Multi-Round Cooperative Search Games with Multiple Players

Amos Korman, Yoav Rodeh

Multi-Round Cooperative Search Games with Multiple Players*

Amos Korman[1] and Yoav Rodeh[2]

[1]CNRS and University Paris Diderot, Paris, France, amos.korman@irif.fr
[2]Ort Braude College, Karmiel, Israel, yoav.rodeh@braude.ac.il

Abstract

Assume that a treasure is placed in one of M boxes according to a known distribution and that k searchers are searching for it in parallel during T rounds. We study the question of how to incentivize selfish players so that group performance would be maximized. Here, this is measured by the *success probability*, namely, the probability that at least one player finds the treasure. We focus on *congestion policies* $C(\ell)$ that specify the reward that a player receives if it is one of ℓ players that (simultaneously) find the treasure for the first time. Our main technical contribution is proving that the *exclusive policy*, in which $C(1) = 1$ and $C(\ell) = 0$ for $\ell > 1$, yields a *price of anarchy* of $(1-(1-1/k)^k)^{-1}$, and that this is the best possible price among all symmetric reward mechanisms. For this policy we also have an explicit description of a symmetric equilibrium, which is in some sense unique, and moreover enjoys the best success probability among all symmetric profiles. For general congestion policies, we show how to polynomially find, for any $\theta > 0$, a symmetric multiplicative $(1+\theta)(1+C(k))$-equilibrium.

Together with an appropriate reward policy, a central entity can suggest players to play a particular profile at equilibrium. As our main conceptual contribution, we advocate the use of symmetric equilibria for such purposes. Besides being fair, we argue that symmetric equilibria can also become highly robust to crashes of players. Indeed, in many cases, despite the fact that some small fraction of players crash (or refuse to participate), symmetric equilibria remain efficient in terms of their group performances and, at the same time, serve as approximate equilibria. We show that this principle holds for a class of games, which we call *monotonously scalable* games. This applies in particular to our search game, assuming the natural *sharing policy*, in which $C(\ell) = 1/\ell$. For the exclusive policy, this general result does not hold, but we show that the symmetric equilibrium is nevertheless robust under mild assumptions.

1 Introduction

Searching in groups is ubiquitous in multiple contexts, including in the biological world, in human populations as well as on the internet [1, 2, 3]. In many cases there is some prior on the distribution of the searched target. Moreover, when the space is large, each searcher typically needs to inspect multiple possibilities, which in some circumstances can only be done sequentially. This paper introduces a game theoretic perspective to such multi-round treasure hunt searches, generalizing a basic collaborative Bayesian framework previously introduced in [1].

Consider the case that a treasure is placed in one of M boxes according to a known distribution f and that k searchers are searching for it in parallel during T rounds, each specifying a box to visit in each round. Assume w.l.o.g. that the boxes are ordered such that lower index boxes have higher probability to host the treasure, i.e., $f(x) \geq f(x+1)$. We evaluate the group performance by the *success probability*, that is, the probability that the treasure is found by at least one searcher.

If coordination is allowed, letting searcher i visit box $(t-1)k+i$ at time t will maximize success probability. However, as simple as this algorithm is, it is very sensitive to faults of all sorts. For example, if an adversary that knows where the treasure is can crash a searcher before the search starts (i.e., prevent it from searching), then it can reduce the search probability to zero.

The authors of [1] suggested the use of *identical non-coordinating* algorithms. In such scenarios all processors act independently, using no communication or coordination, executing the same probabilistic

*This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 648032).

algorithm, differing only by the results of their coin flips. As argued in [1], in addition to their economic use of communication, identical non-coordinating algorithms enjoy inherent robustness to different kind of faults. For example, assume that there are $k + k'$ searchers, and that an adversary can fail up to k' searchers. Letting all searchers run the best non-coordinating algorithm for k searchers guarantees that regardless of which $\ell \leq k'$ searchers fail, the overall search efficiency is at least as good as the non-coordinating one for k players. Of course, since k' players might fail, any solution can only hope to achieve the best performance of k players. As it applies to the group performance we term this property as *group robustness.* Among the main results in [1] is identifying a non-coordinating algorithm, denoted $A^\star$, whose expected running time is minimal among non-coordinating algorithms. Moreover, for every given T, if this algorithm runs for T rounds, it also maximizes the success probability.

The current paper studies the game theoretic version of this multi-round search problem[1]. The setting of [1] assumes that the searchers adhere fully to the instructions of a central entity. In contrast, in a game theoretical context, searchers are self-interested and one needs to incentivize them to behave as desired, e.g., by awarding those players that find the treasure first. For many real world contexts, the competitive setting is in fact the more realistic one to assume. Applications range from crowd sourcing [4], multi-agent searching on the internet [5], grant proposals [6], to even contexts of animals [7] (see Appendix A).

In the competitive setting, choosing a good *rewarding policy* becomes a problem in algorithmic mechanism design [8]. Typically, a reward policy is evaluated by its *price of anarchy (PoA)*, namely, the ratio between the performances of the best collaborative algorithm and the worst equilibrium [9]. Aiming to both accelerate the convergence process to an equilibrium and obtain a preferable one, the announcement of the reward policy can be accompanied by a proposition for players to play particular strategies that form a profile at equilibrium.

This paper highlights the benefits of suggesting (non-coordinating) *symmetric equilibria* in such scenarios, that is, to suggest the same non-coordinating strategy to be used by all players, such that the resulting profile is at equilibrium. This is of course relevant assuming that the *price of symmetric stability (PoSS)*, namely, the ratio between the performances of the best collaborative algorithm and the best symmetric equilibrium, is low. Besides the obvious reasons of fairness and simplicity, from the perspective of a central entity who is interested in the overall success probability, we obtain the group robustness property mentioned above, by suggesting that the $k+k'$ players play according to the strategy that is a symmetric equilibrium for k players. Obviously, this group robustness is valid only provided that the players indeed play according to the suggested strategy. However, the suggested strategy is guaranteed to be an equilibrium only for k players, while in fact, the adversary may keep some of the extra k' players alive. Interestingly, however, in many cases, a symmetric equilibrium for k players also serves as an approximate equilibrium for $k + k'$ players, as long as $k' \ll k$. As we show, this *equilibrium robustness* property is rather general, holding for a class of games, that we call *monotonously scalable* games.

1.1 The Collaborative Search Game

A treasure is placed in one of M boxes according to a known distribution f and k players are searching for it in parallel during T rounds. Assume w.l.o.g. that $f(x) > 0$ for every x and that $f(x) \geq f(x + 1)$.

Strategies. An *execution* of T rounds is a sequence of box visitations $\sigma = x(1), x(2), \ldots x(T)$, one for each round $i \leq T$. We assume that a player visiting a box has no information on whether other players have already visited that box or are currently visiting it. Hence, a *strategy* of a player is a probability distribution over the space of executions of T rounds. Note that the probability of visiting a box x in a certain round may depend on the boxes visited by the player until this round, but not on the actions of other players. A strategy is *non-redundant* if at any given round it always checks a box it didn't check before (as long as there are such boxes).

A *profile* is a collection of k strategies, one for each player. Special attention will be devoted to *symmetric profiles.* In such profiles all players play the same strategy (note that their actual executions may be different, due to different probabilistic choices).

Probability Matrix. While slightly abusing notation, we shall associate each strategy A with its *probability matrix*, $A : \{1, \ldots, M\} \times \{1, \ldots, T\} \to [0, 1]$, where $A(x, t)$ is the probability that strategy A

[1]We concentrate on the normal form version in which players do not receive any feedback during the search (except when the treasure is found in which case the game ends). In particular, we assume that players cannot communicate with each other.

visits x for the first time at round t. We also denote $\widetilde{A}(x,t)$ as the probability that A does not visit x by, and including, time t. That is, $\widetilde{A}(x,t) = 1 - \sum_{s \le t} A(x,t)$ and $\widetilde{A}(x,0) = 1$. For convenience we denote by $\delta_{x,t}$ the matrix of all zeros except 1 at x,t. Its dimensions will be clear from context.

Group Performance. A profile is evaluated by its *success probability*, i.e., the probability that at least one player finds the treasure by time T. Formally, let $\mathbb{P}$ be a profile. Then,

$$\texttt{success}(\mathbb{P}) = \sum_x f(x)\left(1 - \prod_{A\in\mathbb{P}} \widetilde{A}(x,T)\right).$$

The expected running time in the symmetric case, which is $\sum_x f(x) \sum_t \widetilde{A}(x,t)^k$, was studied in [1]. That paper identified a strategy, denoted $A^\star$, that minimizes this quantity. In fact, it does so by minimizing the term $\sum_x f(x)\widetilde{A}(x,t)^k$ for each t separately. Note that minimizing the case $t = T$ is exactly the same as maximizing the success probability. Thus, restricted to the case where all searchers use the same strategy, $A^\star$ simultaneously optimizes the success probability as well as optimizes the expected running time. For completeness, a description of $A^\star$ is provided below.

Algorithm $A^\star$. We note that in [1] the matrix of $A^\star$ is given, and then an algorithm is explicitly described that has its matrix (Section 4.3 in [1]). We describe the matrix only, as its details are necessary for this paper. Denote $q(x) = f(x)^{-1/(k-1)}$. For each t, $\widetilde{A^\star}(x,t) = \min(1, \alpha(t)q(x))$, where $\alpha(t) \ge 0$ is such that $\sum_x \widetilde{A^\star}(x,t) = M - t$. Of course, once $\widetilde{A^\star}$ is known, then so is $A^\star$: $A^\star(x,t) = \widetilde{A^\star}(x,t-1) - \widetilde{A^\star}(x,t)$.

Congestion Policies. A natural way to incentivize players is by rewarding those players that find the treasure before others. A *congestion policy* $C(\ell)$ is a function specifying the reward that a player receives if it is one of ℓ players that (simultaneously) find the treasure for the first time. We assume that $C(1) = 1$, and that C is non-negative and non-increasing. Due to the fact that the policy $C \equiv 1$ is rather degenerate, we henceforth assume that $C \not\equiv 1$. We shall give special attention to the following policies.

- The *sharing policy* is defined by $C_{\texttt{share}}(\ell) = 1/\ell$, namely, the treasure is shared equally among all those who find it first.
- The *exclusive policy* is defined by $C_{\texttt{ex}}(1) = 1$, and $C_{\texttt{ex}}(\ell) = 0$ for $\ell > 1$, namely, the treasure is given to the first one that finds it exclusively; if more than one discover it, they get nothing[2].

A *configuration* is a triplet (C, f, T), where C is a congestion policy, T is a positive integer, and f is a positive non-increasing probability distribution on M boxes.

Values, Utilities and Equilibria. Let (C, f, T) be a configuration. The *value* of box x at round t when playing against a profile $\mathbb{P}$ is the expected gain from visiting x at round t. Formally,

$$\begin{aligned} v_{\mathbb{P}}(x,t) &= f(x)\sum_{\ell=0}^{k-1} C(\ell+1)\Pr\begin{pmatrix} x \text{ was not visited before time } t\text{, and} \\ \text{at time } t \text{ is visited by } \ell \text{ players of } \mathbb{P}\end{pmatrix} \\ &= f(x)\sum_{\ell=0}^{k-1} C(\ell+1) \sum_{\substack{I\subseteq\mathbb{P}\\|I|=\ell}} \prod_{A\in I} A(x,t) \prod_{A\notin I}\widetilde{A}(x,t). \end{aligned}$$

The *utility of A in round* t and the *utility of A* are defined as:

$$U_{\mathbb{P}}(A,t) := \sum_x A(x,t)\cdot v_{\mathbb{P}^{-A}}(x,t), \qquad U_{\mathbb{P}}(A) := \sum_t U_{\mathbb{P}}(A,t), \tag{1}$$

where $\mathbb{P}^{-A}$ is the set of players of $\mathbb{P}$ excluding A. Here are some specific cases we are interested in:

- For symmetric profiles, $v_A(x,t)$ denotes the value when playing against $k-1$ players playing A. Then $v_A(x,t) = f(x)\sum_{\ell=0}^{k-1} C(\ell+1)\binom{k-1}{\ell} A(x,t)^\ell \widetilde{A}(x,t)^{k-\ell-1}$.
- For the exclusive policy, $v_{\mathbb{P}}(x,t) = f(x)\prod_{A\in\mathbb{P}}\widetilde{A}(x,t)$.

[2]In the one round game, the exclusive policy yields a utility for a player that equals its marginal contribution to the social welfare, i.e., the success probability [10]. However, this is not the case in the multi-round game.

- For the exclusive policy in symmetric profiles, $v_A(x,t) = f(x)\widetilde{A}(x,t)^{k-1}$.

A profile $\mathbb{P}$ is a *Nash equilibrium* under a configuration, if for any $A \in \mathbb{P}$ and any other strategy B, $U_{\mathbb{P}^{-A}}(A) \geq U_{\mathbb{P}^{-A}}(B)$. Similarly, a strategy A is called a *symmetric equilibrium* if the profile A^k consisting of all k players playing according to A is an equilibrium. We also use the notion of approximate equilibrium. For $\epsilon > 0$, we say a profile $\mathbb{P}$ is a $(1+\epsilon)$-equilibrium if for every $A \in \mathbb{P}$ and for every strategy B, $U_{\mathbb{P}^{-A}}(B) \leq (1+\epsilon)U_{\mathbb{P}^{-A}}(A)$.

A Game of Doubly-Substochastic Matrices. Both the expressions for the success probability and utility solely depend on the values of the probability matrices associated with the strategies in question. Hence we view all strategies sharing the same matrix as *equivalent*. Note that a matrix does not necessarily correspond to a unique strategy, as illustrated by the following equivalent strategies, for which $A(x,t) = B(x,t) = 1/M$ for every $t \leq M$ and 0 thereafter:

- Strategy A chooses uniformly at every round one of the boxes it didn't choose yet.
- Strategy B chooses once $x \in \{0, \dots, M-1\}$. Then, at round t it visits box $(x + t \mod M) + 1$.

