Supplementary Material for
Estimating α-Rank from A Few Entries with Low Rank Matrix Completion

A. Additional Details on Algorithms

A.1. Additional Background of α-rank

Given a n-player game, where each player $i \in [n]$ has a finite set S_i of pure strategies. Let $S = \Pi_i S_i$ denote the set of joint strategies. For each tuple $s = (s_1, ..., s_n) \in S$ of pure strategies, the game specifies a joint payoffs $\boldsymbol{M}(s)$ of players. The vector of expected payoffs is denoted $\boldsymbol{M}(s) = (\boldsymbol{M}^1(s), ..., \boldsymbol{M}^n(s)) \in \mathbb{R}^n$. α-rank computes rankings following four steps: 1) construct payoff matrix for each player $\boldsymbol{M}^i, i \in [n]$; 2) construct transition matrix by Equation (2); 3) compute the stationary distribution of $\boldsymbol{C}$, as $\boldsymbol{\pi}$; 4) return the ranking of strategies according to probabilities in $\boldsymbol{\pi}$. Below is the computation of transition matrix $\boldsymbol{C}$:

$$\boldsymbol{C}_{s,\sigma} = \begin{cases} \eta \frac{1-\exp\left(-\alpha\left(\boldsymbol{M}^i(\sigma)-\boldsymbol{M}^i(s)\right)\right)}{1-\exp(-\alpha p(\boldsymbol{M}^i(\sigma)-\boldsymbol{M}^i(s)))} & \text{if } \boldsymbol{M}^i(\sigma) \neq \boldsymbol{M}^i(s) \\ \frac{\eta}{p} & \text{otherwise} \end{cases} \tag{2}$$

where the coefficient η is defined as $\eta = \left(\sum_{i=1}^{n}(|S_i|-1)\right)^{-1}$, and $\alpha > 0, p \in \mathbb{N}$ are hyperparameters. Let $C_{\sigma,\tau} = 0$ for all τ that differ from σ in more than a single player's strategy. $C_{\sigma,\sigma} = 1 - \sum_{\tau \neq \sigma} C_{\sigma,\tau}$ ensures that transition distributions are valid.

Our two-player meta-games setting is the single population case of traditional α-rank that two players have a shared pure strategies space S, and the joint strategies space is defined as $S \times S$. The payoffs of joint strategies are saved as a payoff matrix $\boldsymbol{M}$, where $\boldsymbol{M}_{ij}, \boldsymbol{M}_{ji}$ represents the payoffs of strategy S_i and strategy S_j respectively. Thus we could construct the transition matrix $\boldsymbol{C}$ between strategies in S by Equation (1) and get the ranking of strategies in S eventually.

A.2. Supporting Algorithms

Algorithm 3 gives the details of RG-UCB (Rowland et al., 2019) algorithm as a supplement of Algorithm 2. RG-UCB is composed by a sampling scheme $\mathcal{S}$ and a stopping condition $\mathcal{C}(\delta)$. It adopts Uniform-exhaustive (UE) as sampling scheme $\mathcal{S}$. At each time, it uniformly randoms a pair from all pairs need to be estimated to make a simulation. For the stopping condition $\mathcal{C}(\delta)$, Hoeffding (UCB) is considered as confidence-bound for stopping the evaluation of $\boldsymbol{M}_{ij}$. With δ as confidence level and K as interaction times of $\boldsymbol{M}_{ij}$, we can get $\boldsymbol{M}_{ij}$ are bounded in $[\widehat{\boldsymbol{M}}_{ij} - \epsilon, \widehat{\boldsymbol{M}}_{ij} + \epsilon]$, where $\widehat{\boldsymbol{M}}_{ij}$ is empirical estimation and ϵ is a very small quantity calculated by the Hoeffding inequality and $\epsilon < \sqrt{\frac{4M_{\max}^2 \log(2/\delta)}{K}}$.

