

Estimating α -Rank by Maximising Information Gain

Tabish Rashid^{*}

Cheng Zhang⁺

Kamil Ciosek⁺

^{*}University of Oxford

⁺Microsoft Research Cambridge

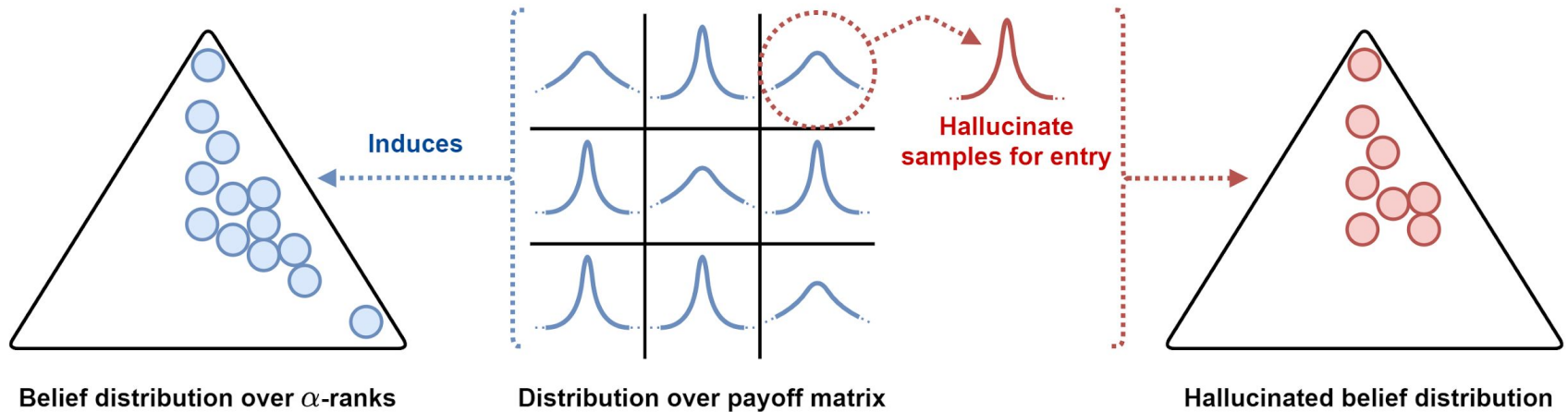


UNIVERSITY OF
OXFORD



Microsoft

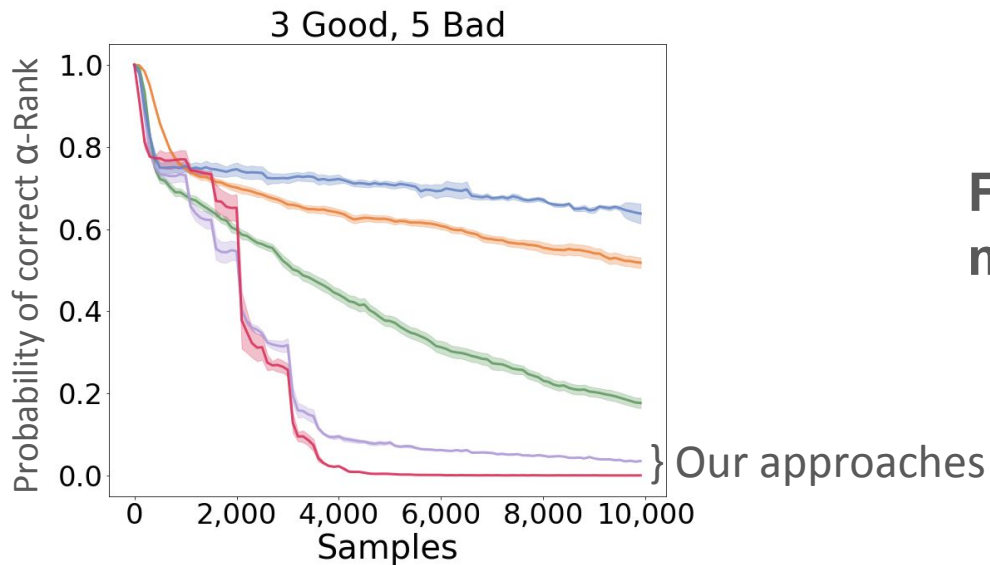
- Our goal is to estimate α -Rank by querying entries of the payoff matrix using as few samples as possible.
- Adopt a Bayesian approach:



- Pick entry that maximises the **Information Gain** about our hallucinated belief distribution and its observed payoff.

- **Theoretically**, we bound:
 - regret attained by greedily maximizing information gain **on the payoffs**.
 - regret attained by sampling sequence of strategy profiles maximizing information gain **on the α -rank**.
- **Empirically**, we show maximising information gain in this way outperforms the baselines on synthetic games.

We also introduce a Wasserstein-based objective (see paper for more details!)



Feedback would be very much appreciated!