Matrices are much simpler to handle than strategies, and so we would rather think of our game as a game of probability matrices than a game of strategies. For this we need to characterize which matrices are indeed probability matrices of strategies. Clearly, a probability matrix is non-negative. Also, by their definition, each row and each column sums to at most 1. Such a matrix is called *doubly-substochastic*. In Appendix C we prove the converse, i.e., that every doubly-substochastic matrix is a probability matrix of some strategy. Furthermore, this strategy is implementable as a polynomial algorithm. We will therefore view our game as a game of doubly-substochastic matrices.

Greediness. Informally, a strategy is greedy at a round if its utility in this round is the maximum possible in this round. Formally, given a profile $\mathbb{P}$ and some strategy A, we say that A is *greedy* w.r.t. $\mathbb{P}$ at time t if for any strategy B such that for every x and $s < t$, $B(x,s) = A(x,s)$, we have $U_{\mathbb{P}}(A,t) \geq U_{\mathbb{P}}(B,t)$. We say A is greedy w.r.t. $\mathbb{P}$ if it is greedy w.r.t. $\mathbb{P}$ for each $t \leq T$. A strategy A is called *self-greedy* (or *sgreedy* for short) if it is greedy w.r.t. the profile with $k-1$ players playing A.

Evaluating Policies. Let (C, f, T) be a configuration. Denote by $\text{Nash}(C,f,T)$ the set of equilibria for this configuration, and by $\text{S-Nash}(C,f,T)$ the subset of symmetric ones. Let $\mathcal{P}(T)$ be the set of all profiles of T-round strategies. We are interested in the following measures.

- The *Price of Anarchy (PoA)* is $\texttt{PoA}(C,f,T) := \frac{\max_{\mathbb{P}\in\mathcal{P}(T)} \texttt{success}(\mathbb{P})}{\min_{\mathbb{P}\in\text{Nash}(C,f,T)} \texttt{success}(\mathbb{P})}$.
- The *Price of Symmetric Stability (PoSS)* is $\texttt{PoSS}(C,f,T) := \frac{\max_{\mathbb{P}\in\mathcal{P}(T)} \texttt{success}(\mathbb{P})}{\max_{A\in\text{S-Nash}(C,f,T)} \texttt{success}(A^k)}$.
- The *Price of Symmetric Anarchy (PoSA)* is $\texttt{PoSA}(C,f,T) = \frac{\max_{\mathbb{P}\in\mathcal{P}(T)} \texttt{success}(\mathbb{P})}{\min_{A\in\text{S-Nash}(C,f,T)} \texttt{success}(A^k)}$.

On the Difficulty of the Multi-Round Game. The setting of multi-rounds poses several challenges that do not exist in the single round game. An important one is the fact that, in contrast to the single round game, the multi-round game is not a potential game. Indeed, being a potential game has several implications, and a significant one is that such a game has always a pure equilibrium. However, we show that multi-round games do not always have pure equilibria and hence they are not potential games. Another important difference is that for policies that incur high levels of competition (such as the exclusive policy), profiles that maximize the success probability are at equilibrium in the single round case, whereas they are not in the multi-round game. See Appendix B for more details.

1.2 Our Results

Equilibrium Robustness. We first provide a simple, yet general, robustness result, that holds for symmetric (approximate) equilibria in a family of games, termed *monotonously scalable*. Informally, these are games in which the sum of utilities of players can only increase when more players are added, yet for each player, its individual utility can only decrease. Our search game with the sharing policy is one such example.

Theorem 1. *Consider a symmetric monotonously scalable game. If A is a symmetric $(1+\epsilon)$-equilibrium for k players, then it is an $(1+\epsilon)(1+\ell/k)$-equilibrium when played by $k+\ell$ players.*

Theorem 1 is applicable in fault tolerant contexts. Consider a monotonously scalable game with $k+k'$ players out of which at most k' may fail. Let A_k be a symmetric (approximate) equilibrium designed for k players and assume that its social utility is high compared to the optimal profile with k players. The theorem implies that if players play A_k, then regardless of which $\ell \leq k'$ players fail (or decline to participate), the incentive to switch strategy would be very small, as long as $k' \ll k$. Moreover, due to symmetry, if the social utility of the game is monotone, then the social utility of A_k when played with k players is guaranteed when playing with more. Thus, in such cases we obtain both group robustness and equilibrium robustness.

General Congestion Policies. Coming back to our search game, we consider general policies, focus on symmetric profiles, and specifically, on the properties of sgreedy strategies.

Theorem 2. *For every policy C there exists a non-redundant sgreedy strategy. Moreover, all such strategies are equivalent and are symmetric $(1 + C(k))$-equilibria.*

When $C(k) = 0$ this shows that a non-redundant sgreedy strategy is actually a symmetric equilibrium. We next claim that this is the only symmetric equilibrium (up to equivalence).

Claim 3. *For any policy such that $C(k) = 0$, all symmetric equilibria are equivalent.*

Theorem 2 is non-constructive because it requires calculating the inverse of non-trivial functions. Therefore, we resort to an approximate solution.

Theorem 4. *Given $\theta > 0$, there exists an algorithm that takes as input a configuration, and produces a symmetric $(1 + C(k))(1 + \theta)$-equilibrium. The algorithm runs in polynomial time in T, k, M, $\log(1/\theta)$, $\log(1/(1 - C(k)))$, and $\log(1/f(M))$.*

The Exclusive Policy. Recall that the exclusive policy is defined by $C_{\mathtt{ex}}(1) = 1$ and $C_{\mathtt{ex}}(\ell) = 0$ for every $\ell > 1$. We show that $A^\star$ is a non-redundant and sgreedy strategy in the exclusive policy. Hence, Theorem 2 implies the following.

Theorem 5. *Under the exclusive policy, Strategy $A^\star$ of [1] is a symmetric equilibrium.*

Claim 3 together with the fact (established in [1]) that $A^\star$ has the highest success probability among symmetric profiles, implies that both the PoSS and the PoSA of $C_{\mathtt{ex}}$ are optimal (and equal) on any f and T when compared to any other policy. The next theorem considers general equilibria.

Theorem 6. *Consider the exclusive policy. For any profile $\mathbb{P}_{\mathtt{nash}}$ at equilibrium and any symmetric profile A, $\mathtt{success}(\mathbb{P}_{\mathtt{nash}}) \geq \mathtt{success}(A)$.*

Observe that, as $A^\star$ is a symmetric equilibrium, Theorem 6 provides an alternative proof for the optimality of $A^\star$ (established in [1]). Interestingly, this alternative proof is based on game theoretic considerations, which is a very rare approach in optimality proofs.

Combining Theorems 5 and 6, we obtain:

Corollary 7. *For any f and T, $\mathtt{PoA}(C_{\mathtt{ex}}, f, T) = \mathtt{PoSA}(C_{\mathtt{ex}}, f, T)$. Moreover, for any f, T and policy C, $\mathtt{PoA}(C_{\mathtt{ex}}, f, T) \leq \mathtt{PoA}(C, f, T)$.*

At first glance the effectiveness of C_{ex} might not seem so surprising. Indeed, it seems natural that high levels of competition would incentivize players to disperse. However, it is important to note that $C_{\mathtt{ex}}$ is not extreme in this sense, as one may allow congestion policies to also have negative values upon collisions. Moreover, one could potentially define more complex kinds of policies, e.g., policies that depend on time, and reward early finds more. However, the fact that $A^\star$ is optimal among all symmetric profiles combined with the fact that any symmetric policy has a symmetric equilibrium [11] implies that no symmetric reward mechanism can improve either the PoSS, the PoSA, or the PoA of the exclusive policy.

We proceed to show a tight upper bound on the PoA of $C_{\mathtt{ex}}$. Note that as k goes to infinity the bound converges to $e/(e - 1) \approx 1.582$.

Theorem 8. *For every T, $\sup_f \mathtt{PoA}(C_{\mathtt{ex}}, f, T) = (1 - (1 - 1/k)^k)^{-1}$.*

Concluding the results on the exclusive policy, we study the robustness of $A^\star$ in Appendices F.3 and F.4. Let $A^\star_k$ denote algorithm $A^\star$ when set to work for k players. Unfortunately, for any ϵ, there are cases where $A^\star_k$ is not a $(1 + \epsilon)$-equilibrium even when played by $k + 1$ players. However, as indicated below, $A^\star$ is robust to failures under reasonable assumptions regarding the distribution f.

Theorem 9. *If* $\frac{f(1)}{f(M)} \leq (1+\epsilon)^{\frac{k-1}{k'}}$*, then* $A_k^\star$ *is a* $(1+\epsilon)$*-equilibrium when played by* $k+k'$ *players.*

The Sharing Policy. Another important policy to consider is the sharing policy. This policy naturally arises in some circumstances, and may be considered as a less harsh alternative to the exclusive one. Although not optimal, it follows from Vetta [12] that its PoA is at most 2 (see Appendix G). Furthermore, as this policy yields a monotonously scalable game, a symmetric equilibrium under it is also robust. Therefore, the existence of a symmetric profile which is both robust and has a reasonable success probability is guaranteed.

Unfortunately, we did not manage to find a polynomial algorithm that generates a symmetric equilibrium for this policy. However, Theorem 4 gives a symmetric $(1+\theta)(1+1/k)$-equilibrium in polynomial time for any $\theta > 0$. This strategy is also robust thanks to Theorem 1. Moreover, the proof in [12] regarding the PoA can be extended to hold for approximate equilibria. In particular, if $\mathbb{P}$ is some $(1+\epsilon)$-equilibrium in the sharing policy, then for every f and T, $\texttt{success}(\mathbb{P}) \geq \frac{1}{2+\epsilon} \max_{\mathbb{P}' \in \mathcal{P}(T)} \texttt{success}(\mathbb{P}')$ (see Appendix G).

1.3 Related Works

Fault tolerance has been a major topic in distributed computing for several decades, and in recent years more attention has been given to these concepts in game theory [13, 14]. For example, Gradwohl and Reingold studied conditions under which games are robust to faults, showing that equilibria in anonymous games are fault tolerant if they are "mixed enough" [15].

Restricted to a single round the search problem becomes a coverage problem, which has been investigated in several papers. For example, Collet and Korman studied in [16] (one-round) coverage while restricting attention to symmetric profiles only. The main result therein is that the exclusive policy yields the best coverage among symmetric profiles. Gairing [17] also considered the single round setting, but studied the optimal PoA of a more general family of games called *covering games* (see also [10, 18]). Motivated by policies for research grants, Kleinberg and Oren [6] considered a one-round model similar to that in [16]. Their focus however was on pure strategies only. The aforementioned papers give a good understanding of coverage games in the single round setting. As mentioned, however, the multi-round setting studied here is substantially more complex than the single-round setting.

The area of "incentivizing exploration" also studies the tradeoff between exploration, exploitation and incentives [19, 20, 4, 21]. This area often focuses on different variants of the Multi-Armed Bandit problem. The settings of selfish routing, job scheduling, and congestion games [22, 23] all bear similarities to the search game studied here, however, the social welfare measurements of success probability or running time are very different from the measures studied in these frameworks, such as makespan or latency [24, 25, 8, 26].

2 Robustness in Symmetric Monotonously Scalable Games

Consider a symmetric game where the number of players is not fixed. Let $U_{\mathbb{P}}(A)$ denote the utility that a player playing A gets if the other players play according to $\mathbb{P}$ and let $\sigma(\mathbb{P}) = \sum_{A \in \mathbb{P}} U_{\mathbb{P}-A}(A)$. We say that such a game is *monotonously scalable* if:

1. Adding more players can only increase the sum of utilities, i.e., if $\mathbb{P} \subseteq \mathbb{P}'$ then $\sigma(\mathbb{P}) \leq \sigma(\mathbb{P}')$.
2. Adding more players can only decrease the individual utilities, i.e., if $\mathbb{P} \subseteq \mathbb{P}'$ then for all $A \in \mathbb{P}$, $U_{\mathbb{P}-A}(A) \geq U_{\mathbb{P}'-A}(A)$.

Theorem 1. *Consider a symmetric monotonously scalable game. If* A *is a symmetric* $(1+\epsilon)$*-equilibrium for* k *players, then it is an* $(1+\epsilon)(1+\ell/k)$*-equilibrium when played by* $k+\ell$ *players.*

Proof. On the one hand by symmetry,

$$U_{A^{k+\ell-1}}(A) = \frac{\sigma(A^{k+\ell})}{k+\ell} \geq \frac{\sigma(A^k)}{k+\ell},$$

where the last step is because σ is non-decreasing. On the other hand, if B is some other strategy,

$$U_{A^{k+\ell-1}}(B) \leq U_{A^{k-1}}(B) \leq (1+\epsilon) U_{A^{k-1}}(A) = (1+\epsilon)\frac{\sigma(A^k)}{k}.$$

The first inequality is because U is non-increasing, and the second is because A is a $(1+\epsilon)$-equilibrium for k players. Therefore, what a player can gain by switching from A to B is at most a multiplicative factor of $(1+\epsilon)(k+\ell)/k = (1+\epsilon)(1+\ell/k)$. $\square$

An example of such a game is our setting with the sharing policy. Note however, that our game with the exclusive policy does not satisfy the first property, as adding more players can actually deteriorate the sum of utilities. Another example is a generalization known as *covering games* [17]. This sort of game is the same as our game in a single-round version, except that each player chooses not necessarily one element, but a set of elements, from a prescribed set of sets. Again, to be a monotonously scalable game, the congestion policy should be the sharing policy. Note that one may consider a multi-round version of these games, which will be monotonously scalable as well.

3 General Policies

The proofs of this section appear in Appendix D.

3.1 Non-Redundancy and Monotonicity

A doubly-substochastic matrix A is called *non-redundant at time* t if $\sum_x A(x,t) = 1$. It is *non-redundant* if it is non-redundant for every $t \leq M$. In the algorithmic view, as $\sum_x A(x,t)$ is the probability that a new box is opened at time t, then a strategy's matrix is non-redundant iff it never checks a box twice, unless it already checked all boxes.

