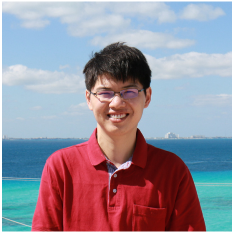
Top-K Ranking with a Monotone Adversary

Yuepeng Yang
Department of Statistics, UChicago

COLT, July. 2024



Cong Ma

UChicago Stat

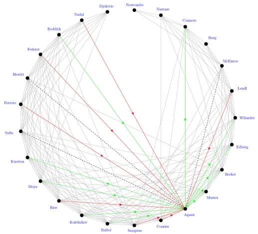
Antares Chen

UChicago CS

Lorenzo Orecchia

UChicago CS

Ranking from pairwise comparisons



pairwise comparisons for ranking top tennis players

figure credit: Bozóki, Csató, Temesi

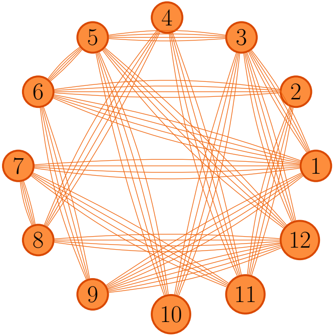
Bradley-Terry-Luce model: Assign latent score to each of n items $\boldsymbol{\theta}^\star = [\theta_1^\star, \cdots, \theta_n^\star]$ with

$$\mathbb{P}\left\{\text{item } j \text{ beats item } i\right\} = \frac{e^{\theta_i^\star}}{e^{\theta_i^\star} + e^{\theta_j^\star}}$$

Goal: identify the set of top-K items under minimal sample size

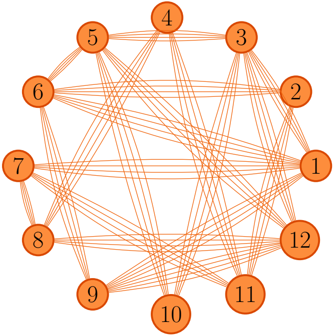
Sampling model

Sampling on comparison graph $\mathcal{G} = ([n], \mathcal{E})$: i, j are compared iff $(i, j) \in \mathcal{E}$



Sampling model

Sampling on comparison graph $\mathcal{G} = ([n], \mathcal{E})$: i, j are compared iff $(i, j) \in \mathcal{E}$



For each $(i, j) \in \mathcal{E}$, obtain L paired comparisons

$$y_{i,j}^{(l)} \overset{\text{ind.}}{=} \begin{cases} 1, & \text{with prob. } \frac{e^{\theta_j^\star}}{e^{\theta_i^\star} + e^{\theta_j^\star}} \\ 0, & \text{else} \end{cases} \qquad 1 \leq l \leq L$$

Prior art: MLE works for uniform sampling

- Uniform comparison graph: Erdős–Rényi graph $\mathcal{G}_{\mathrm{ER}} \sim \mathcal{G}(n, p)$

Theorem 1 (CFMW, AoS '19; CGZ, AoS '22)

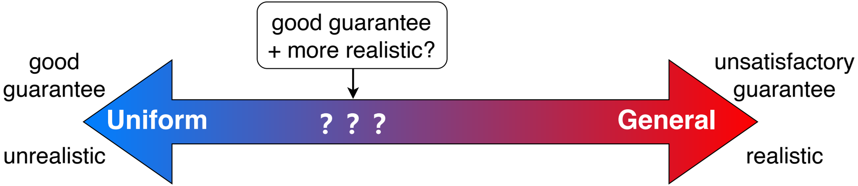
When $p \gtrsim \frac{\log n}{n}$, regularized or unregularized MLE achieves optimal sample complexity for top-K ranking

Prior art: MLE works for uniform sampling

- Uniform comparison graph: Erdős–Rényi graph $\mathcal{G}_{\mathrm{ER}} \sim \mathcal{G}(n,p)$

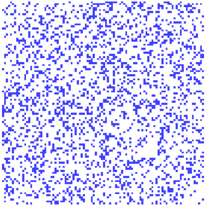
Theorem 1 (CFMW, AoS '19; CGZ, AoS '22)

When $p \gtrsim \frac{\log n}{n}$, regularized or unregularized MLE achieves optimal sample complexity for top-K ranking

