely

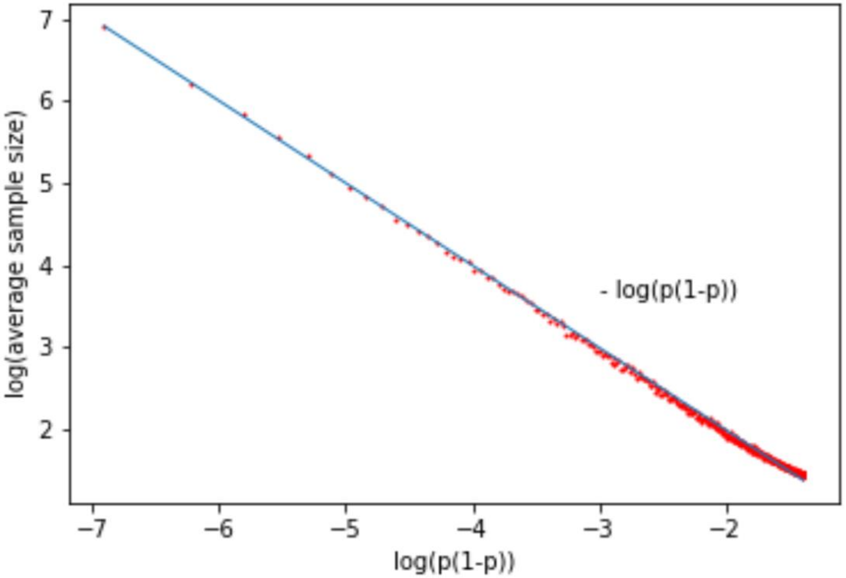


Main result

- Theorem: For any $\delta > 0$, we can identify the axis $<$ using $O(\log 1/\delta)$ samples from $\mathcal{U}(<)$ w.p. at least $1 - \delta$
- Proof sketch:
  - a vote from $\mathcal{U}(<)$ = a uniform random walk in 1D
  - a sample of k votes admits a cut iff the respective k random walks all meet at the same point
  - Lemma (hard): with constant probability, four random walks never meet
  - Algorithm: draw $O(\log 1/\delta)$ lots of 4 votes each
- Empirically: 5 votes always suffice

Skewed distribution

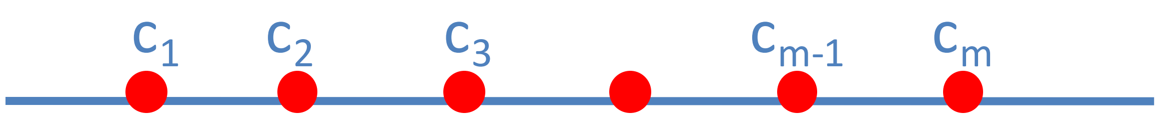
- $\mathcal{U}_p(<)$: skewed uniform distribution (L w.p. p, R w.p. 1-p)
- <u>Theorem</u>: For any $\delta > 0$, and any $0 < p < 1$ we can identify the axis $<$ using $O(\log 1/\delta)$ samples from $\mathcal{U}_p(<)$ w.p. at least $1 - \delta$
  - skewed random walk
- Dependence on p (empirical): $1/(p(1-p))$



Random peak distribution

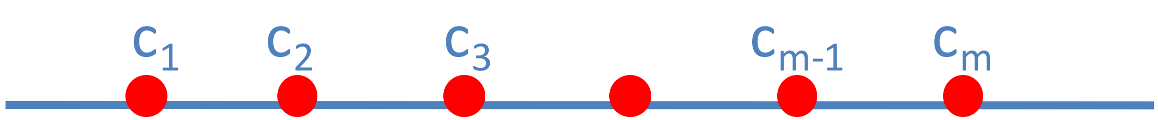
- Uniformly **random peak** distribution $\mathcal{RP}(<)$:
  - generate the vote top to bottom
  - each candidate is **equally likely** to be ranked **first**
  - then move left or right on $<$ w.p. $1/2$