Algorithm 4 gives the OptSpace algorithm (Keshavan & Oh, 2009; Keshavan et al., 2009; 2010) as a supplement to Algorithm 1 and 2. OptSpace reconstructs a low rank matrix from a small subset of entries. Given incomplete observations $\boldsymbol{M}^\Omega$, OptSpace aims to find $\overline{\boldsymbol{M}}$, such that $\overline{\boldsymbol{M}} = \boldsymbol{U\Sigma V}$, and $\|\overline{\boldsymbol{M}}^\Omega - \boldsymbol{M}^\Omega\|_F$ is minimized. It relies on singular value decomposition for an initial guess and then adopts local manifold optimization. Two important steps are Trimming and Rank-r projection. Trimming eliminates over-represented rows and columns in $\boldsymbol{M}^\Omega$, which are those containing more than $2|\Omega|/n$ observed entries. Let $\tilde{\boldsymbol{M}}^\Omega$ denote the trimmed matrix. Rank-r projection is then applied to find the initialization of $\boldsymbol{U}_0, \boldsymbol{V}_0$. The singular value decomposition of the trimmed matrix $\tilde{\boldsymbol{M}}^\Omega$ is defined as: $\tilde{\boldsymbol{M}}^\Omega = \sum_{i=1}^{n} \Sigma_i \boldsymbol{U}_i \boldsymbol{V}_i^T$, where $\Sigma_1 \geq \Sigma_2 ... \geq \Sigma_n$ are singular values. Then the rank-r projection of $\tilde{\boldsymbol{M}}^\Omega$ is defined as: $P_r(\tilde{\boldsymbol{M}}^\Omega) = \frac{n^2}{|\Omega|} \sum_{i=1}^{r} \Sigma_i \boldsymbol{U}_i \boldsymbol{V}_i^T$. Then we get the reconstructed matrix $\overline{\boldsymbol{M}}$ through gradient descent on the Grassman manifold, with initial condition $(\boldsymbol{U}_0, \boldsymbol{V}_0)$. For more detailed descriptions, see (Keshavan & Oh, 2009; Keshavan et al., 2009; 2010).

Algorithm 3 ResponseGraphUCB$(\delta, \mathcal{S}, \mathcal{C}(\delta))$

Construct list L of pairs of strategy profiles to compare;
Initialize tables $\widehat{\boldsymbol{M}}, \boldsymbol{N}$ to store empirical means and interaction counts while L is not empty do;
while L is not empty **do**
  Select a strategy profile s appearing in an edge in L using sampling scheme S;
  Simulate one interaction for s and update $\widehat{\boldsymbol{M}}, \boldsymbol{N}$ accordingly;
  Check whether any edges are resolved according to $\mathcal{C}(\delta)$, remove them from L if so return empirical table $\widehat{M}$.
end while

Algorithm 4 OptSpace(Matrix completion of $\boldsymbol{M}^{\Omega}$)

Input: A chosen rank r, sampling operator $\Omega \in [n] \times [n]$
Output: The recovered matrix $\overline{\boldsymbol{M}}$
Trim $\boldsymbol{M}^{\Omega}$, and let $\tilde{\boldsymbol{M}}^{\Omega}$ be the output;
Compute the rank-r projection of $\tilde{\boldsymbol{M}}^{\Omega}$, $P_r(\tilde{\boldsymbol{M}}^{\Omega}) = \boldsymbol{U}_0 \boldsymbol{\Sigma}_0 \boldsymbol{V}_0^T$;
Minimize $\tilde{F}(\boldsymbol{U}, \boldsymbol{V})$ through gradient descent, with initial condition $(\boldsymbol{U}_0, \boldsymbol{V}_0)$.
Return $\overline{\boldsymbol{M}} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$