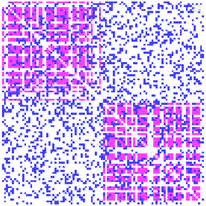


Top-K ranking with a monotone adversary

—aka semi-random adversary



$\mathcal{G}_{\mathrm{ER}} = ([n], \mathcal{E}_{\mathrm{ER}})$

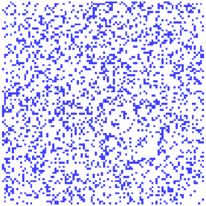


$\mathcal{G}_{\mathrm{SR}} = ([n], \mathcal{E}_{\mathrm{SR}})$ with added edges

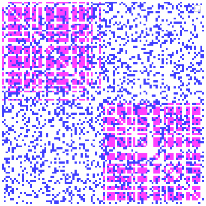
Special case: non-uniform sampling $(i, j) \in \mathcal{E}$ with probability $p_{ij} \geq p$

Top-K ranking with a monotone adversary

—aka semi-random adversary



$\mathcal{G}_{\mathrm{ER}} = ([n], \mathcal{E}_{\mathrm{ER}})$



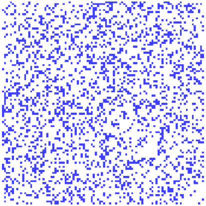
$\mathcal{G}_{\mathrm{SR}} = ([n], \mathcal{E}_{\mathrm{SR}})$ with added edges

Special case: non-uniform sampling $(i, j) \in \mathcal{E}$ with probability $p_{ij} \geq p$

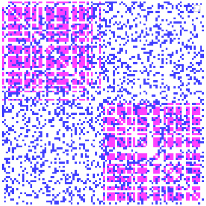
Can we identify top-K items under monotone adversary?

Top-K ranking with a monotone adversary

—aka semi-random adversary



$\mathcal{G}_{\mathrm{ER}} = ([n], \mathcal{E}_{\mathrm{ER}})$



$\mathcal{G}_{\mathrm{SR}} = ([n], \mathcal{E}_{\mathrm{SR}})$ with added edges

Special case: non-uniform sampling $(i, j) \in \mathcal{E}$ with probability $p_{ij} \geq p$

Can we identify top-K items under monotone adversary? Not clear!

Intuition: mimicking oracle

If we have oracle knowledge of $\mathcal{E}_{\mathrm{ER}}$

Intuition: mimicking oracle

If we have oracle knowledge of $\mathcal{E}_{\mathrm{ER}}$

$$\Downarrow$$

We would run MLE using edges in $\mathcal{E}_{\mathrm{ER}}$

Intuition: mimicking oracle

If we have oracle knowledge of $\mathcal{E}_{\mathrm{ER}}$

$$\Downarrow$$

We would run MLE using edges in $\mathcal{E}_{\mathrm{ER}}$

$$\Downarrow$$

Equivalent to weighted MLE with unit weight on $\mathcal{E}_{\mathrm{ER}}$

Intuition: mimicking oracle

If we have oracle knowledge of $\mathcal{E}_{\mathrm{ER}}$

$$\Downarrow$$

We would run MLE using edges in $\mathcal{E}_{\mathrm{ER}}$

$$\Downarrow$$

Equivalent to weighted MLE with unit weight on $\mathcal{E}_{\mathrm{ER}}$

$$\Downarrow$$

We don't know $\mathcal{E}_{\mathrm{ER}}$

Intuition: mimicking oracle

If we have oracle knowledge of $\mathcal{E}_{\mathrm{ER}}$

$$\Downarrow$$

We would run MLE using edges in $\mathcal{E}_{\mathrm{ER}}$

$$\Downarrow$$

Equivalent to weighted MLE with unit weight on $\mathcal{E}_{\mathrm{ER}}$

$$\Downarrow$$

We don't know $\mathcal{E}_{\mathrm{ER}}$

$$\Downarrow$$

Can we find weights that mimic the above?

Optimal control of entrywise error

Theorem 2 (Yang, Chen, Oreccia, Ma, 2024)

When $p \gtrsim \frac{\log(n)}{n}$ *and* $npL \gtrsim \log^3(n)$*, with some proper reweighting, weighted MLE* $\widehat{\boldsymbol{\theta}}_w$ *obeys*

$$\|\widehat{\boldsymbol{\theta}}_w - \boldsymbol{\theta}^\star\|_\infty \lesssim \sqrt{\frac{\log(n)}{npL}}$$

will come back later to explain what is proper reweighting

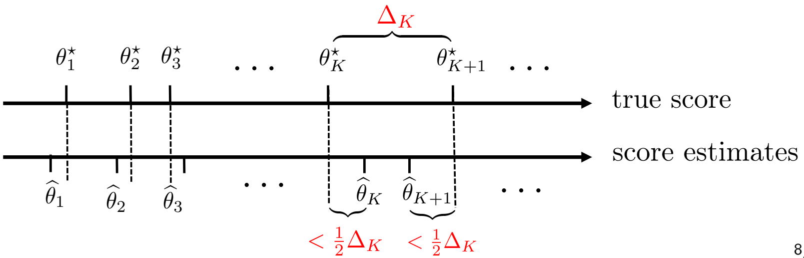
Optimal control of entrywise error

Theorem 2 (Yang, Chen, Oreccia, Ma, 2024)