Lemma 10. *If a profile $\mathbb{P}$ is at equilibrium and* $\texttt{success}(\mathbb{P}) < 1$ *then every player is non-redundant.*

We will later see that in the symmetric case the condition in the lemma is not needed. However, the following example shows it is necessary in general. Let $M = k$, $T > 1$, and assume that for every $1 \leq i \leq k$, player i goes to box i in every round. Under the exclusive policy, this strategy is an equilibrium, whereas each player is clearly redundant. The following monotonicity lemmas hold under any congestion policy C.

Lemma 11. *Consider two doubly-substochastic matrices A and B. If $A(x,t) > B(x,t)$, and for all $s < t$, $A(x,s) = B(x,s)$ then $v_A(x,t) < v_B(x,t)$.*

Lemma 12. *Let A be doubly-substochastic. For every x and t, $v_A(x,t+1) \leq v_A(x,t)$. Moreover, if $A(x,t+1) > 0$ then the inequality is strict.*

Using the above, we prove a stronger result than Lemma 10 for the symmetric case:

Lemma 13. *If A is a symmetric equilibrium then it is non-redundant.*

Proof. Because of Lemma 10 it is sufficient to consider only the case where $\texttt{success}(A) = 1$. Let $T' = \min\{M, T\}$, and assume by contradiction that A is redundant. Thus there is some $t \leq T'$ where $\sum_x A(x,t) < 1$. Hence, $\sum_{s \leq T'} \sum_x A(x,s) < T'$. Therefore, there is some x such that $\sum_{s \leq T'} A(x,s) < 1$ and so $\sum_{s \leq t} A(x,s) < 1$. As $\texttt{success}(A) = 1$, there is some $t' > t$ such that $A(x,t') > 0$. Define $A' = A + \epsilon(\delta_{x,t} - \delta_{x,t'})$. Taking $\epsilon > 0$ small enough, A' is doubly-substochastic. Also, $U_A(A') - U_A(A) = \epsilon(v_A(x,t) - v_A(x,t'))$, which is strictly positive by Lemma 12. Contradicting the fact that A is an equilibrium. $\square$

3.2 Greedy Strategies

Lemma 14. *A non-redundant strategy A is greedy w.r.t. $\mathbb{P}$ at time t iff for every x and y, if $A(x,t) > 0$ and $v_{\mathbb{P}}(x,t) < v_{\mathbb{P}}(y,t)$ then $\widetilde{A}(y,t) = 0$.*

The lemma above gives a useful equivalent definition for greediness. We can then prove:

Theorem 2. *For every policy C there exists a non-redundant sgreedy strategy. Moreover, all such strategies are equivalent and are symmetric $(1 + C(k))$-equilibria.*

Proof. Proving the existence of a strategy A that is non-redundant and sgreedy is deferred to the appendix (see Lemma 24). We prove here that such a strategy is a $(1 + C(k))$-equilibrium. Consider a strategy B. We compare the utility of B to that of A when both play against $k - 1$ players playing A. By non-redundancy, all of $v_A(x, t)$ are 0 when $t > M$, and so we can assume $T \leq M$.

Denote $\mathtt{maxv}(t) = \max_x v_A(x, t)$. Since the utility of B in any round t is a convex combination of $v_A(x, t)$, we have $U_A(B, t) \leq \mathtt{maxv}(t)$. We say that A *fills* box x at round t if $A(x, t) > 0$ and $\widetilde{A}(x, t) = 0$. The following four claims hold for any round t:

1. If A does not fill any box at round t then $U_A(A, t) = \mathtt{maxv}(t)$. This is because $U_A(A, t)$ is a convex combination of $v_A(x, t)$ for the boxes where $A(x, t) > 0$, which by the characterization of greediness in Lemma 14, all have the same value at time t.

2. $U_A(A, 1) = \mathtt{maxv}(1)$. Why? if no box is filled in round 1, then Item 1 applies. Otherwise, for some box x, $A(x, 1) = 1$, and all other boxes have $A(\cdot, 1) = 0$. The result follows again by Lemma 14.

3. For any $s < t$, $U_A(A, s) \geq \mathtt{maxv}(t)$. We prove this by showing that for every x, $U_A(A, s) \geq v_A(x, t)$. If $v_A(x, t) = 0$, then the claim is clear. Otherwise, $A(x, t) > 0$ or $\widetilde{A}(x, t) > 0$ or both. Either way, $\widetilde{A}(x, s) > 0$. Therefore, as A is sgreedy, for every y such that $A(y, s) > 0$, $v_A(y, s) \geq v_A(x, s) \geq v_A(x, t)$. The last inequality follows from monotonicity, i.e., Lemma 12. As $v_A(A, s)$ is a convex combination of such y's we conclude.

4. If A fills box x at time $t > 1$ then for any $s < t$, $v_A(x, t) \leq C(k)v_A(x, s)$. To see why, first note that $v_A(x, s) \geq f(x)C(1)\widetilde{A}(x, s)^{k-1} = f(x)\widetilde{A}(x, s)^{k-1}$. On the other hand, since $\widetilde{A}(x, t) = 0$, $v_A(x, t) = f(x)C(k)A(x, t)^{k-1} \leq f(x)C(k)\widetilde{A}(x, s)^{k-1}$, because $A(x, t) \leq \widetilde{A}(x, t - 1) \leq \widetilde{A}(x, s)$. Combining the above two inequalities gives the result.

Denote by X_1 the set of rounds for which there is no box x that is filled by A. Let X_2 be the rest of the rounds, except for $t = 1$ which is in neither. Also denote $t_0 = \min X_2$, and $t_1 = \max X_2$. Since $U_A(B) \leq \sum_t \mathtt{maxv}(t)$, by Items 1,2 and 3 above,

$$U_A(B) \leq \sum_{t \in X_1 \cup \{1\}} U_A(A, t) + \mathtt{maxv}(t_0) + \sum_{t \in X_2 \setminus \{t_1\}} U_A(A, t) \leq U_A(A) + \mathtt{maxv}(t_0).$$

We conclude by using Items 4 and 2 and showing:

$$\mathtt{maxv}(t_0) = \max_x v_A(x, t_0) \leq \max_x C(k)v_A(x, 1) = C(k)U_A(A, 1) \leq C(k)U_A(A). \qquad \square$$

In Appendix D.3.2 we provide an example showing that in the sharing policy, a non-redundant sgreedy strategy is not necessarily at equilibrium. On the other hand, it is worth noting that for any policy, the existence of a symmetric equilibrium follows from [11], and for $C(k) = 0$ we can get a full characterization of such equilibria:

Claim 3. *For any policy such that $C(k) = 0$, all symmetric equilibria are equivalent.*

Interestingly, this result does not extend to non-symmetric profiles even for the exclusive policy, as is demonstrated by the following example of a non-greedy non-redundant equilibrium. Consider three players and two rounds. $f(1) = f(2) = f(3) = (1 - \epsilon)/3, f(4) = \epsilon$, for some small positive ϵ. Player 1 plays first 4 and then 1. Player 2 plays 2 and then 3, and player 3 plays 3 and then 2. This can be seen to be an equilibrium, yet player 1 is not greedy.

Finally, the proof of Theorem 4, which shows how to construct an approximate equilibrium in polynomial time, is deferred to Appendix E. This proof involves defining notions of approximate greediness and non-redundancy, proving an equivalent of Theorem 2 for them, and then using bounds on the rate of change that $v_A(x, t)$ goes through as a function of $A(x, t)$. This allows us to polynomially find an approximate sgreedy and non-redundant matrix, thus giving a polynomial strategy with our use of the Birkhoff von-Neumann theorem (Appendix C).

4 The Exclusive Policy

Missing proofs of this section appear in Appendix F. There, we first prove that under the exclusive policy, $A^\star$ is sgreedy and non-redundant. Hence, according to Theorem 2,

Theorem 5. *Under the exclusive policy, Strategy $A^\star$ of [1] is a symmetric equilibrium.*

According to Claim 3 all symmetric equilibria under the exclusive policy are equivalent, and thus equivalent to $A^\star$. Hence, the optimality of $A^\star$ (w.r.t. symmetric profiles) implies that both the PoSA and PoSS of the exclusive policy are optimal. That is, for every f, T, and policy C,

$$\texttt{PoSA}(C_{\texttt{ex}}, f, T) = \texttt{PoSS}(C_{\texttt{ex}}, f, T) \leq \texttt{PoSS}(C, f, T).$$

Our next goal is to establish the PoA of the exclusive policy. For this purpose, we first prove that the success probability of any equilibrium is at least as large as that of any symmetric profile. Since $A^\star$ is a symmetric equilibrium, its optimality among symmetric profiles follows. Hence, the proof provides as alternative proof to the one in [1].

Theorem 6. *Consider the exclusive policy. For any profile $\mathbb{P}_{\texttt{nash}}$ at equilibrium and any symmetric profile A, $\texttt{success}(\mathbb{P}_{\texttt{nash}}) \geq \texttt{success}(A)$.*

Proof. Let A be a strategy and $\mathbb{P}_{\texttt{nash}}$ be a profile at equilibrium with respect to $C_{\texttt{ex}}$. If $\texttt{success}(\mathbb{P}_{\texttt{nash}}) = 1$, then the inequality is trivial. According to Lemma 10, we can therefore assume that all players of $\mathbb{P}_{\texttt{nash}}$ are non-redundant and that $T \leq M$. Denote the probability of visiting x in profile $\mathbb{P}$ by

$$\texttt{success}(\mathbb{P}, x) = 1 - \prod_{B \in \mathbb{P}} \widetilde{B}(x, T).$$

We say that box x is *high* with respect to a profile $\mathbb{P}$ if $\texttt{success}(\mathbb{P}, x) > \texttt{success}(A, x)$, *low* if $\texttt{success}(\mathbb{P}, x) < \texttt{success}(A, x)$, and *saturated* if they are equal. The next lemma uses the fact that A is symmetric.

Lemma 15. *If a profile $\mathbb{P}$ is non-redundant and contains no high boxes, then all boxes are saturated.*

We proceed to prove a weak greediness property for equilibria. Denote a box x *full* for player B if $\sum_t B(x,t) = 1$. Also, for readability of what follows, when $\mathbb{P}$ is clear from the context, we shall denote $v_B(x,t) = v_{\mathbb{P}^{-B}}(x,t) = f(x) \cdot \prod_{A \in \mathbb{P} \setminus \{B\}} \widetilde{A}(x,t)$.

Lemma 16. *Consider a profile $\mathbb{P}_{\texttt{nash}}$ at equilibrium. For every $B \in \mathbb{P}_{\texttt{nash}}$ and t, x, y such that y is not full in B, if $B(x,t) > 0$ then $v_B(x,t) \geq v_B(y,t)$.*

Proof. Assume otherwise. Define an alternative matrix B' for player B, as $B' = B + \epsilon(\delta_{y,t} - \delta_{x,t})$. For a sufficiently small $\epsilon > 0$, B' is a doubly-substochastic matrix because y is not full in B. Then, $U_{\mathbb{P}^{-B}_{\texttt{nash}}}(B') - U_{\mathbb{P}^{-B}_{\texttt{nash}}}(B) = \epsilon(v_B(y,t) - v_B(x,t)) > 0$, in contradiction. □

Let us define a process that starts with the profile $\mathbb{P}_{\texttt{nash}}$ and changes it by a sequence of *alterations*, each shifting some amount of probability between two boxes. Importantly, we make sure that each alteration can only decrease the success probability. Hence, the proof is concluded once we show that the final profile has a success probability that is higher than that of A.

We first describe the alternations. Each alteration considers the current profile $\mathbb{P}$, and changes it to $\mathbb{P}'$. It takes some high box x, some low box y (both w.r.t. $\mathbb{P}$), and the maximal t such that there is a player $B \in \mathbb{P}$ with $B(x,t) > 0$. It defines $B' = B + \epsilon(\delta_{y,t} - \delta_{x,t})$, and lets the player that played B play B' instead. This ϵ is taken to be the largest so that x does not become low, y does not become high, and such that $\epsilon \leq B(x,t)$, so that the entries remain non-negative. Note that B' is doubly sub-stochastic, because taking care that y remains low, also means that y's row in B' still sums to less than 1.

After this alteration, either x is saturated, y is saturated, or $B'(x,t) = 0$. Clearly, in a finite number of alterations a profile $\mathbb{P}_{\texttt{final}}$ is obtained, for which either no box is high or no box is low.

Lemma 17. $\texttt{success}(\mathbb{P}_{\texttt{final}}) \geq \texttt{success}(A)$.

Proof. By Lemma 15, $\mathbb{P}_{\texttt{final}}$ can only contain high and saturated boxes, that is, for every box x, $\texttt{success}(\mathbb{P}_{\texttt{final}}, x) \geq \texttt{success}(A, x)$. However, $\texttt{success}(\mathbb{P}) = \sum_x f(x)\, \texttt{success}(\mathbb{P}, x)$, and therefore $\texttt{success}(\mathbb{P}_{\texttt{final}}) \geq \texttt{success}(A)$. □

Lastly, the following lemma concludes the proof of Theorem 6.

Lemma 18. *An alteration can only decrease the probability of success.* □

Since $A^\star$ is a symmetric equilibrium, we immediately get that for every f and T, the PoA is attained by $A^\star$, that is, $\texttt{PoA}(C_{ex}, f, T) = \max_{\mathbb{P}\in\mathcal{P}(T)} \texttt{success}(\mathbb{P})/\texttt{success}(A^\star)$. Since $A^\star$ has the best success probability among symmetric profiles, and that every policy has some symmetric equilibrium, we get Corollary 7. To make this more concrete, we show that in the worst case,

Theorem 8. *For every T, $\sup_f \texttt{PoA}(C_{\texttt{ex}}, f, T) = (1-(1-1/k)^k)^{-1}$.*

Note that as k goes to infinity the PoA converges to $e/(e-1) \approx 1.582$.