B. Theories and Proofs

B.1. Details of definition and theorem for Proposition 1

Definition 1 ((μ_0, μ_1)-Incoherence(Keshavan et al., 2009))**.** *Let matrix $\boldsymbol{M} \in \mathbb{R}^{n\times n}$ of rank r and the singular value decomposition is $\boldsymbol{M} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$. $\boldsymbol{U}, \boldsymbol{V} \in \mathbb{R}^{n\times r}$ are orthogonal matrices and $\boldsymbol{\Sigma} \in \mathbb{R}^{r\times r}$ is a diagonal matrix. In matrix $\boldsymbol{\Sigma}$, $\Sigma_{\min} = \Sigma_r \leq ... \leq \Sigma_1 = \Sigma_{\max}$, and define $\kappa = (\Sigma_{\max}/\Sigma_{\min})$. If $\boldsymbol{M}$ meet the following two conditions:*
(i) $\forall i \in [n] : \sum_{k=1}^{r} \boldsymbol{U}_{ik}^2 \leq \mu_0 r$, $\sum_{k=1}^{r} \boldsymbol{V}_{ik}^2 \leq \mu_0 r$
(ii) $\forall i, j \in [n] : \left|\sum_{k=1}^{r} \boldsymbol{U}_{ik} \left(\frac{\Sigma_k}{\Sigma_1}\right) \boldsymbol{V}_{jk}\right| \leq \mu_1 \sqrt{r}$
then $\boldsymbol{M}$ is defined as (μ_0, μ_1)-incoherent.

This condition describes that one cannot expect to recover the payoff matrix if the meaningful payoffs are in the null space of the sampling operator. Let $\|\cdot\|_*$ denote the nuclear norm, which is a summation of all singular values. The following theorem supports the result in Proposition 1.

Theorem 3. *(Keshavan et al., 2010) Assume $\boldsymbol{M} \in \mathbb{R}^{n\alpha\times n}$ of rank r that satisfies the incoherence conditions with (μ_0, μ_1). Let $\mu = \max\{\mu_0, \mu_1\}$. Further, assume $\Sigma_{\min} \leq \Sigma_1, \ldots, \Sigma_r \leq \Sigma_{\max}$ with $\Sigma_{\min}, \Sigma_{\max}$ bounded away from 0 and ∞. Then there exists a numerical constant C such that, if*

$$|\Omega| \geq Cnr\sqrt{\alpha}\left(\frac{\Sigma_{\max}}{\Sigma_{\min}}\right)^2 \max\left\{\mu_0 \log n, \mu^2 r\sqrt{\alpha}\left(\frac{\Sigma_{\max}}{\Sigma_{\min}}\right)^4\right\}$$

then the output of OptSpace $\overline{\boldsymbol{M}}$ converges, with high probability, to the matrix $\boldsymbol{M}$.

The proof of Proposition 1 directly follows by applying Theorem 3 with $\alpha = 1$.

B.2. Proof of Theorem 1

We first give the necessary lemmas and theorems for our proof.

Lemma 1. *(Rowland et al., 2019) Suppose there are n strategies and all payoffs are bounded in the interval $[-M_{\max}, M_{\max}]$, and define $L(\alpha, M_{\max}) = 2\alpha\exp(2\alpha M_{\max})$, and $g(\alpha, \eta, p, M_{\max}) = \eta\frac{\exp(2\alpha M_{\max})-1}{\exp(2p\alpha M_{\max})-1}$, where α, η, p are all hyperparameters in α-rank. Let $\epsilon \in (0, 18\times 2^{-n}\sum_{i=1}^{n-1}\binom{n}{i} i^n)$. If $sup_{(i,j)\in[n]\times[n]} |\widehat{\overline{\boldsymbol{M}}}_{i,j} - \boldsymbol{M}_{i,j}| \leq \frac{\epsilon g(\alpha,\eta,p,M_{\max})}{18L(\alpha,M_{\max})\sum_{i=1}^{n-1}\binom{n}{i} i^n}$, then we have $\max_{i\in[n]} |\bar{\tilde{\boldsymbol{\pi}}}(i) - \boldsymbol{\pi}(i)| \leq \epsilon$.*