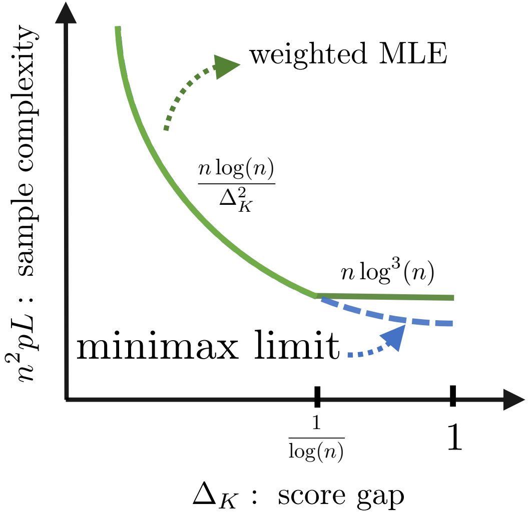
When $p \gtrsim \frac{\log(n)}{n}$ *and* $npL \gtrsim \log^3(n)$*, with some proper reweighting, weighted MLE* $\widehat{\boldsymbol{\theta}}_w$ *obeys*

$$\|\widehat{\boldsymbol{\theta}}_w - \boldsymbol{\theta}^\star\|_\infty \lesssim \sqrt{\frac{\log(n)}{npL}}$$

will come back later to explain what is proper reweighting



Near-optimal sample complexity



A few words about analysis

- ℓ_2 loss vs. ℓ_∞ loss
- Prior analysis e.g., leave-one-out analysis relies heavily on independence of edges, and also is not transparent in terms of graph properties

Master theorem for weighted MLE

- $w_{\max} := \max_{i,j} w_{ij}$ be the maximum weight
- $d_{\max} := \max_{i\in[n]} \sum_{j:j\neq i} w_{ij}$ be the maximum (weighted) degree
- Weighted graph Laplacian

$$\boldsymbol{L}_w := \sum_{(i,j):i>j} w_{ij}(\boldsymbol{e}_i - \boldsymbol{e}_j)(\boldsymbol{e}_i - \boldsymbol{e}_j)^\top$$

Theorem 3 (Yang, Chen, Oreccia, Ma, 2024)

When graph is connected, as long as

$$L \gg \frac{w_{\max}(d_{\max})^4 \log^3(n)}{(\lambda_{n-1}(\boldsymbol{L}_w))^5},$$

with high probability, we have

$$\|\widehat{\boldsymbol{\theta}}_w - \boldsymbol{\theta}^\star\|_\infty \lesssim \sqrt{\frac{w_{\max}\log(n)}{\lambda_{n-1}(\boldsymbol{L}_w)L}}$$

Master theorem for weighted MLE

- LOO-free analysis that allows dependent edges
- Depends explicitly only on graph properties
- Applicable to other settings:
  "Random pairing MLE for estimation of item parameters in Rasch model", Yang and Ma, 2024

Optimization-based reweighting

- Master theorem motivates us to consider following SDP

$$\begin{aligned} \max_{\boldsymbol{w}} \quad & \lambda_{n-1}(\boldsymbol{L}_w) \\ \text{s.t.} \quad & \sum_i w_{ij} \le 2np \quad \text{for all } j \\ & 0 \le w_{ij} \le 1 \quad \text{for all } i, j \end{aligned}$$

- Since unit weights on $\mathcal{E}_{\mathrm{ER}}$ is feasible, we know the maximizer is at least as good as that for Erdős–Rényi graph

- Approximately solvable in near-linear time

Concluding remarks

Weighted MLE is statistically and computationally efficient for top-K ranking with monotone adversary

- Novel analysis of weighted MLE with general weights
- Efficient algorithm to approximately solve SDP-based reweighting

Future directions:

- Is weighted MLE necessary?
- Stronger adversary?

Papers:

- Y. Yang, A. Chen, L. Orecchia, C. Ma, "Top-K ranking with a monotone adversary," COLT, 2024
- Y. Yang, C. Ma, "Random pairing MLE for estimation of item parameters in Rasch model" arXiv, 2024