5 Future Work and Open Questions

In [1], the main complexity measure was actually the running time and not the success probability. Our results about equilibria are also relevant to this measure, but the social gain is different. For example, it is still true that $A^\star$ is an equilibrium under the exclusive policy, and that all other symmetric equilibria in the exclusive policy are equivalent to it. As $A^\star$ is optimal among symmetric profiles w.r.t. the running time, the PoSA of $C_{\texttt{ex}}$ is equal to the PoSS, and it is also the best among all policies. Furthermore, importing from [1], we know that the PoSA (w.r.t. the running time) is about 4. However, showing the analogue of Corollary 7, namely, that the PoA of $C_{\texttt{ex}}$ is that achieved by $A^\star$, seems difficult, especially because general equilibria are not necessarily greedy. Moreover, the results of Vetta [12] do not apply when analyzing the running time, and finding the PoA, PoSA, and PoSS of the sharing policy, for example, remains open.

Another interesting variant would be to consider feedback during the search. For example, assuming that a player visiting a box x knows whether or not other players were there before. Such a feedback can help in the case that the players collaborate [27], but seems to significantly complicate the analysis in the game theoretic variant.

Finally, we would like to encourage game theoretical studies of other frameworks of collaborative search, e.g., [28, 29, 30, 31].

References

[1] P. Fraigniaud, A. Korman, and Y. Rodeh, "Parallel Bayesian Search with no Coordination." https://hal.archives-ouvertes.fr/hal-01865469, To appear in *J.ACM.* (Abstract preliminary versions appeared as [**?**] and [**?**]), Aug. 2018.

[2] L.-A. Giraldeau and T. Caraco, "Social foraging theory," *Princeton University Press*, 2000.

[3] T. T. Hills, P. M. Todd, D. Lazer, A. D. Redish, I. D. Couzin, C. S. R. Group, *et al.*, "Exploration versus exploitation in space, mind, and society," *Trends in cognitive sciences*, vol. 19, no. 1, pp. 46–54, 2015.

[4] Y. Papanastasiou, K. Bimpikis, and N. Savva, "Crowdsourcing exploration," *Management Science*, 2017.

[5] U. of California Berkeley, "Boinc." `https://boinc.berkeley.edu/`, 2017.

[6] J. Kleinberg and S. Oren, "Mechanisms for (mis) allocating scientific credit," in *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pp. 529–538, ACM, 2011.

[7] M. Kennedy and R. D. Gray, "Can ecological theory predict the distribution of foraging animals? a critical analysis of experiments on the ideal free distribution," *Oikos*, vol. 68, no. 1, pp. 158–166, 1993.

[8] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic Game Theory.* Cambridge University Press, 2007.

[9] E. Koutsoupias and C. Papadimitriou, "Worst-case equilibria," *Computer Science Review*, vol. 3, no. 2, pp. 65 – 69, 2009.

[10] A. C. Pigou, "The economics of welfare," *Library of Economics and Liberty*, 1932.

[11] J. Nash, "Non-cooperative games," *Annals of Mathematics*, vol. 54, no. 2, pp. 286–295, 1951.

[12] A. Vetta, "Nash equilibria in competitive societies, with applications to facility location, traffic routing and auctions," in *Foundations of Computer Science, 2002. Proceedings. The 43rd Annual IEEE Symposium on*, pp. 416–425, IEEE, 2002.

[13] J. Y. Halpern, "A computer scientist looks at game theory," *Games and Economic Behavior*, vol. 45, no. 1, pp. 114–131, 2003.

[14] J. Y. Halpern, "Computer science and game theory: A brief survey," *CoRR*, vol. abs/cs/0703148, 2007.

[15] R. Gradwohl and O. Reingold, "Fault tolerance in large games," *Games and Economic Behavior*, vol. 86, pp. 438–457, 2014.

[16] C. Scheideler and J. T. Fineman, eds., *Proceedings of the 30th on Symposium on Parallelism in Algorithms and Architectures, SPAA 2018, Vienna, Austria, July 16-18, 2018*, ACM, 2018.

[17] M. Gairing, "Covering games: Approximation through non-cooperation," in *Internet and Network Economics, 5th International Workshop, WINE 2009, Rome, Italy, December 14-18, 2009. Proceedings*, pp. 184–195, 2009.

[18] V. Ramaswamy, D. Paccagnan, and J. R. Marden, "The impact of local information on the performance of multiagent systems," *CoRR*, vol. abs/1710.01409, 2017.

[19] Y. Mansour, A. Slivkins, and V. Syrgkanis, "Bayesian incentive-compatible bandit exploration," in *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pp. 565–582, ACM, 2015.

[20] P. Frazier, D. Kempe, J. Kleinberg, and R. Kleinberg, "Incentivizing exploration," in *Proceedings of the fifteenth ACM conference on Economics and computation*, pp. 5–22, ACM, 2014.

[21] I. Kremer, Y. Mansour, and M. Perry, "Implementing the "wisdom of the crowd"," *Journal of Political Economy*, vol. 122, no. 5, pp. 988–1012, 2014.

[22] D. Monderer and L. S. Shapley, "Potential games," *Games and economic behavior*, vol. 14, no. 1, pp. 124–143, 1996.

[23] R. W. Rosenthal, "A class of games possessing pure-strategy nash equilibria," *International Journal of Game Theory*, vol. 2, no. 1, pp. 65–67, 1973.

[24] S. Albers and M. Hellwig, "Online makespan minimization with parallel schedules," *CoRR*, vol. abs/1304.5625, 2013.

[25] A. Czumaj and B. Vöcking, "Tight bounds for worst-case equilibria," *ACM Trans. Algorithms*, vol. 3, pp. 4:1–4:17, Feb. 2007.

[26] S. Bhattacharya, S. Im, J. Kulkarni, and K. Munagala, "Coordination mechanisms from (almost) all scheduling policies," in *Innovations in Theoretical Computer Science, ITCS'14, Princeton, NJ, USA, January 12-14, 2014*, pp. 121–134, 2014.

[27] S. Dobrev, R. Královic, and D. Pardubská, "Treasure hunt with barely communicating agents," in *21st International Conference on Principles of Distributed Systems, OPODIS 2017, Lisbon, Portugal, December 18-20, 2017*, pp. 14:1–14:16, 2017.

[28] D. Dereniowski, Y. Disser, A. Kosowski, D. Pajak, and P. Uznanski, "Fast collaborative graph exploration," *Inf. Comput.*, vol. 243, pp. 37–49, 2015.

[29] Y. Emek, T. Langner, J. Uitto, and R. Wattenhofer, "Solving the ANTS problem with asynchronous finite state machines," in *Automata, Languages, and Programming - 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part II*, pp. 471–482, 2014.

[30] O. Feinerman and A. Korman, "The ANTS problem," *Distributed Computing*, vol. 30, no. 3, pp. 149–168, 2017.

[31] Y. Higashikawa, N. Katoh, S. Langerman, and S.-I. Tanigawa, "Online graph exploration algorithms for cycles and trees by multiple searchers," *J. Comb. Optim.*, vol. 28, pp. 480–495, Aug. 2014.

[32] S. D. Fretwell and H. L. L. Jr., "On territorial behavior and other factors influencing habitat distribution in birds," *Acta biotheoretica*, vol. 19, no. 1, pp. 16–32, 1969.

[33] T. Tregenza, "Building on the ideal free distribution," *Adv. Ecol. Res*, pp. 253–307, 1995.

[34] A. Nicholson, "An outline of the dynamics of animal populations.," *Australian Journal of Zoology*, pp. 9–65, 1954.

[35] N. Pinter-Wollman, T. Dayan, D. Eilam, and N. Kronfeld-Schor, "Can aggression be the force driving temporal separation between competing common and golden spiny mice?," *Journal of Mammalogy*, vol. 87, no. 1, pp. 48–53, 2006.

[36] R. A. Horn, *Topics in Matrix Analysis.* New York, NY, USA: Cambridge University Press, 1986.

[37] G. Birkhoff, "Three observations on linear algebra," *Univ. Nac. Tacuman, Rev. Ser. A*, vol. 5, pp. 147–151, 1946.

A Animals Searching for Food

The way animals disperse in their environment is a cornerstone of ecology [2]. In these contexts, dispersal is typically governed by two contradicting forces: The bias towards selecting the higher quality patches, and the need to avoid costly collisions or overlaps. In the ecology discipline, the (single-round) setting of animals competing over patches of resources has been extensively studied through the concept of *Ideal Free Distribution* [32, 7, 33]. Unfortunately, however, although many animals engage in search over multiple sites, much less is known about relevant game theoretical aspects is such multi-round settings.

In animal contexts, increased collision costs (the analogy to reward policies) can be caused by various factors, including aggressive behavior, or merely due to equally sharing the patch between the colliding individuals (a.k.a., scramble competition [34]). Plausibly, collision costs have emerged by evolution due to dynamics that is governed by multiple parameters. Without excluding other factors, one of the possible evolutionary driving forces may relate to competition between groups. Indeed, from the perspective of the group, consuming large quantities of food by all members together (coverage, in our terminology) can indirectly increase the fitness of individuals and hence become significant for their survival. To see why, consider for example, a setting in which two species compete over the same patched food resource, each acting in a different time period of the day (so there is no direct interaction between the two species) [35]. Assume that one species is more aggressive towards conspecifics. All other factors being similar, at first glance, it may appear that this species would be inferior to the more peaceful one as it induces unnecessary waste of energy and risks of injury. However, perhaps counter-intuitively, our results suggest that it might actually be the converse. Indeed, the higher collision costs of the aggressive species may drive its members to better cover the food resources, on the expense of the more peaceful species.

B On the Difficulty of the Multi-Round Game.

The Multi-Round Game is Not a Potential Game. It is interesting to note that the single round game is an exact potential game, yet the multi-round game is not. Indeed, for the single round, assume that $\mathbb{P}$ is a deterministic profile, and let $x_{\mathbb{P}}$ be the number of players that choose box x in $\mathbb{P}$. Denote

$$\Phi(\mathbb{P}) = \sum_x f(x) \sum_{\ell=1}^{x_{\mathbb{P}}} C(\ell).$$

If a player changes strategy and chooses (deterministically) some box y instead of box x, then the change in its utility is $f(y)C(y_{\mathbb{P}}+1) - f(x)C(x_{\mathbb{P}})$. This is also the change that $\Phi(\mathbb{P})$ sees. This extends naturally to mixed strategies, and so the single round game is a potential game. This observation has several consequences, and in particular that there always exists a pure Nash equilibrium.

On the other hand, the multi-round game does not always have a pure equilibrium, and so is not a potential game. For example, the following holds for any policy C. Consider the case where $M = 3$, $T = 2$, $k = 2$, and all boxes have $f(x) = 1/3$. Note that $C(2) < 1$ since here $k = 2$ and $C \not\equiv 1$. Assume there is some deterministic profile that is at equilibrium, and w.l.o.g. assume player 1's first pick is box 1. There are two cases:

1. Player 1 picks it again in the second round. Player 2's strictly best response is to pick box 2 and then 3 (or the other way around). In this case, player 1 would earn more by first picking box 3 (box 2) and then box 1. In contradiction.

2. Player 1 picks a different box in the second round. W.l.o.g. assume it is box 2. Player 2's strictly best response is to first take box 2 and then take box 3. However, player 1 would then prefer to start with box 3 and then box 1. Again a contradiction.

Optimal Profiles may not be at Equilibrium. A second notable difference concerns profiles that maximize the success probability. In the single-round game, when $M \geq k$, the success probability is maximized when each player exclusively visits one box in $1, 2, \ldots, k$ with probability 1. Under the exclusive policy, for example, such a profile is also at equilibrium. In fact, if $f(k) \geq f(1)/2$ then the same is true also for the sharing policy. For the multi-round setting, when $M \geq Tk$, an optimal scenario is also achieved by a deterministic profile, e.g., when player i visits box $i + (t-1)k$ in round t. However, this profile would typically not be an equilibrium, even under the exclusive policy. Indeed, when $f(x)$ is strictly decreasing, player 2 for example, can gain more by stealing box $k+1$ from player 1 in the first round, then safely taking box 2 in the second round, and continuing from there as scheduled originally. This shows that in the multi-round game, the best equilibrium has only sub-optimal success probability.

C Every Doubly-Substochastic Matrix is a Probability Matrix

A matrix is called *doubly-substochastic* if it is non-negative, and each of its rows and columns sum to at most 1. Also, a doubly-substochastic matrix is a *partial permutation* if it consists of only 0 and 1 values. The following is a generalization of the Birkhoff - von Neumann theorem, proved for example in [36].

Theorem 19. *A matrix is doubly-substochastic iff it is a convex combination of partial permutations.*

Furthermore, Birkhoff's construction [37] finds this decomposition in polynomial time, and guarantees it contains at most a number of terms as the number of positive elements of the matrix. The generalization of [36] does not change this claim significantly, as it embeds the doubly-substochastic matrix in a doubly-stochastic one which is at most 4 times larger.

Corollary 20. *If matrix A is doubly-substochastic then there is some strategy such that A is its probability matrix. Furthermore, this strategy can be found in polynomial time, and is implementable as a polynomial algorithm.*

Proof. First note that the claim is true if A is a partial permutation. The strategy in this case will be a deterministic strategy, which may sometimes choose not to visit any box. In the general case, Theorem 19 states that there exist $\theta_1, \theta_2, \ldots, \theta_k$, such that $\sum_{i=1}^{k} \theta_i = 1$ and partial permutations $A_1, A_2, \ldots, A_k$, such that $A = \sum_{i=1}^{k} \theta_i A_i$. As mentioned, each A_i is the probability matrix of some strategy B_i. Define the following strategy B as follows: with probability θ_i run strategy B_i. Then, the probability matrix of B is $\sum_i \theta_i A_i = A$, as required. $\square$

D Proofs for Section 3 - General Policies

A first general observation is that if a box has some probability of not being chosen, then it has a positive value. This is clear from the definition of utility, from the fact that $C(1) = 1$ and because C is non-negative.