Theorem 4. *(Keshavan et al., 2009) Let $\boldsymbol{M} \in \mathbb{R}^{n\times n}$ be a (μ_0, μ_1)-incoherent matrix of rank r and the singular value decomposition is $\boldsymbol{M} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$, where $\boldsymbol{U}, \boldsymbol{V} \in \mathbb{R}^{n\times r}$ are orthogonal matrices and $\boldsymbol{\Sigma} \in \mathbb{R}^{r\times r}$ is a diagonal matrix. In* *matrix $\boldsymbol{\Sigma}$, $\Sigma_{\min} = \Sigma_r \leq ... \leq \Sigma_1 = \Sigma_{\max}$, and define $\kappa = (\Sigma_{\max}/\Sigma_{\min})$. Let $\widehat{\boldsymbol{M}} = \boldsymbol{M} + \boldsymbol{Z}$ be the observed matrix with noise $\boldsymbol{Z}$. Define $\Omega \subseteq [n] \times [n]$ is the sampling operator in which m entries are randomly selected for observation from all n^2 entries. Therefore, the matrix with noise observed by the sampling operator Ω is $\widehat{\boldsymbol{M}}^\Omega = \boldsymbol{M}^\Omega + \boldsymbol{Z}^\Omega$. There exist constants C, C' such that if the number of sampled entries satisfies*

$$|\Omega| \geq C\kappa^2 n \max(\mu_0 r \log(n), \mu_0^2 r^2 \kappa^4, \mu_1^2 r^2 \kappa^4)$$

and get $\overline{\widehat{\boldsymbol{M}}}$ through performing matrix completion algorithm ***OptSpace*** *(Keshavan et al., 2009) on $\widehat{\boldsymbol{M}}^\Omega$ then we have*

$$\frac{1}{n}\|\overline{\widehat{\boldsymbol{M}}} - \boldsymbol{M}\|_F \leq C'\kappa^2 \frac{n\sqrt{r}}{|\Omega|}\|\boldsymbol{Z}^\Omega\|_2$$

with probability at least $1 - \frac{1}{n^3}$. The right hand side above is less than $\Sigma_{\min}$.

Theorem 5. *(Keshavan et al., 2009) For any matrix $\boldsymbol{M} \in \mathbb{R}^{n\times n}$ and any set $\Omega \subseteq [n] \times [n]$,*

$$\|\boldsymbol{M}^\Omega\|_2 \leq \frac{2|\Omega|}{n} \max_{(i,j)\in\Omega} |\boldsymbol{M}_{ij}|.$$

Now we are ready to provide the proof for Theorem 1.

Proof of Theorem 1. According to Theorem 4 and 5, we have

$$\begin{aligned}
\|\overline{\widehat{\boldsymbol{M}}} - \widehat{\boldsymbol{M}}\|_F &\leq \|\overline{\widehat{\boldsymbol{M}}} - \boldsymbol{M}\|_F + \|\boldsymbol{M} - \widehat{\boldsymbol{M}}\|_F && (3)\\
&\leq C'\kappa^2 \frac{n^2\sqrt{r}}{|\Omega|}\|\boldsymbol{Z}^\Omega\|_2 + \|\boldsymbol{Z}\|_F && (4)\\
&\leq C'\kappa^2 \frac{n^2\sqrt{r}}{|\Omega|} \cdot \frac{2|\Omega|}{n} \max_{(i,j)\in\Omega} |\boldsymbol{Z}_{ij}| + n\|\boldsymbol{Z}\|_{\max} && (5)\\
&\leq (2C'\kappa^2\sqrt{r} + 1)n\|\boldsymbol{Z}\|_{\max}. && (6)
\end{aligned}$$

Recall that, $\tau = \frac{\epsilon g(\alpha,\eta,p,M_{\max})}{18L(\alpha,M_{\max})\sum_{i=1}^{n-1}\binom{n}{i} i^n (2C'\kappa^2\sqrt{r}+1)n}$. Thus we have

$$\sup_{(i,j)\in[n]\times[n]} |\overline{\widehat{\boldsymbol{M}}}_{i,j} - \widehat{\boldsymbol{M}}_{i,j}| \leq \|\overline{\widehat{\boldsymbol{M}}} - \widehat{\boldsymbol{M}}\|_F \leq \frac{\epsilon g(\alpha,\eta,p,M_{\max})}{18L(\alpha,M_{\max})\sum_{i=1}^{n-1}\binom{n}{i} i^n}. \qquad (7)$$

By applying Lemma 1, we have $\max_{i\in[n]} |\bar{\hat{\boldsymbol{\pi}}}(i) - \hat{\boldsymbol{\pi}}(i)| \leq \epsilon$. Thus the proof of Theorem 1 is completed.

□

B.3. Proofs of Theorem 2

Now we are ready to prove Theorem 2.

