We thank the reviewers for their positive feedback and will make the suggested improvements in the paper. Given the extra page allowed for accepted papers, we will include the discussions and experiments resulting from this feedback. We kindly hope the reviewers take this into account when finalizing their scores. Responses below are ordered w.r.t. the reviewer comments (e.g., **R2.3** refers to Reviewer 2's 3rd comment).

R1.1 **Figure 6.2 presentation:** Agreed, Fig. 6.2's presentation will be improved. We thank the reviewer for the suggestions. We will move extraneous plots from Fig. 6.2 to the appendix and increase the plot sizes for readability. We will do the same for other figures where applicable. **R1.2** **Sec. 5, move CSSP details to the appendix:** Agreed, per the reviewer's suggestion, we will expound on the higher-level details & intuitions in the section itself, and move the technical details to the appendix. **R1.3** **Theorem 4.1 & sampling strategy comparisons:** Indeed, Theorem 4.1 is agnostic of the chosen sampling scheme; while we note this in the text preceding the theorem, we will update the theorem statement itself to make this property explicit. Note that Fig 6.2 visualizes empirical differences between the different sampling strategies. Overall, CP-UCB seems to attain best empirical performance by a small margin, which can be explained due to the Bernoulli meta-game outcomes (which are win/loss in nature). **R1.4 & 1.5** **Fig. 6.3:** Good point, we will correct the wording regarding the positive correlation of errors and tolerances here. Regarding the higher variance of Poker results, let us consider the distribution of payoffs gaps, which play a key role in determining response graph reconstruction errors. Let $\Delta(s, \sigma) = |\mathbf{M}^k(s) - \mathbf{M}^k(\sigma)|$, the payoff difference corresponding to the edge of the response graph where player k deviates, causing a transition between strategy profiles $s, \sigma \in S$ (see paper lines 91–92 for precise definitions). Figure R1 plots the ground truth distribution of these gaps for all response graph edges in Soccer & Poker. The higher ranking variance may be explained by these gaps tending to be more heavily distributed near 0 for Poker. **R1.6** **Bernoulli game payoffs:** We agree that it would be interesting to evaluate bound tightness, and plan to investigate this in future work.

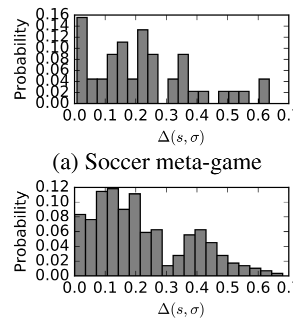

Figure R1: Distribution of payoff gaps $\Delta(s, \sigma)$.

R2.1 & 2.2 **Alternative approaches, e.g., collaborative filtering:** We thank the reviewer for the interesting and important question regarding alternative approaches. The pairing of bandit algorithms and α-Rank is a natural means of computing rankings in settings where, e.g., one has a limited budget for adaptively sampling match outcomes. Our use of bandit algorithms also leads to analysis which is flexible enough to be able to deal with K-player general-sum games. However, approaches such as collaborative filtering may indeed fare well in their own right. We provide discussion of one such application below, specifically for the case of two-player win-loss games.

For such games, the meta-payoff table is given by a matrix $\mathbf{M}$ with all entries lying in $(0, 1)$ (encoding loss as payoff 0 and win as payoff 1). Taking a matrix completion approach, we might attempt to reconstruct a low-rank approximation of the payoff table from an incomplete list of (possible noisy) payoffs, and then run α-Rank on the reconstructed payoffs. Possible candidates for the low-rank structure include: (i) the payoff matrix itself; (ii) the *logit matrix* $\mathbf{L}_{ij} = \log(\mathbf{M}_{ij}/(1 - \mathbf{M}_{ij}))$; and (iii) the *odds matrix* $\mathbf{O}_{ij} = \exp(\mathbf{L}_{ij})$. In particular, Balduzzi et al. (2018) make an argument for the (approximate) low-rank structure of the logit matrix in many applications of interest.

Per the reviewer's suggestion, we have now conducted preliminary experiments on this in Fig. R2, implementing matrix completion calculations via Alternating Minimization (Jain et al., 2013). We compare here the resulting α-Rank errors for the three reconstruction approaches for the Soccer meta-game. We sweep across the observation rates of payoff matrix entries and the matrix rank assumed in the reconstruction. Interestingly, conducting low-rank approximation on the logits (as opposed to the odds) matrix generally yields the lowest ranking error. Overall, the bandit-based approach may be more suitable when one can afford to play all strategy profiles at least once, whereas matrix completion is perhaps more so when this is not feasible. We will append these discussions and results to the paper.

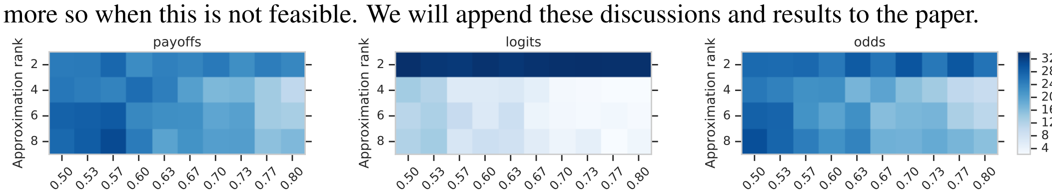

Figure R2: Ranking errors (Kendall's distance w.r.t. ground truth) from completion of, respectively, the sparse payoffs, logits, and odds matrices for Soccer dataset. 20 trials per combo of assumed matrix ranks and observation rates/density.

R3.1 We thank the reviewer for the positive feedback. Indeed, the bounds in Sec. 3 are not directly related to the final algorithm, in contrast to the bound in Sec. 4; we also plan to investigate the possibility of tightening the former bounds.

References: Balduzzi, D., Tuyls, K., Perolat, J., & Graepel, T. (2018). *Re-evaluating evaluation*.
Jain, P., Netrapalli, P., & Sanghavi, S. (2013). *Low-rank matrix completion using alternating minimization*.