Observation 21. *If for all $A \in \mathbb{P}$, $\widetilde{A}(x,t) > 0$, then $v_{\mathbb{P}}(x,t) > 0$.*

D.1 Non-Redundancy

First a simple observation:

Observation 22. *If A is non-redundant then for all $t \geq M$, and for all x, $\widetilde{A}(x,t) = 0$.*

Lemma 10. *If a profile $\mathbb{P}$ is at equilibrium and* $\texttt{success}(\mathbb{P}) < 1$ *then every player is non-redundant.*

Proof. As $\texttt{success}(\mathbb{P}) < 1$, there is some x such that $\prod_{B \in \mathbb{P}} \widetilde{B}(x,T) > 0$, and so for all $B \in \mathbb{P}$, $\widetilde{B}(x,T) > 0$. Fix such an x. Assume that some player plays a redundant matrix A. This means that there is some time s where $\sum_y A(y,s) < 1$. Define a new matrix $A' = A + \epsilon \delta_{x,s}$, for a small $\epsilon > 0$. Taking it small enough will ensure that A' is doubly-substochastic since (1) in column s, there is space because of the redundancy of A at time s, and (2) in row x there is space because $\sum_t A(x,t) = 1 - \widetilde{A}(x,T) < 1$.

Therefore our player can play according to A' instead of A, and

$$U_{\mathbb{P}^{-A}}(A') - U_{\mathbb{P}^{-A}}(A) = \epsilon \cdot v_{\mathbb{P}^{-A}}(x,s).$$

Recall that by how we chose x, for all $B \in \mathbb{P}^{-A}$, $\widetilde{B}(x,T) > 0$, and as $\widetilde{B}$ is weakly decreasing in t, $\widetilde{B}(x,s) > 0$. By Observation 21, $v_{\mathbb{P}^{-A}}(x,s) > 0$, and so the utility strictly increases, in contradiction to $\mathbb{P}$ being at equilibrium. $\square$

D.2 Monotonicity

Lemma 11. *Consider two doubly-substochastic matrices A and B. If $A(x,t) > B(x,t)$, and for all $s < t$, $A(x,s) = B(x,s)$ then $v_A(x,t) < v_B(x,t)$.*

Proof. Denote:

$$\Psi(p,n) = \sum_{\ell=0}^{k-1} C(\ell+1) \binom{k-1}{\ell} p^{\ell} n^{k-1-\ell}.$$

Then,

$$v_A(x,t) = f(x)\Psi(A(x,t), \widetilde{A}(x,t)), \tag{2}$$

and similarly for B. Denote $R(\ell) = \binom{k-1}{\ell} p^{\ell} n^{k-1-\ell}$, and $Q(\ell) = \sum_{s=0}^{\ell} R(\ell)$.

$$\begin{aligned}
\Psi(p,n) &= C(1)R(0) + C(2)R(1) + \ldots + C(\ell)R(\ell-1) \\
&= C(1)Q(0) + C(2)(Q(1) - Q(0)) + \ldots + C(k)(Q(k-1) - Q(k-2)) \\
&= (C(1) - C(2))Q(0) + \ldots + (C(k-1) - C(k))Q(k-2) + C(k)Q(k-1)
\end{aligned}$$

First, by the properties of congestion policies, all the $C(\ell) - C(\ell+1) \geq 0$, and, since $C \not\equiv 1$, then at least one is strictly positive. Now,

$$Q(\ell) = \sum_{s=0}^{\ell} \binom{k-1}{s} p^s n^{k-1-s} = (p+n)^{k-1} \sum_{s=0}^{\ell} \binom{k-1}{s} \left(\frac{p}{p+n}\right)^s \left(\frac{n}{p+n}\right)^{k-1-s}. \tag{3}$$

The sum is the probability that there are at most ℓ ones out of $k-1$ Bernoulli random variables sampled with probability $p/(p+n)$. Therefore it is strictly decreasing in $p/(p+n)$ for any $\ell < k-1$. In the case of A,

$$p + n = \widetilde{A}(x,t) + A(x,t) = \widetilde{A}(x,t-1).$$

For B it is the same, and by the conditions of the lemma, they are equal. As p is larger in A, we get that $Q(\ell)$ is strictly larger in B for any $\ell < k-1$. By the interpretation of the sum in Eq. (3), $Q(k-1) = (p+n)^{k-1}$, hence this term is the same in both A and B. Therefore, $\Psi(p,n)$ is strictly larger in B, and so $v_B(x,t) > v_A(x,t)$. □

Lemma 12. *Let A be doubly-substochastic. For every x and t, $v_A(x,t+1) \leq v_A(x,t)$. Moreover, if $A(x,t+1) > 0$ then the inequality is strict.*

Proof. First Assume $A(x,t+1) = 0$. The value at round $t+1$ is:

$$v_A(x,t+1) = f(x) \sum_{\ell=0}^{k-1} C(\ell+1) \binom{k-1}{\ell} A(x,t+1)^{\ell} \widetilde{A}(x,t+1)^{k-\ell-1} = f(x)C(1)\widetilde{A}(x,t+1)^{k-1}$$

On the other hand,

$$v_A(x,t) = f(x) \sum_{\ell=0}^{k-1} C(\ell+1) \binom{k-1}{\ell} A(x,t)^{\ell} \widetilde{A}(x,t)^{k-\ell-1} \geq f(x)C(1)\widetilde{A}(x,t)^{k-1},$$

because all the $C(\ell) \geq 0$. As $\widetilde{A}(x,t) = \widetilde{A}(x,t+1)$, we get that $v_A(x,t+1) \leq v_A(x,t)$, as required.

Next, assume $\widetilde{A}(x,t) > 0$. Consider strategy A and let B be the same as A except that for every $s > t$, $B(x,s) = 0$. By the above, $v_B(x,t+1) \leq v_B(x,t) = v_A(x,t)$. By Lemma 11,

$$v_A(x,t+1) < v_B(x,t+1),$$

and we conclude. □

D.3 Greediness

For a profile $\mathbb{P}$ and some strategy A, denote $v_{\mathbb{P}}(A,t) = \min_x \{v_{\mathbb{P}}(x,t) \mid A(x,t) > 0\}$. Clearly, if A is non-redundant then $U_{\mathbb{P}}(A,t) \geq v_{\mathbb{P}}(A,t)$. For symmetric profiles, we simply denote $v_A(t)$ for $v_{A^{k-1}}(A,t)$. A simple and useful observation is,

Observation 23. *Let A be greedy at time t w.r.t. profile $\mathbb{P}$. If $A(x,t) < \widetilde{A}(x,t-1)$ then $v_{\mathbb{P}}(x,t) \leq v_{\mathbb{P}}(A,t)$.*

Proof. Assume otherwise, that is, $v_{\mathbb{P}}(x,t) > v_{\mathbb{P}}(A,t)$. Take some y s.t. $v_{\mathbb{P}}(y,t) = v_{\mathbb{P}}(A,t)$, and let B coincide with $A + \epsilon(\delta_{x,t} - \delta_{y,t})$ for the first t columns, and 0 for the rest. Taking $\epsilon > 0$ small enough makes B doubly-substochastic. But

$$U_{\mathbb{P}}(B,t) - U_{\mathbb{P}}(A,t) = \epsilon\left(v_{\mathbb{P}}(x,t) - v_{\mathbb{P}}(y,t)\right) > 0,$$

contradicting the greediness of A. □

D.3.1 Equivalent Definition for Greediness

Lemma 14. *A non-redundant strategy A is greedy w.r.t. $\mathbb{P}$ at time t iff for every x and y, if $A(x,t) > 0$ and $v_{\mathbb{P}}(x,t) < v_{\mathbb{P}}(y,t)$ then $\widetilde{A}(y,t) = 0$.*

Proof. If A is greedy at time t, then assume by contradiction, that there are some x and y where $A(x,t) > 0$, $v_{\mathbb{P}}(x,t) < v_{\mathbb{P}}(y,t)$, and yet $\widetilde{A}(y,t) > 0$. This means that $A(y,t) < \widetilde{A}(y,t-1)$, and so by Observation 23, $v_{\mathbb{P}}(y,t) \leq v_{\mathbb{P}}(A,t) \leq v_{\mathbb{P}}(x,t)$, in contradiction.

Conversely, if the condition in the lemma is true, then let B be some matrix as required by the definition of greediness. We partition the set of boxes into two sets. *High-value* boxes are those x such that $v_{\mathbb{P}}(x,t) > v_{\mathbb{P}}(A,t)$, and *low-value* boxes are all the rest. By the condition in the lemma, we know that for every high-value box x, $\widetilde{A}(x,t) = 0$, and hence $A(x,t)$ is set to its maximal quantity $\widetilde{A}(x,t-1)$. Therefore, for each such box, its contribution to the utility of A is at least as large as its contribution to B's utility, i.e., $A(x,t)v_{\mathbb{P}}(x,t) \geq B(x,t)v_{\mathbb{P}}(x,t)$. The total contribution of high-value boxes to the utility of A is therefore also at least as large as their contribution to B's utility. On the other hand, the total contribution of low-value boxes to the utility of B is at most $v_{\mathbb{P}}(A,t)$ times the probability that B goes to a low-value box in round t. Since $v_{\mathbb{P}}(A,t)$ is the lowest value that A checks, and since A is non-redundant, we get that in total, $\sum_x A(x,t)v_{\mathbb{P}}(x,t) \geq \sum_x B(x,t)v_{\mathbb{P}}(x,t)$, as required. □

D.3.2 An Example where a Non-Redundant Sgreedy Strategy is not an Equilibrium

Consider the sharing policy, and take $k = 2$, $T = 2$, and $f(1) = 4/7$, $f(2) = 3/7$. Denote by A the matrix of the sgreedy non-redundant strategy guaranteed to exist by Lemma 24. By non-redundancy, we know $A = \left(\begin{smallmatrix} p & 1-p \\ 1-p & p \end{smallmatrix}\right)$. Which gives values:

$$v_A = \begin{pmatrix} \frac{4}{7}\left(1 - p + \frac{p}{2}\right) & \frac{4}{7} \cdot \frac{1-p}{2} \\ \frac{3}{7}\left(p + \frac{1-p}{2}\right) & \frac{3}{7} \cdot \frac{p}{2} \end{pmatrix} = \frac{1}{14} \begin{pmatrix} 8\left(1 - \frac{p}{2}\right) & 4(1-p) \\ 3\left(1+p\right) & 3p \end{pmatrix}$$

Let us put the $1/14$ aside as it only changes utilities by a multiplicative constant. Setting $p = 1$ gives $v_A = \left(\begin{smallmatrix} 4 & 0 \\ 6 & 3 \end{smallmatrix}\right)$. This means that A is not sgreedy, because in $t = 1$, choosing the second box gives a higher utility for this round than A gets. Setting $p = 0$ gives $v_A = \left(\begin{smallmatrix} 8 & 4 \\ 3 & 0 \end{smallmatrix}\right)$. Again, A is not sgreedy, because choosing box 1 in the first round improves the utility at this round.

Otherwise, $0 < p < 1$, and since A is sgreedy, then $v_A(1,1) = v_A(2,1)$. This means:

$$8(1 - p/2) = 3(1 + p) \implies 5 = 7p \implies p = 5/7.$$

Therefore, $v_A = \begin{pmatrix} 36/7 & 8/7 \\ 36/7 & 15/7 \end{pmatrix}$. Let strategy B first choose box 1 and then box 2. Then,

$$\frac{U_A(B)}{U_A(A)} = \frac{36 + 15}{36 + \frac{2}{7}8 + \frac{5}{7}15} > 1.$$

So A is not a symmetric equilibrium.

D.3.3 Existence and Uniqueness of SGreedy Strategies

Lemma 24. *For every policy C there exists a non-redundant sgreedy strategy A. Moreover, all such strategies are equivalent.*

Proof. As our construction will be non-redundant, for any $t > M$, all boxes will already be checked, and so setting $A(x,t) = 0$ is fine. Fix a time $t \leq \min\{T, M\}$. Assume that we already defined all the $A(x,s)$ for $s < t$ such that A is non-redundant and sgreedy for all such s. We will find values for $A(x,t)$ that are the same, and thus prove the lemma by induction.

Consider $v_A(x,t)$ as a function of $A(x,t)$, with $\widetilde{A}(x,t-1) = 1 - \sum_{s<t} A(x,s)$ fixed. By Lemma 12, it is a strictly decreasing function, it is continuous, and is defined between 0 and $\widetilde{A}(x,t-1)$. This means it has a continuous inverse function with this range. Denote by $A_w(x,t)$ the extended inverse. That is, given w, it is the $A(x,t)$ such that setting it gives $v_A(x,t) = w$ if such $A(x,t)$ exists. If this is not possible, it means either w is too large, and so $A_w(x,t) = 0$, or too small, and then $A_w(x,t) = \widetilde{A}(x,t-1)$.

Next, we choose w so that $\sum_x A_w(x,t) = 1$. This would be needed to ensure that A is doubly-substochastic and non-redundant. For that, we note that $\sum_x A_0(x,t) = \sum_x \widetilde{A}(x,t-1) \geq 1$, as $t-1 < M$. On the other hand, 1 is larger than all $f(x)$, and so is an upper bound on the values boxes can take. Therefore, $\sum_x A_1(x,t) = 0$. By the fact that the $A_w(x,t)$ are a continuous function of w, this means there is some w such that $\sum_x A_w(x,t) = 1$. Moreover, it is easy to see (no matter what w is) that these $A_w(x,t)$'s satisfy the greediness condition of Lemma 14. In other words, setting $A(x,t) = A_w(x,t)$, extends A to time t in that it is sgreedy and non-redundant, as required.