Proof of Theorem 2. Define $\boldsymbol{Z} = \widehat{\boldsymbol{M}} - \boldsymbol{M}$. Let $\tau = \frac{\epsilon g(\alpha,\eta,p,M_{\max})|\Omega|}{18L(\alpha,M_{\max})\sum_{i=1}^{n-1}\binom{n}{i} i^n C'\kappa^2 n^2\sqrt{r}}$. Denote $\widehat{\boldsymbol{M}}_{ij} = \frac{1}{K}\sum_{k=1}^{K} \widehat{\boldsymbol{M}}_{ij}^k$, then we have:

$$\begin{aligned}
&P(\|\boldsymbol{Z}^{\Omega}\|_2 > \tau) \le P(\frac{2|\Omega|}{n}\max_{(i,j)\in\Omega}|\boldsymbol{Z}_{ij}| > \tau) \quad \text{(by Theorem 5)} \\
&= P(\max_{(i,j)\in\Omega}|\boldsymbol{Z}_{ij}| > \frac{\tau n}{2|\Omega|}) \\
&= P(\exists (i,j)\in\Omega : |\widehat{\boldsymbol{M}}_{ij} - \boldsymbol{M}_{ij}| > \frac{\tau n}{2|\Omega|}) \\
&\le \sum_{(i,j)\in\Omega} P(|\widehat{\boldsymbol{M}}_{ij} - \boldsymbol{M}_{ij}| > \frac{\tau n}{2|\Omega|}) \qquad (8) \\
&\le \sum_{i,j\in\Omega} \frac{1}{mn^3} \quad (\text{since } K > \frac{8M_{\max}^2 \log(2mn^3)m^2}{\tau^2 n^2}) \qquad (9) \\
&= \frac{1}{n^3}
\end{aligned}$$

Here (8) holds because of union bound theorem (Shalev-Shwartz & Ben-David, 2014). (9) holds because of Hoeffding's Inequality: let $X_1, X_2, ..., X_n$ be i.i.d random variables bounded in $[a, b]$, then for any $\epsilon > 0$, $P\left(\left|\frac{1}{K}\sum_{i=1}^{K} X_i - \mathbb{E}(X_i)\right| > \epsilon\right) \le 2e^{-2K\epsilon^2/(b-a)^2}$. So we get that with probability at least $1 - \frac{1}{n^3}$,

$$\|\boldsymbol{Z}^{\Omega}\|_2 \le \frac{\epsilon g(\alpha, \eta, p, M_{\max})|\Omega|}{18L(\alpha, M_{\max})\sum_{i=1}^{n-1}\binom{n}{i} i^n C' \kappa^2 n^2 \sqrt{r}}$$

Thus, combined with Theorem 4 and the union bound, the probability that the first inequality (in Theorem 4) is true is $1 - 1/n^3$, the probability that the second inequality(above) is true is $1 - 1/n^3$, we can get with probability at least $1 - \frac{2}{n^3}$, that:

$$\|\overline{\widehat{\boldsymbol{M}}} - \boldsymbol{M}\|_F \le \frac{\epsilon g(\alpha, \eta, p, M_{\max})}{18L(\alpha, M_{\max})\sum_{i=1}^{n-1}\binom{n}{i} i^n}$$

Obviously, $\sup_{(i,j)\in[n]\times[n]} |\overline{\widehat{\boldsymbol{M}}}_{i,j} - \boldsymbol{M}_{i,j}| \le \|\overline{\widehat{\boldsymbol{M}}} - \boldsymbol{M}\|_F$. By applying Lemma 1, the proof of Theorem 2 is completed. □