- <u>Theorem</u>: For any $\delta > 0$,
  we can **identify** the axis $<$ using $O(\log 1/\delta)$ samples from $\mathcal{RP}(<)$ w.p. at least $1 - \delta$



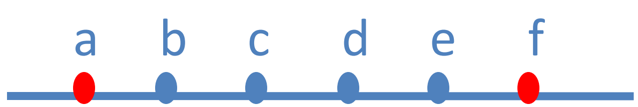
Sampling pairwise comparisons

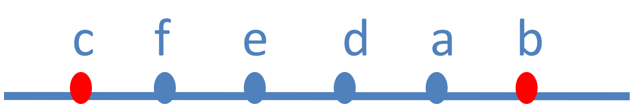
- Suppose we ask each voter about one pair of candidates
  - impossible to learn with certainty
  - but possible to learn w.h.p., both for $\mathcal{U}(<)$ and $\mathcal{RP}(<)$
- <u>Theorem</u>: Suppose that we sample pairwise comparisons from $\mathcal{U}(<)$. For any $\delta > 0$, we can learn $<$ w.p. at least $1 - \delta$ using $O(m^3 \log(m/\delta))$ samples (and $O(m^4 \log(m/\delta))$ for $\mathcal{RP}(<)$)
  - order candidates by number of wins
  - we obtain $\{c_1, c_m\} < \{c_2, c_{m-1}\} < \{c_3, c_{m-2}\} < \ldots$
  - break "ties" moving from the center outwards



Two axes

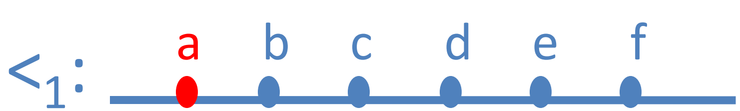
- Suppose there are two axes on C ($<_1$ and $<_2$)
- Half of the votes come from $U(<_1)$, half from $U(<_2)$
  - or, 80% and 20%
- $<_1$: political left-right axis
- $<_2$: conflict in Ukraine
- Can we learn $<_2$ if $<_1$ is known?
- Can we learn $<_1$ and $<_2$ if both are unknown?

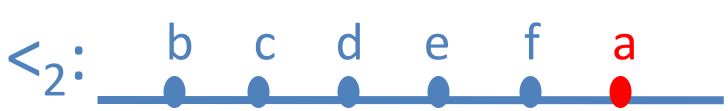




One known, one unknown axis

- Suppose $<_1$ is known, $<_2$ is unknown
- If all endpoints of $<_1$ and $<_2$ are distinct, we can discard votes s.p. on $<_1$
- But if $<_1$ and $<_2$ share one endpoint, we may be unable to identify $<_2$ even given $\mathcal{P}(<_1) \cup \mathcal{P}(<_2)$
- Which votes are s.p. on $<_2$, but not on $<_1$?
  - exactly the votes in $\mathcal{P}(<_2)$ that rank b last
  - but such votes are also compatible with $c < d < e < f < a < b$





Two axes: results

- Theorem: For any $\delta > 0$, given $<_1$, we can identify $<_2$ using $O(m \log m/\delta)$ samples from $1/2 \times \mathcal{U}(<_1) + 1/2 \times \mathcal{U}(<_2)$ w.p. at least $1 - \delta$ as long as $<_1$, $<_2$ have 0 or 2 common endpoints

- Theorem: For any $\delta > 0$, if $m > 4$, we can identify the pair of axes $<_1$, $<_2$ using $O(m \log m/\delta)$ samples from $1/2 \times \mathcal{U}(<_1) + 1/2 \times \mathcal{U}(<_2)$ w.p. at least $1 - \delta$ as long as $<_1$, $<_2$ have 0 or 2 common endpoints

Future work

- Can we characterize distributions that enable axis identification?
- Single-peakedness on richer graphs?
- Learning partial information about the axis?