Next we prove uniqueness. Assume by contradiction that $A \neq B$ and both are sgreedy and non-redundant. Take the first column $t \geq 1$ where they differ. As the matrices are non-redundant, each column sums to 1, and therefore there is some x where $A(x,t) > B(x,t)$ and some y where $B(y,t) > A(y,t)$, which by Lemma 11, imply that $v_A(x,t) < v_B(x,t)$ and $v_B(y,t) < v_A(y,t)$. Now,

$$v_A(t) \leq v_A(x,t) < v_B(x,t) \leq v_B(t),$$

where the last inequality is because $B(x,t) < A(x,t) \leq \widetilde{A}(x,t-1) = \widetilde{B}(x,t-1)$, and so Observation 23 applies. On the other hand, in the same manner,

$$v_B(t) \leq v_B(y,t) < v_A(y,t) \leq v_A(t),$$

in contradiction. □

D.3.4 Symmetric Equilibirum is Sgreedy when $C(k) = 0$

Claim 3. *For any policy such that $C(k) = 0$, all symmetric equilibria are equivalent.*

Proof. Let A be some symmetric equilibrium. We will show it is non-redundant and sgreedy, and so by Theorem 2 we get that it is unique up to equivalence. By Lemma 13, A is non-redundant, and we may assume that $T \leq M$. If A is not sgreedy, then let x, y and t consist of a counter example, i.e., $A(x,t) > 0$ and yet $v_A(x,t) < v_A(y,t)$, where $\widetilde{A}(y,t) > 0$. There are two cases, and in both we will construct some doubly-substochastic matrix B such that $U_A(B) > U_A(A)$, thus contradicting the claim that A is a symmetric equilibrium.

Case 1. $\sum_{s \leq T} A(y,s) < 1$. In this case, let $B = A + \epsilon(\delta_{y,t} - \delta_{x,t})$. Taking a small enough ϵ ensures that B is a doubly-substochastic matrix, since (1) y's row sums to strictly less than 1 in A, (2) the sum of column t is the same as in A, and (3) x's row only decreased in value compared to A. Lastly,

$$U_A(B) - U_A(A) = \epsilon(v_A(y,t) - v_A(x,t)) > 0.$$

Case 2. $\sum_{s \leq T} A(y,s) = 1$. This means that there is some first t', where $\widetilde{A}(y,t') = 0$. As $\widetilde{A}(y,t) > 0$, we have $t' > t$. Now:

$$v_A(y,t') = f(y) \sum_{\ell=0}^{k-1} C(\ell+1) \binom{k-1}{\ell} A(y,t')^{\ell} \widetilde{A}(y,t')^{k-\ell-1} = f(y) C(k) A(y,t')^{k-1} = 0,$$

because $C(k) = 0$. Define:

$$B = A + \epsilon(-\delta_{x,t} + \delta_{y,t} - \delta_{y,t'}).$$

Taking a small enough ϵ, as both $A(x,t)$ and $A(y,t')$ are strictly positive, B is non-negative. As A is doubly-substochastic, B is doubly-substochastic as well.

$$U_A(B) - U_A(A) = \epsilon(v_A(y,t) - v_A(x,t) - v_A(y,t')).$$

However, $v_A(y,t') = 0$, and $v_A(y,t) > v_A(x,t)$, and so this quantity is strictly positive. □

E Constructing Approximate Equilibria

Our goal in this section is to prove the following.

Theorem 4. *Given $\theta > 0$, there exists an algorithm that takes as input a configuration, and produces a symmetric $(1 + C(k))(1 + \theta)$-equilibrium. The algorithm runs in polynomial time in T, k, M, $\log(1/\theta)$, $\log(1/(1 - C(k)))$, and $\log(1/f(M))$.*

Towards that, we define approximate notions on non-redundancy and greediness.

Definition 25. We say A is *ϵ-sgreedy* at time t if whenever $A(x,t) > 0$ and $v_A(y,t) > v_A(x,t) + \epsilon$ then $\widetilde{A}(y,t) = 0$. It is ϵ-sgreedy if it is ϵ-sgreedy for each time t.

Definition 26. Let $0 \leq \delta \leq 1$. We say A is *δ-redundant* if for every $t \leq M$, $\sum_x A(x,t) \geq 1 - \delta$.

We then prove the following two lemmas (in Appendices E.1 and E.2 below)

Lemma 27. *If A is δ-redundant and ϵ-sgreedy then for any B,*

$$U_A(B) \leq \frac{1+C(k)}{1-\delta} U_A(A) + (T+1)\epsilon.$$

Lemma 28. *There is an algorithm that given a configuration, $\epsilon > 0$ and $\delta > 0$, finds a matrix that is δ-redundant and ϵ-sgreedy. The algorithm runs in polynomial time in parameters T, k, M, $\log(1/\epsilon)$, $\log(1/\delta)$, $\log(1/(1-C(k)))$, and $\log(1/f(M))$.*

Using these two lemmas, the theorem is not difficult to prove.

Proof. (of Theorem 4) First, if $M = 1$, the strategy at equilibrium would be to pick box 1 at $t = 1$. We therefore assume $M \geq 2$. Set $\epsilon = \theta/2(T+1) \cdot f(2)/2^{k+1}$ and $\delta = \frac{\theta/2}{1+\theta/2}$. Hence, $1/(1-\delta) = 1+\theta/2$. We use Lemma 28 to construct a matrix A which is δ-redundant and ϵ-sgreedy. As $\delta = \Theta(\theta)$, this construction takes polynomial time as required. Also, according to Appendix C, this strategy is implementable by a polynomial algorithm in the size of the matrix. According to Lemma 27, for any strategy B,

$$U_A(B) \leq (1 + C(k)) \left(1 + \frac{\theta}{2}\right) U_A(A) + \frac{\theta}{2} \cdot \frac{f(2)}{2^{k+1}}. \tag{4}$$

As $M \geq 2$, then either $A(1,1) \leq 1/2$ or $A(2,1) \leq 1/2$. As $v_A(x,1) \geq f(x)C(1)\widetilde{A}(x,1)^{k-1} = f(x)\widetilde{A}(x,1)^{k-1}$, we get that $\texttt{maxv}(1) \geq f(2)/2^{k-1}$. Since $\epsilon < f(2)/2^k$, and δ can be assumed to be at most $1/2$, we get that $U_A(A) \geq f(2)/2^{k+1}$. Therefore,

$$\frac{\theta}{2} \cdot \frac{f(2)}{2^{k+1}} \leq (1 + C(k)) \frac{\theta}{2} U_A(A),$$

which combined with Eq. (4) gives the result. □

E.1 Approximate Redundant and Sgreedy implies Approximate Equilibrium

Recall that $v_A(t) = \min \{v_A(x,t) \mid A(x,t) > 0\}$. This is the minimal value gained by A in round t when played against $k-1$ other players playing A.

Lemma 27. *If A is δ-redundant and ϵ-sgreedy then for any B,*

$$U_A(B) \leq \frac{1+C(k)}{1-\delta} U_A(A) + (T+1)\epsilon.$$

Proof. The proof follows the same steps as that of Theorem 2, except it deals with the approximate redundancy and sgreediness. Consider a strategy B. We compare the utility of B versus that of A when both play against $k-1$ players playing A.

By Lemma 12, $v_A(x,t)$ is non-increasing in t, and so we can assume w.l.o.g. that B is non-redundant, and hence does not choose any box beyond time M. Therefore, we shall let A run at most M rounds, and prove the result assuming this new A and $T \leq M$. This can only decrease the r.h.s. in the statement of the theorem, and thus is enough.

Denote $\texttt{maxv}(t) = \max_x v_A(x,t)$. Since the utility of B in any round t is a convex combination of $v_A(x,t)$, we have:

$$U_A(B,t) \leq \texttt{maxv}(t). \tag{5}$$

Definition 29. We say that A *fills* box x at round t if $A(x,t) > 0$ and $\widetilde{A}(x,t) = 0$.

Claim 30. *If A does not fill any box at round t then $U_A(A,t) \geq (1-\delta)(\texttt{maxv}(t) - \epsilon)$.*

Proof. We first argue that under the assumption of the claim, for every x such that $A(x,t) > 0$, we have $v_A(x,t) \geq \texttt{maxv}(t) - \epsilon$. Indeed, assume by contradiction that $v_A(x,t) < v_A(y,t) - \epsilon$ for some y. Since A is ϵ-sgreedy then $\widetilde{A}(y,t) = 0$. By our assumption, this means that $A(y,t) = 0$. This implies that $v_A(y,t) = 0$, in contradiction.

The claim then follows by δ-redundancy, as $U_A(A,t)$ is a convex combination of such $v_A(x,t)$'s where the coefficients sum to at least $1-\delta$. □

The conclusion in Claim 30 holds for the first round without any condition:

Claim 31. $U_A(A,1) \geq (1-\delta)(\mathtt{maxv}(1)-\epsilon)$.

Proof. If the condition of Claim 30 does not hold, it means that there is an x where $A(x,1)=1$, and so, since A is ϵ-sgreedy, $v_A(x,1) \geq v_A(y,1)-\epsilon$ for all other y's. Therefore, in this case $U_A(A,1) = v_A(x,1) \geq \mathtt{maxv}(1)-\epsilon \geq (1-\delta)(\mathtt{maxv}(1)-\epsilon)$. □

For the other rounds, we cannot prove the same, but looking at previous rounds, it is true:

Claim 32. *For any $s<t$, $U_A(A,s) \geq (1-\delta)(\mathtt{maxv}(t)-\epsilon)$.*

Proof. We will show that for any $s<t$, and every x, $U_A(A,s) \geq (1-\delta)(v_A(x,t)-\epsilon)$. Take some x. If $v_A(x,t)=0$, then the claim is clear. Otherwise, $A(x,t)>0$ or $\widetilde{A}(x,t)>0$ or both. Either way, $\widetilde{A}(x,s)>0$. Therefore, as A is ϵ-sgreedy, for every y such that $A(y,s)>0$,

$$v_A(y,s) \geq v_A(x,s)-\epsilon \geq v_A(x,t)-\epsilon,$$

where the last inequality is because $v_A(x,\cdot)$ is non-increasing (Lemma 12). As A is δ-redundant, $U_A(A,s)$ is a sum of such $v_A(y,s)$'s with coefficients that sum to at least $1-\delta$. Hence $U_A(A,s) \geq (1-\delta)(v_A(x,t)-\epsilon)$, as required. □

Lastly, we claim that the utility decreases considerably between rounds for the cases we are interested in:

Claim 33. *Assume that A fills box x in time $t>1$. Then, for any time $s<t$, $v_A(x,t) \leq C(k)v_A(x,s)$.*

Proof. First,

$$v_A(x,s) \geq f(x)C(1)\widetilde{A}(x,s)^{k-1} = f(x)\widetilde{A}(x,s)^{k-1}$$

On the other hand, since $\widetilde{A}(x,t)=0$,

$$v_A(x,t) = f(x)C(k)A(x,t)^{k-1} \leq f(x)C(k)\widetilde{A}(x,s)^{k-1},$$

because $A(x,t) \leq \widetilde{A}(x,t-1) \leq \widetilde{A}(x,s)$. Combining the above two inequalities gives the result. □

Denote by X_1 the set of rounds for which there is no box x that is filled by A. Let X_2 be the rest of the rounds, except for $t=1$ which is in neither. Also denote $t_0 = \min X_2$. By Eq. (5) and Claims 30, 31 and 32:

$$\begin{aligned}
U_A(B) &\leq \sum_{t\in X_1\cup\{1\}} \mathtt{maxv}(t) + \mathtt{maxv}(t_0) + \sum_{t\in X_2\setminus\{t_0\}} \mathtt{maxv}(t) \\
&\leq \sum_{t\in X_1\cup\{1\}} \left(\frac{U_A(A,t)}{1-\delta}+\epsilon\right) + \mathtt{maxv}(t_0) + \sum_{t\in X_2} \left(\frac{U_A(A,t)}{1-\delta}+\epsilon\right) \\
&= \frac{U_A(A)}{1-\delta} + \epsilon T + \mathtt{maxv}(t_0).
\end{aligned}$$

By Claims 33 and 31,

$$\begin{aligned}
\mathtt{maxv}(t_0) &= \max_x v_A(x,t_0) \leq \max_x C(k)v_A(x,1) \\
&\leq C(k)\left(\frac{U_A(A,1)}{1-\delta}+\epsilon\right) \leq C(k)\left(\frac{U_A(A)}{1-\delta}+\epsilon\right).
\end{aligned}$$

Therefore,

$$U_A(B) \leq \frac{1+C(k)}{1-\delta}U_A(A) + (T+1)\epsilon.$$

□

E.2 Polynomial Construction of Approximate Sgreedy and Non-redundancy Strategy

Lemma 28. *There is an algorithm that given a configuration, $\epsilon > 0$ and $\delta > 0$, finds a matrix that is δ-redundant and ϵ-sgreedy. The algorithm runs in polynomial time in parameters T, k, M, $\log(1/\epsilon)$, $\log(1/\delta)$, $\log(1/(1-C(k)))$, and $\log(1/f(M))$.*

Proof. We will say a quantity is polynomial if it is as stated in the lemma. First, if $T > M$, we set $A(\cdot, t) = 0$ for any $t > M$. This trivially satisfies both δ-redundancy and ϵ-sgreediness for these rounds. For $t \leq M$, we show how to calculate $A(x,t)$'s assuming all of $A(x,s)$'s for $s < t$ are already calculated.

This proof is a constructive version of the proof of Lemma 24, and so they bear many similarities. Consider $v_A(x,t)$ as a function of $A(x,t)$, with $\widetilde{A}(x,t-1) = 1 - \sum_{s<t} A(x,s)$ fixed. As we know it is a continuous strictly decreasing function, and is defined between 0 and $\widetilde{A}(x,t-1)$. Denote by $A_w(x,t)$ the extended inverse. That is, given w, it is the $A(x,t)$ such that setting it gives $v_A(x,t) = w$ if such $A(x,t)$ exists. If this is not possible, it means either w is too large, and so $A_w(x,t) = 0$, or too small, and then $A_w(x,t) = \widetilde{A}(x,t-1)$.