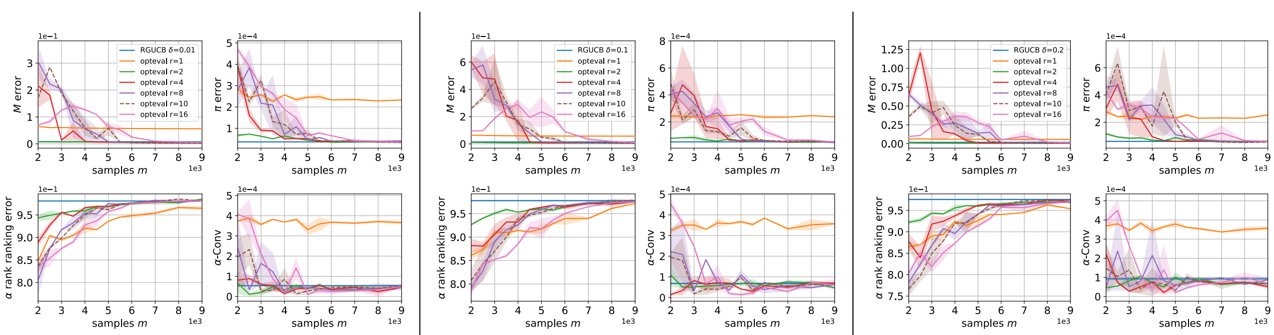
C. Further Experiments

Additional results Figure 7 and 8 show the results with $\alpha = 0.001$ and $\delta \in \{0.01, 0.1, 0.2\}$ on Bern(100) and soccer meta-game, as a supplement for Figure 5. Similarly, Figure 9 and 10 show the results with $\alpha = 0.01$ and $\delta \in \{0.01, 0.1, 0.2\}$ on Bern(100) and soccer meta-game, as a supplement for Figure 5. The results show that, across different choices of α-rank parameters, our algorithm can estimate α-rank with much fewer sample of pairs.

Table 3 shows the statistics of real world games that is used in Figure 1. Table 4 shows results of twelve real world games with α-conv metric, as a supplement of Table 2, which demonstrates that higher rank will lead to lower approximation error on payoff matrices and better convergence to α-rank.

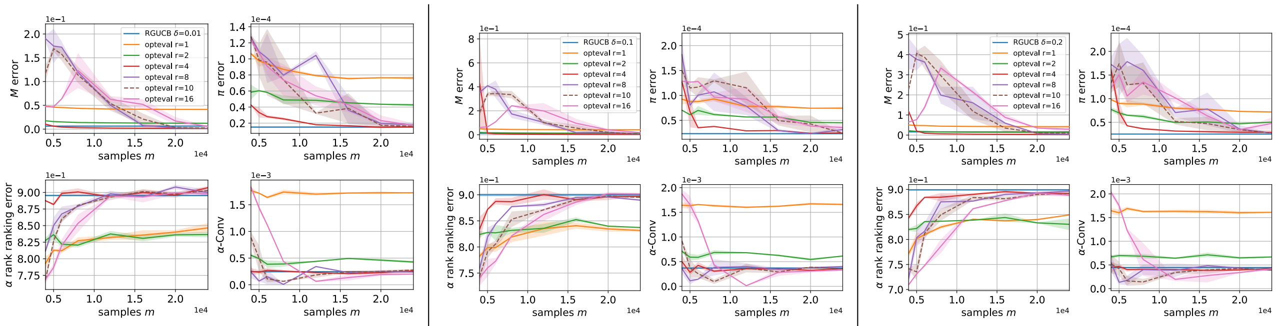
Table 3. Statistics of payoffs from real world games from (Czarnecki et al., 2020). k denote the number of dominant singular values such that $\sum_i^k \Sigma_i / \sum_i^n \Sigma_i \geq 80\%$.

Game	# policies	rank	k
10,3-Blotto	66	30	12
10,4-Blotto	286	40	14
10,5-Blotto	1001	50	16
3-move parity game 2	160	14	9
5,3-Blotto	21	12	7
5,4-Blotto	56	16	8
5,5-Blotto	126	20	10
AlphaStar	888	888	238
Blotto	1001	50	16
Disc game	1000	2	2
Elo game + noise=0.1	1000	1000	396
Elo game + noise=0.5	1000	1000	507
Elo game + noise=1.0	1000	1000	524
Elo game	1000	38	2
Kuhn-poker	64	64	24
Normal Bernoulli game	1000	1000	499
Rock-Paper-Scissors	3	2	2
Random game of skill	1000	1000	515
Transitive game	1000	2	2
Triangular game	1000	1000	137
connect_four	1470	1464	297
go(board_size=3,komi=6.5)	1933	1924	516
go(board_size=4,komi=6.5)	1679	1668	380
hex(board_size=3)	766	764	232
misere(game=tic_tac_toe())	926	926	295
quoridor(board_size=3)	1404	1306	244
quoridor(board_size=4)	1540	1464	343
tic_tac_toe	880	880	285



