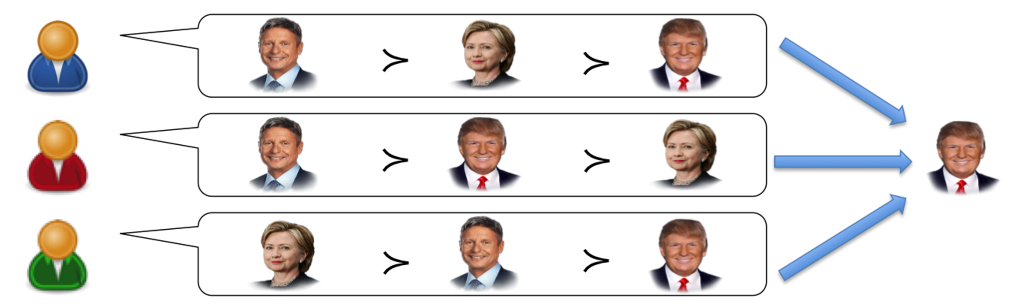
LEARNING MIXTURES OF RANDOM UTILITY MODELS

Zhibing Zhao, Tristan Villamil and Lirong Xia
Computer Science Department, Rensselaer Polytechnic Institute

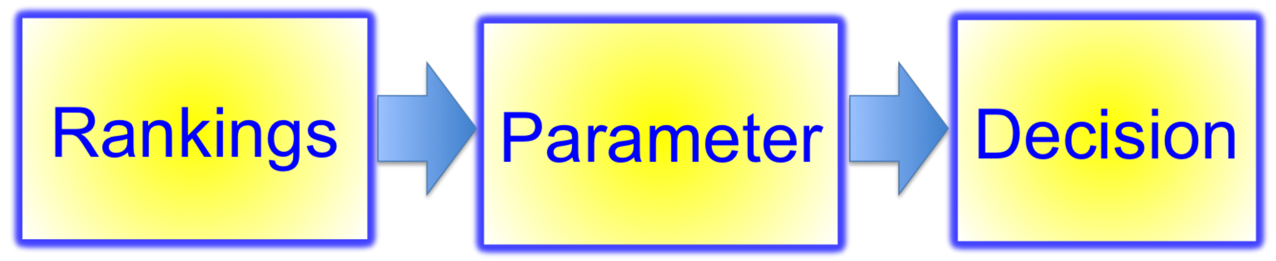
RANK AGGREGATION

- Alternatives: {Donald Trump, Hillary Clinton, Gary Johnson}.
- Mechanism: *plurality* rule.
- Decision: Donald Trump.