The idea is to do a binary search for a good enough $v_A(t)$. Denote our current guess as w. We need some procedure to say whether w is too large, too small, or sufficient as a guess for $v_A(t)$. For this purpose, we find $A(x,t)$'s that approximate the $A_w(x,t)$'s. That is,

1. If assigning $A(x,t) = 0$ gives $v_A(x,t) \leq w$, we set $A(x,t) = 0 = A_w(x,t)$.
2. If assigning $A(x,t) = \widetilde{A}(x,t)$ gives $v_A(x,t) \geq w$, we set $A(x,t) = \widetilde{A}(x,t) = A_w(x,t)$.
3. Otherwise, we use binary search to find $A(x,t)$ such $|A(x,t) - A_w(x,t)| < \delta/4M$. As $v_A(x,t)$ as a function of $A(x,t)$ is strictly monotone and continuous this simply involves running a binary search, comparing $v_A(x,t)$ to w, until the size of the interval of $A(x,t)$'s we consider is less than $\delta/4M$. This can be done polynomially.

   In fact, we run the binary search even more, so as to be able to guarantee that $|v_A(x,t) - w| < \epsilon/2$. In the terminology of Lemma 34, $v_A(x,t) = f(x)\phi(A(x,t))$, and so, if the end points of our current interval are $A_1(x,t)$ and $A_2(x,t)$, then:

$$|f(x)\phi(A_1(x,t)) - f(x)\phi(A_2(x,t))| \leq f(1)4^{k-1}|A_1(x,t) - A_2(x,t)|.$$

   Thus, taking $O(k + \log(1/\epsilon))$ steps of the binary search can make this at most $\epsilon/2$. As w is between the values $v_A(x,t)$'s we get for $A_1(x,t)$ and $A_2(x,t)$, we can conclude our search.

Surely, this gives an ϵ-sgreedy matrix at time t (for now it is not doubly-substochastic, but this will be fixed soon). Now, consider the sum of the $A(x,t)$'s we got:

1. If it is between $1-\delta$ and 1 we are done.
2. If it is greater than 1 then we considered w as too small, and continue with the next step of binary search.
3. If it is smaller than $1-\delta$, then we say w is too large and continue.

If the process concludes, then the A we get is doubly-substochastic by the fact that always $A(x,t) \in [0, \widetilde{A}(x,t)]$, $\sum_x A(x,t) \leq 1$, and for $t > M$, $A(x,t) = 0$. As this A is both ϵ-sgreedy and δ-redundant, we are done.

Next, we claim that the process concludes and analyze its time complexity. Taking $w = 0$ will set each $A(x,t)$ to $\widetilde{A}(x,t-1)$, and taking w to be larger than $f(1)$ will set all of them to be 0. Therefore $\sum_x A_w(x,t)$ can range between 0 and at least $\sum_x \widetilde{A}(x,t-1) \geq 1$ (as $t \leq M$), and by continuity, there is some $w^\star$ such that this sum is exactly $1 - \delta/2$.

Always, $|\sum_x A(x,t) - \sum_x A_w(x,t)| < \delta/4$. This in particular means that for any w such that $\sum_x A_w(x,t) \in [1 - 3\delta/4, 1 - \delta/4]$ we are guaranteed to stop. Also, it means that if $w > w^\star$ then it will never be considered small, and if $w < w^\star$ it will never be considered large. Thus our binary search is valid and is guaranteed to stop.

The time it will take to stop is the time until the interval between the w's it considers guarantees that the difference between $\sum_x A_w(x,t)$'s is smaller than $\delta/4$. Say the current interval is $[w_1, w_2]$. In the terminology of Lemma 34, $p = A_{w_2}(x,t)$, $p+\epsilon = A_{w_1}(x,t)$, $w_1 = f(x)\phi(p+\epsilon)$, and $w_2 = f(x)\phi(p)$. Therefore, according the the l.h.s. of the lemma,

$$A_{w_1}(x,t) - A_{w_2}(x,t) = \epsilon \le \left(\frac{(k-1)(w_2-w_1)}{f(x)(1-C(k))}\right)^{1/k-1},$$

which is at most $\delta/4M$ if

$$w_2 - w_1 \le \frac{f(x)(1-C(k))}{k-1}\left(\frac{\delta}{4M}\right)^{k-1}.$$

This can be guaranteed with $O(\log(1/f(M)) + \log(1/(1-C(k))) + k\log(1/\delta) + k\log(M))$ binary search steps, as required. □

E.2.1 Upper and Lower bounds on the Value

This in fact generalizes the monotonicity lemma.

Lemma 34. *Let*

$$\phi(p) = \sum_{i=0}^{k-1} C(i+1)\binom{k-1}{i} p^i (q-p)^{k-1-i},$$

where $q \in (0,1]$ and $p \in [0,q]$. For $\epsilon > 0$, where $p+\epsilon \le q$,

$$\frac{1-C(k)}{k-1}\epsilon^{k-1} \le \phi(p) - \phi(p+\epsilon) \le 4^{k-1}\epsilon.$$

Proof. Let us first prove the upper bound. Dropping the $C(i+1)$'s as they are at most 1, and considering the sum as a sum of 2^{k-1} terms:

$$\begin{aligned}\phi(p) - \phi(p+\epsilon) &\le 2^{k-1}\max_{i=0}^{k-1}\left|(p+\epsilon)^i(q-p-\epsilon)^{k-1-i} - p^i(q-p)^{k-1-i}\right| \\ &\le 2^{k-1}\max_{i=0}^{k-1}\left((p+\epsilon)^i(q-p)^{k-1-i} - p^i(q-p-\epsilon)^{k-1-i}\right).\end{aligned}$$

Denoting $a = p$ and $b = q-p-\epsilon$, the term we are maximizing is:

$$(a+\epsilon)^i(b+\epsilon)^{k-1-i} - a^i b^{k-1-i} \le 2^{k-1}\epsilon,$$

where the last inequality is because opening the left term to its 2^{k-1} terms, each one is a multiplication of some powers of a, b and ϵ, all of them are at most 1. The only one that does not contain ϵ is canceled out by the right term. This establishes the upper bound part.

Next, we prove the lower bound. Denote

$$B_i(p) := \sum_{j=0}^{i}\binom{k-1}{j}p^j(q-p)^{k-1-j} = q^{k-1}\sum_{j=0}^{i}\binom{k-1}{j}\left(\frac{p}{q}\right)^j\left(1-\frac{p}{q}\right)^{k-1-j}.$$

Then,

$$\phi(p) = C(k)B_{k-1}(p) + (C(k-1)-C(k))B_{k-2}(p) + \cdots + (C(1)-C(2))B_0(p).$$

All of the B_i are non-increasing in p, as it is q^{k-1} times the probability that at most i of $k-1$ Bernoulli random variables each of probability p/q are 1. Also, all of the $C(i-1)-C(i)$ are non-negative, and at least one of them is strictly positive, and is at least $(C(1)-C(k))/(k-1)$. Therefore,

$$\phi(p) - \phi(p+\epsilon) \ge \frac{1-C(k)}{k-1}(B_j(p) - B_j(p+\epsilon)), \tag{6}$$

where $j \le k-2$. For any integers $i \le n$, denote by $X_i^n(p)$ the probability that at least i of n i.i.d. Bernoulli random variables of probability p are 1. Since $B_j(p)$ can be seen as $q^{k-1}(1 - X_{j+1}^{k-1}(p/q))$, Claim 35 (see below) implies that the r.h.s. of Eq (6) is at least:

$$\frac{1-C(k)}{k-1}q^{k-1}\left(\frac{p+\epsilon}{q} - \frac{p}{q}\right)^{k-1} = \frac{1-C(k)}{k-1}\epsilon^{k-1},$$

as required. □

Claim 35. *If $i \geq 1$, and p and $p+\epsilon$ are in $[0,1]$, then*

$$X_i^n(p+\epsilon) - X_i^n(p) \geq \epsilon^n.$$

Proof. We prove it by a double induction on $i+n$. One base case if when $i=1$ and the second is when $i=n$. First, regarding the former base case:

$$\begin{aligned}X_1^n(p+\epsilon) - X_1^n(p) &= (1-(1-p-\epsilon)^n) - (1-(1-p)^n) \\ &= (1-p)^n - (1-p-\epsilon)^n = (a+\epsilon)^n - a^n,\end{aligned}$$

where $a = 1-p-\epsilon \geq 0$. This last expression is easily seen to be at least ϵ^n. Now, for the latter base case:

$$X_n^n(p+\epsilon) - X_n^n(p) = (p+\epsilon)^n - p^n \geq \epsilon^n.$$

For the induction step, first note that if $n > i$,

$$X_i^n(p) = pX_{i-1}^{n-1}(p) + (1-p)X_i^{n-1}(p).$$

Therefore, in this case,

$$\begin{aligned}X_i^n(p+\epsilon) - X_i^n(p) &= p\left(X_{i-1}^{n-1}(p+\epsilon) - X_{i-1}^{n-1}(p)\right) \\ &\quad + (1-p)\left(X_i^{n-1}(p+\epsilon) - X_i^{n-1}(p)\right) \\ &\quad + \epsilon\left(X_{i-1}^{n-1}(p+\epsilon) - X_i^{n-1}(p+\epsilon)\right) \\ &\geq p\epsilon^{n-1} + (1-p)\epsilon^{n-1} + \epsilon\cdot 0 = \epsilon^{n-1} \geq \epsilon^n,\end{aligned}$$

where for the first two terms we used the induction hypothesis, and the last term is non-negative by the definition of X. $\square$

F Missing Proofs for Section 4 - Exclusive Policy

F.1 $A^\star$ is Sgreedy and non-Redundant

Lemma 36. *Under the exclusive policy, $A^\star$ restricted to $T \leq M$ rounds is sgreedy and non-redundant.*

Proof. See Section 1.1 for a description of $A^\star$. To see it is non-redundant:

$$\sum_x A^\star(x,t) = \sum_x \widetilde{A^\star}(x,t-1) - \sum_x \widetilde{A^\star}(x,t) = (M-t-1) - (M-t) = 1.$$

To see that it is sgreedy, let us fix t. The value of a box x is

$$v_{A^\star}(x,t) = f(x)\widetilde{A^\star}(x,t)^{k-1} = f(x)\min\left(1, \alpha(t)q(x)\right)^{k-1} = \min(f(x), \alpha(t)^{k-1}).$$

If $A^\star(x,t) > 0$ then $\widetilde{A^\star}(x,t) < 1$, which means that $v_{A^\star}(x,t) = \alpha(t)^{k-1}$, and so is constant for all such x. If $A^\star(x,t) = 0$ then $\widetilde{A^\star}(x,t) = 1$, and so $\alpha(t) \geq 1/q(x) = f(x)^{1/(k-1)}$. Then $v_{A^\star}(x,t) = f(x) \leq \alpha(t)^{k-1}$, as required from a sgreedy strategy. $\square$

F.2 The Price of Anarchy

F.2.1 Proof of Lemma 15

Lemma 15. *If a profile $\mathbb{P}$ is non-redundant and contains no high boxes, then all boxes are saturated.*

Proof. If x is not high, then

$$\widetilde{A}(x,T)^k \leq \prod_{B\in\mathbb{P}} \widetilde{B}(x,T) \leq \left(\frac{1}{k}\sum_{B\in\mathbb{P}} \widetilde{B}(x,T)\right)^k, \tag{7}$$

which means that

$$\sum_{B\in\mathbb{P}} \widetilde{B}(x,T) \geq k\widetilde{A}(x,T).$$

This is the same as

$$\sum_{B,t} B(x,t) \leq k \sum_{t} A(x,t).$$

As there are no high boxes, summing over all x's:

$$\sum_{B,x,t} B(x,t) \leq k \sum_{x,t} A(x,t). \tag{8}$$

As all players in $\mathbb{P}$ are non-redundant, $\sum_x B(x,t) = 1 \geq \sum_x A(x,t)$ for every t and every player B. Hence, Eq. (8) is actually an equality. On the other hand, Eq. (7) is a strict inequality if box x is low and not saturated. Therefore, if even one box is low, we get that Eq. (8) is strict as well, in contradiction. □

Lemma 18. *An alteration can only decrease the probability of success.*

Proof. We first make the following sequence of claims:

1. Consistency: If box x is low (high) w.r.t. to some intermediate profile then it was low (high) in all profiles preceding it.
2. Monotonicity: Alterations can only increase the value of high boxes, and can only decrease the value of low boxes. This is true in the eyes of all players.
3. If an alteration is made to an intermediate profile $\mathbb{P}$ with x, y, B, t as above (that is, shifting some probability mass from $B(x,t)$ to $B(y,t)$), then w.r.t. to $\mathbb{P}$, $v_B(x,T) = v_B(x,t)$.
4. If an alteration is made to an intermediate profile $\mathbb{P}$ with x, y, B, t as above, then w.r.t. to $\mathbb{P}$, $v_B(x,T) \geq v_B(y,T)$

Here are the proofs:

1. This one is clear from the way alterations are defined.
2. Recall $v_B(x,t) = f(x) \prod_{B' \neq B} \widetilde{B'}(x,t)$. If x is high, then alterations only decrease its $B'(x,t)$ and so increase $\widetilde{B'}(x,t)$, thus increasing its value in the eyes of the different players. This works in the same way, the other way around, for low boxes.
3. By the way t is chosen in an alteration, $B'(x,s) = 0$ for every player B' and every $s > t$. Therefore, for all B', $\widetilde{B'}(x,t) = \widetilde{B'}(x,T)$, and so $v_B(x,T) = v_B(x,t)$.
4. We know that $B(x,t) > 0$ in $\mathbb{P}$. Since x is high for the current profile, then by Item 1 it was also high in all preceding profiles. In particular, $B(x,t)$ was never increased, and so $B(x,t) > 0$ in $\mathbb{P}_{\texttt{nash}}$ as well. Since y is low, then by Item 1, it was also low in $\mathbb{P}_{\texttt{nash}}$, and so not full. Therefore, since $\mathbb{P}_{\texttt{nash}}$ is an equilibrium, we can apply Lemma 16, and get $v_B(x,t) \geq v_B(y,t)$ w.r.t. $\mathbb{P}_{\texttt{nash}}$. By Item 2, this is also true in $\mathbb{P}$.