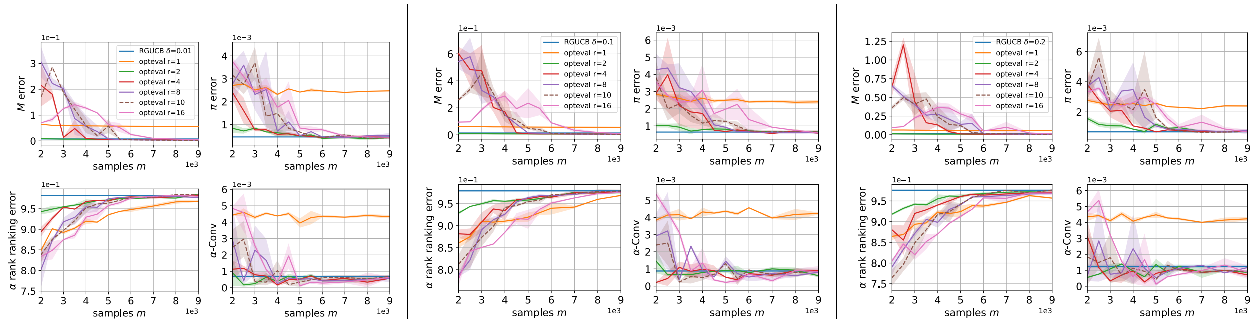
(a) Noisy evaluations, δ=0.01 (b) Noisy evaluations, δ=0.1 (c) Noisy evaluations, δ=0.2

Figure 7. Bernoulli game with $n = 100, r = 10, \alpha = 0.001$ with noisy evaluations.



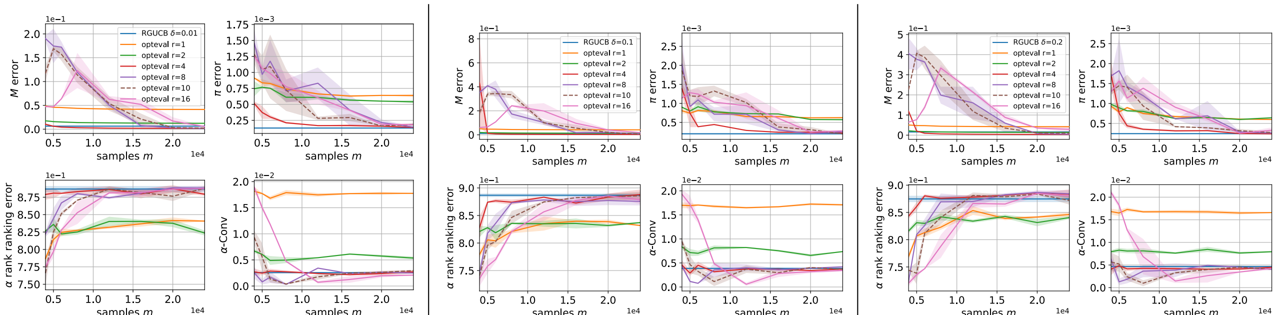
(a) Noisy evaluations, δ=0.01 (b) Noisy evaluations, δ=0.1 (c) Noisy evaluations, δ=0.2

Figure 8. Soccer meta-game with $n = 200, r = 10, \alpha = 0.001$ with noisy evaluations.



(a) Noisy evaluations, δ=0.01 (b) Noisy evaluations, δ=0.1 (c) Noisy evaluations, δ=0.2

Figure 9. Bernoulli game with $n = 100, r = 10, \alpha = 0.01$ with noisy evaluations.



(a) Noisy evaluations, δ=0.01 (b) Noisy evaluations, δ=0.1 (c) Noisy evaluations, δ=0.2

Figure 10. Soccer meta-game with $n = 200, r = 10, \alpha = 0.01$ with noisy evaluations.