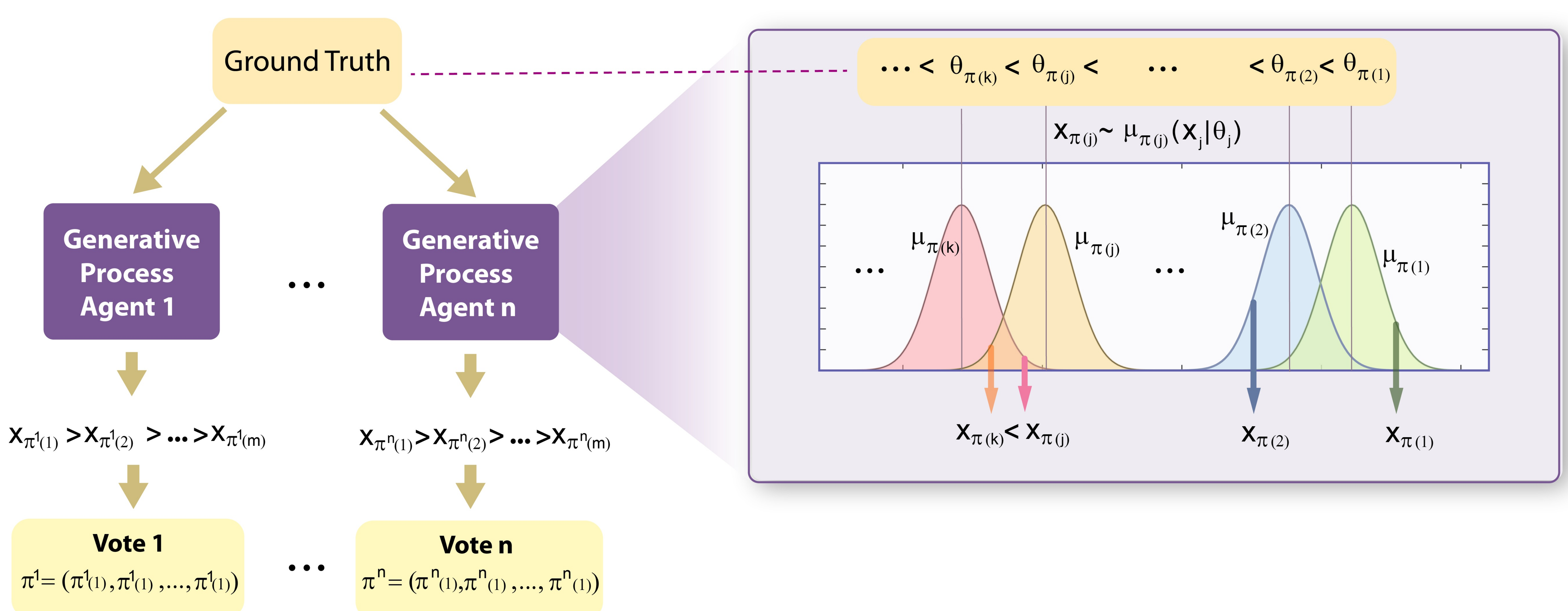
RANKING MODELS

- Set of alternatives: $\mathcal{A} = \{a_1, a_2, \ldots, a_m\}$.
- Parameter space: Θ.
- Sample space: i.i.d. rankings over $\mathcal{A}$.
- Distributions: $\{\Pr_\theta : \theta \in \Theta\}$.



RANDOM UTILITY MODELS (RUMS)

- Sample a random utility for each alternative independently.
- Rank the alternatives w.r.t. these utilities.
- The Plackett-Luce model: all utility distributions are Gumbel distributions.



MIXTURES OF RANDOM UTILITY MODELS (k-RUMS)

- Parameter: mixing coefficients $(\alpha_1, \alpha_2, \ldots, \alpha_k)$; RUM components: RUM_1, RUM_2, …, RUM_k.
- Sample a component r w.r.t. mixing coefficients.
- Sample a ranking from component r.

MOTIVATION: MODEL FITNESS ON PREFLIB DATA

Given the number of parameters d, the number of rankings n and the likelihood of the estimate L:

- $\text{AIC} = 2d - 2\ln(L)$.
- $\text{AICc} = \text{AIC} + \frac{2d(d+1)}{n-d-1}$.
- $\text{BIC} = d\ln(n) - 2\ln(L)$.

Observation:

k-RUM $\succ$ k-PL $\succ$ RUM $\succ$ PL,

where $A \succ B$ means that the number of datasets where A beats B is more than that where B beats A.

THEORETICAL RESULTS: IDENTIFIABILITY

Def: For all $\vec{\theta}_1, \vec{\theta}_2 \in \Theta$, $\Pr_{\vec{\theta}_1} = \Pr_{\vec{\theta}_1} \Rightarrow \vec{\theta}_1 = \vec{\theta}_2$.

Theorem 1. Let $\mathcal{M}$ be any symmetric RUM from the location family. When $m \le 2k-1$, k-$\text{RUM}_\mathcal{M}$ over m alternatives is **non-identifiable**.

Figure 1: The label switching problem

Theorem 2. For any $\text{RUM}_\mathcal{M}$ where all utility distributions have support $(-\infty, \infty)$, when $m \ge \max\{4k-2, 6\}$, k-$\text{RUM}_\mathcal{M}$ over m alternatives is **generically identifiable**.

FIRST ALGORITHMS TO LEARN k-RUM

Generalizd Method of Moments (GMM).

1. Choose q events: all pairwise comparisons and selected triple-wise comparisons.
2. Compute the empirical probabilities $b_1, \ldots, b_q$.
3. Find the parameter that minimize $\sum_{i=1}^{q}(b_i - p_i(\vec{\theta}))^2$, where $p_i(\vec{\theta})$ is the probability computed from the model.

The Sandwich Algorithm (GMM-E-GMM).

1. Run GMM for a good starting point.
2. Run E-GMM to improve accuracy.

E-GMM.

1. E-step: compute the probabilities that a ranking belongs to each component.
2. M-step: compute the parameter of each component using the GMM algorithm by Azari Soufiani (2014).