   By Item 3, $v_B(x,t) = v_B(x,T)$, and as always $v_B(y,T) \leq v_B(y,t)$, we get the result.

Now, considering an alteration, let us examine the success probability of a profile $\mathbb{P}$, and express it as a function of the matrix and values of the player B involved in the alteration:

$$\begin{aligned}\texttt{success}(\mathbb{P}) &= \sum_x f(x) \left(1 - \prod_{B' \in \mathbb{P}} \widetilde{B'}(x,T)\right) = \sum_x f(x) - \sum_x f(x) \prod_{B' \in \mathbb{P}} \widetilde{B'}(x,T) \\ &= \sum_x f(x) - \sum_x \widetilde{B}(x,T) v_B(x,T).\end{aligned}$$

The alteration will increase $B(y,t)$ by ϵ and decrease $B(x,t)$ by ϵ. As changes are made only to player B, all the v_B's are not affected. Thus, the change in the success probability as a result of such an alteration is:

$$\epsilon(v_B(y,T) - v_B(x,T)).$$

By Item 4 this value is non-positive, and so the alteration can only decrease the success probability. □

F.2.2 Calculating the PoA

Theorem 8. *For every T,* $\sup_f \texttt{PoA}(C_{\texttt{ex}}, f, T) = (1 - (1 - 1/k)^k)^{-1}$.

Proof. Denote $m = \min\{kT, M\}$. Let $A_{\texttt{unif}}$ be the very simple strategy that chooses at each step uniformly among all the boxes in $\{1, \ldots, m\}$ that it did not choose yet. The probability that a player checked a specific box x in $\{1, \ldots, m\}$ is T/m. Therefore,

$$\texttt{success}(A_{\texttt{unif}}) = \sum_{x=1}^{m} f(x) \left(1 - \left(1 - \frac{T}{m}\right)^k\right) \geq \sum_{x=1}^{m} f(x) \left(1 - \left(1 - \frac{1}{k}\right)^k\right).$$

Now, by the optimality of $A^\star$, $\texttt{success}(A^\star) \geq \texttt{success}(A_{\texttt{unif}})$, and so:

$$\texttt{PoA}(C_{\texttt{ex}}, f, T) = \frac{\sum_{x=1}^{m} f(x)}{\texttt{success}(A^\star)} \leq \left(1 - \left(1 - \frac{1}{k}\right)^k\right)^{-1},$$

where the first equality is by the fact that the PoA of the exclusive policy is attained by $A^\star$.

Next, consider the situation where f is the uniform distribution on $\{1, \ldots, M\}$, where $M = kT$. In this specific case, $A_{\texttt{unif}}$ is trivially sgreedy, as the value of all boxes at each point in time is the same. Therefore, as it is also non-redundant, by Theorem 2, it is a symmetric equilibrium and in particular, an equilibrium. Hence,

$$\texttt{PoA}(C_{\texttt{ex}}, f, T) \geq \frac{\sum_{x=1}^{M} f(x)}{\texttt{success}(A_{\texttt{unif}})} = \left(1 - \left(1 - \frac{T}{M}\right)^k\right)^{-1} = \left(1 - \left(1 - \frac{1}{k}\right)^k\right)^{-1}.$$

□

F.3 Example of Non-Robustness for the Exclusive Policy

Lemma 37. *For every $\epsilon > 0$ there is a configuration where $A_k^\star$ is not an $(1+\epsilon)$-equilibrium when played by $k+1$ players.*

Proof. Set $T = 1$, and let all boxes except the first have the same f. Denote by $p(x)$ the probability that $A_k^\star$ plays x. Clearly, p is the same for all boxes except the first. Also, by how $A^\star$ is defined $W = M$, and so $p(x) \geq 0$ for all boxes. Denote by $v(1)$ the value boxes get when $A_k^\star$ is played by k players, as we know this is equal for all boxes. Therefore, when played by $k+1$ players, for all x,

$$v(x, 1) = f(x)(1 - p(x))^k = v(1)(1 - p(x)).$$

Denote by B the algorithm that plays box 2 with probability 1. Using the equation above we get:

$$\begin{aligned} \frac{U_{A^\star}(B)}{U_{A^\star}(A^\star)} &= \frac{1 - p(2)}{p(1)(1 - p(1)) + (1 - p(1))(1 - p(2))} = \frac{1 - p(2)}{1 - p(1)} \cdot \frac{1}{1 + p(1) - p(2)} \\ &\geq \frac{1 - p(2)}{1 - p(1)} \cdot \frac{1}{2} = \frac{1}{2} \cdot \frac{\alpha(1)q(2)}{\alpha(1)q(1)} = \frac{1}{2}\left(\frac{f(1)}{f(2)}\right)^{\frac{1}{k-1}}. \end{aligned}$$

Therefore, if for example $f(1) = 1/2$, we can set $f(x) = f(2)$ to be as small as we want (by increasing the number of boxes M), and this ratio to be as large as we wish. □

F.4 Robustness of $A^\star$

Theorem 9. *If $\frac{f(1)}{f(M)} \leq (1+\epsilon)^{\frac{k-1}{k'}}$, then $A_k^\star$ is a $(1+\epsilon)$-equilibrium when played by $k + k'$ players.*

We first prove a simple lemma:

Lemma 38. *For non-negative $a_1, \ldots, a_n$ and strictly positive $b_1, \ldots, b_n$,* $\sum_i a_i / \sum_i b_i \leq \max_i a_i/b_i$.

Proof. First the case $n = 2$. Assume by contradiction that $(a_1 + a_1)/(b_1 + b_2)$ is greater than both a_1/b_1 and a_2/b_2.

$$\frac{a_1 + a_2}{b_1 + b_2} > \frac{a_1}{b_1} \quad\Longrightarrow\quad a_1b_1 + a_2b_1 > a_1b_1 + a_1b_2 \quad\Longrightarrow\quad a_2b_1 > a_1b_2.$$

Also:

$$\frac{a_1 + a_2}{b_1 + b_2} > \frac{a_2}{b_2} \quad\Longrightarrow\quad a_1b_2 + a_2b_2 > a_2b_1 + a_2b_2 \quad\Longrightarrow\quad a_1b_2 > a_2b_1,$$

in contradiction. By induction:

$$\frac{\sum_i a_i}{\sum_i b_i} \le \max\left\{\frac{a_1}{b_1}, \frac{\sum_{i\ge 2} a_i}{\sum_{i\ge 2} b_i}\right\} \le \max_i \frac{a_i}{b_i}. \qquad \square$$

We can now proceed to prove Theorem 9:

Proof. Recall the definition of $A^\star_k$'s matrix (we will henceforth drop the subscript k) given in Section 1.1. Let $q(x) = f(x)^{-1/(k-1)}$. For each t, $\widetilde{A^\star}(x,t) = \min(1, \alpha(t)q(x))$, where $\alpha(t) \ge 0$ is such that $\sum_x 1 - \widetilde{A^\star}(x,t) = t$.

As the $q(x)$ are non-decreasing, then for every t there is some $W_t \le M$, such that $\widetilde{A^\star}(x,t) < 1$ for every $x \le W_t$, and $\widetilde{A^\star}(x,t) = 1$ for larger x. If $\alpha(t) \in [1/q(W_t+1), 1/q(W_t))$, then

$$t = \sum_x 1 - \widetilde{A^\star}(x,t) = \sum_{x \le W_t} 1 - q(x)\alpha(t),$$

and so W_t is the largest such that

$$\sum_{x \le W_t} 1 - q(x)/q(W_t) < t.$$

Note this characterization works also for the case $W_t = M$. In particular, if $W_t < M$ then

$$\alpha(t) \ge 1/q(W_t+1),$$

which we will need later on. When played with $k+k'$ players, the value w.r.t. $A^\star$ is

$$v_{A^\star}(x,t) = f(x)\widetilde{A^\star}(x,t)^{k+k'-1}.$$

Running with k players, for $x \le W_t$ this is equal to $v(t)\widetilde{A^\star}(x,t)^{k'}$, where $v(t)$ is the value of all boxes in $\{1, \dots, W_t\}$ when running $A^\star$ with k players, in which case we know they all have the same value. That is because $A^\star$ is sgreedy, and all of these boxes have $A^\star(x,t) > 0$ (because if $\widetilde{A^\star}(x,t) < 1$, since $\alpha(t)$ is strictly decreasing, we get $A^\star(x,t) > 0$). We also know that for all other boxes, which have $A^\star(x,t) = 0$, by the fact that $A^\star$ is sgreedy when running with k players, $v_{A^\star}(x,t) = f(x) \le v(t)$.

Consider now an alternative strategy B played by one of the $k+k'$ players. The relative utility this player gains is:

$$\begin{aligned}\frac{U_{A^\star}(B)}{U_{A^\star}(A^\star)} &\le \frac{\sum_t \max_x v_{A^\star}(x,t)}{\sum_t \sum_x A^\star(x,t) v_{A^\star}(x,t)} \\ &\le \frac{\sum_t \max_x v_{A^\star}(x,t)}{\sum_t \min_{x \le W_t} v_{A^\star}(x,t)} = \max_t \left\{\frac{\max_x v_{A^\star}(x,t)}{\min_{x\le W_t} v_{A^\star}(x,t)}\right\},\end{aligned}$$

where the last inequality is by Lemma 38.

Fix t. If $W_t = M$, then as mentioned, for all x, $v^\star_A(x,t) = v(t)\widetilde{A^\star}(x,t)^{k'} = v(t)(\alpha(t)q(x))^{k'}$. As the $q(x)$ are non-decreasing in x,

$$\frac{\max_x v_{A^\star}(x,t)}{\min_x v_{A^\star}(x,t)} = \left(\frac{\alpha(t)q(M)}{\alpha(t)q(1)}\right)^{k'} = \left(\frac{f(1)}{f(M)}\right)^{\frac{k'}{k-1}} \le 1+\epsilon,$$

as desired. If $W_t < M$, then as mentioned before, $\alpha(t) \ge 1/q(W_t+1)$. Also, as said, $v_{A^\star}(x,t) = v(t)\widetilde{A^\star}(x,t)^{k'}$ for $x \le W_t$, and is at most $v(t)$ for $x > W_t$. Therefore:

$$\begin{aligned}\frac{\max_x v_{A^\star}(x,t)}{\min_{x\le W_t} v_{A^\star}(x,t)} &\le \frac{v(t)}{v(t)\widetilde{A^\star}(1,t)^{k'}} = \frac{1}{(\alpha(t)q(1))^{k'}} \\ &\le \left(\frac{q(W_t+1)}{q(1)}\right)^{k'} = \left(\frac{f(1)}{f(W_t+1)}\right)^{\frac{k'}{k-1}} \le 1+\epsilon,\end{aligned}$$

as desired. $\square$

G The PoA of the Sharing Policy

The following is practically the same as the proof of [12], where it is shown that the price of anarchy for valid increasing utility systems is at most 2. We include it here for completeness, and extend it slightly to hold for approximate equilibria.

Lemma 39. *Consider any policy C such that for all ℓ, $C(\ell) \leq 1/\ell$. Let $\mathbb{P} = \{A_1, \ldots, A_k\}$ be some $(1+\epsilon)$-equilibrium. Then, for every f and T,*

$$\texttt{success}(\mathbb{P}) \geq \frac{1}{2+\epsilon} \cdot \max_{\mathbb{P}' \in \mathcal{P}(T)} \texttt{success}(\mathbb{P}').$$

In particular, $\texttt{PoA}(C, f, T) \leq 2$.

Proof. Let $\mathbb{P} = \{A_1, \ldots, A_k\}$ be some $(1+\epsilon)$-equilibrium, and let $\mathbb{O} = \{O_1, \ldots, O_k\}$ be the optimal profile in terms of success probability. Our goal is to show $\texttt{success}(\mathbb{P}) \geq \frac{1}{2+\epsilon}\texttt{success}(\mathbb{O})$. The trick is to play all of them together.

$$\begin{aligned}\texttt{success}(\mathbb{P}, \mathbb{O}) = \texttt{success}(\mathbb{P}) &+ (\texttt{success}(\mathbb{P}, O_1) - \texttt{success}(\mathbb{P})) \\ &+ \ldots + (\texttt{success}(\mathbb{P}, O_1, \ldots, O_k) - \texttt{success}(\mathbb{P}, O_1, \ldots, O_{k-1})).\end{aligned}$$

We wish to bound the difference $(\texttt{success}(\mathbb{P}, O_1, \ldots, O_i) - \texttt{success}(\mathbb{P}, O_1, \ldots, O_{i-1}))$. For that, note that when adding a player (O_i in this case) to an existing profile, its utility is always at least what it contributes to the total success probability of the existing profile. This is because for whatever it contributes it gets a utility of 1, and if it steps on a box that another player already checks, it adds nothing to the success probability, yet may gain some utility (depending on C and the time it checks the box relative to the other players). Therefore,

$$\begin{aligned}&\texttt{success}(\mathbb{P}, O_1, \ldots, O_i) - \texttt{success}(\mathbb{P}, O_1, \ldots, O_{i-1}) \\ &\qquad \leq U_{\mathbb{P}, O_1, \ldots, O_{i-1}}(O_i) \leq U_{\mathbb{P}^{-A_i}}(O_i) \leq (1+\epsilon) U_{\mathbb{P}^{-A_i}}(A_i),\end{aligned}$$

where the last step is because $\mathbb{P}$ is a $(1+\epsilon)$-equilibrium. As $C(\ell) \leq 1/\ell$ for every ℓ, $\sum_i U_{\mathbb{P}^{-A_i}}(A_i) \leq \texttt{success}(\mathbb{P})$, and so:

$$(2+\epsilon)\,\texttt{success}(\mathbb{P}) \geq \texttt{success}(\mathbb{P}, \mathbb{O}) \geq \texttt{success}(\mathbb{O}),$$

as required. □