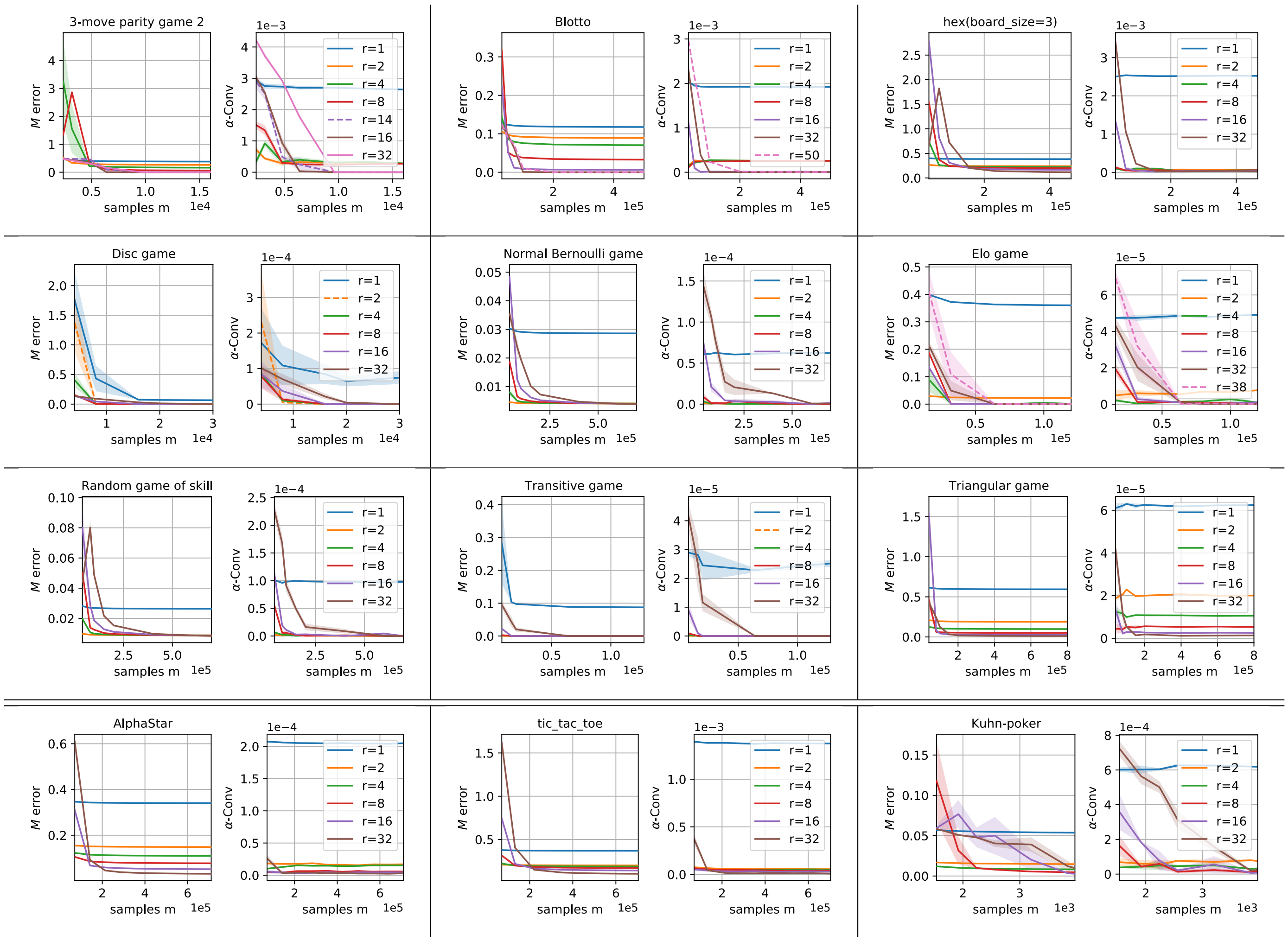


Table 4. Results on twelve real world games with noise free evaluations. (Left of plot) Recovery error on the payoff matrices. (Right of the plot) α-conv error showing the convergence to α-rank.