Figure 2: The sandwich algorithm

EXPERIMENTS

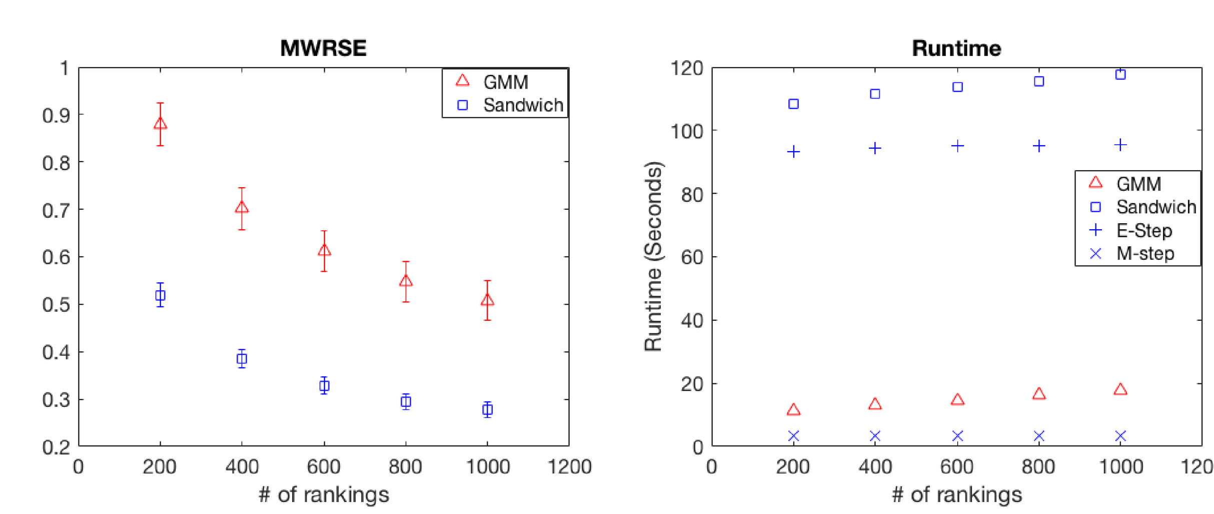


Figure 3: 2-RUM over 6 alternatives

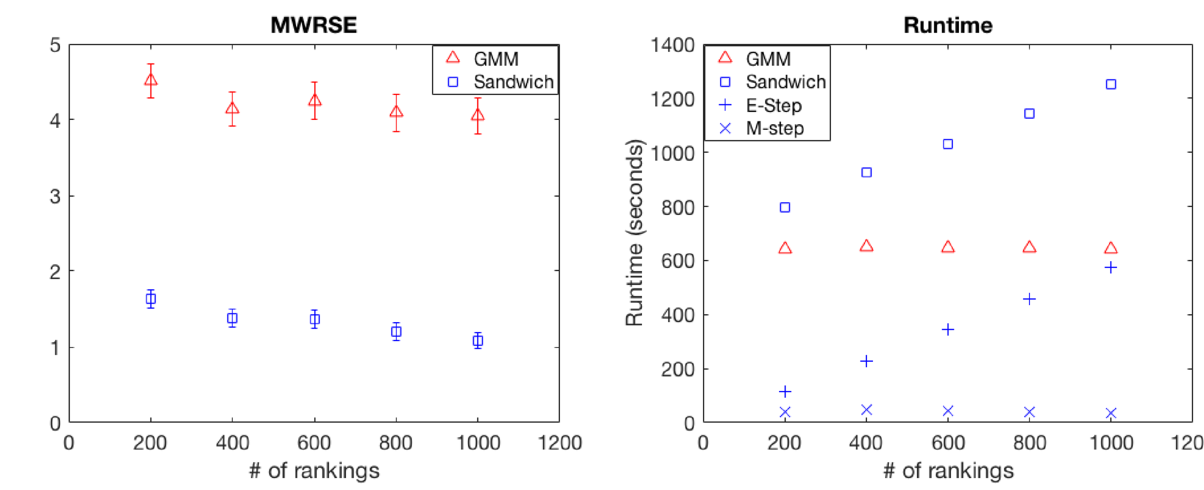


Figure 4: 4-RUM over 15 alternatives

COMPARISON WITH k-PLS

	k-PLs	k-RUMs
Closed-form likelihood?	Yes	No
Proving identifiability?	Hard	Harder

REFERENCES

Hossein Azari Soufiani, David C. Parkes, and Lirong Xia, "Computing Parametric Ranking Models via Rank-Breaking". ICML-14.

Zhibing Zhao, Peter Piech, and Lirong Xia, "Learning Mixtures of Plackett-Luce Models". ICML-16