Bounds for the Communication Complexity of Approximate Correlated Equilibria

Young Kun Ko * Ariel Schvartzman †

May 8, 2017

Abstract

In the recent paper of [BR16], the authors show that, for any constant $10^{-15} > \varepsilon > 0$ the communication complexity of $\varepsilon$-approximate Nash equilibria in 2-player $n \times n$ games is $n^{\Omega(\varepsilon)}$, resolving the long open problem of whether or not there exists a polylogarithmic communication protocol. In this paper we address an open question they pose regarding the communication complexity of 2-player $\varepsilon$-approximate correlated equilibria.

For our upper bounds, we provide a communication protocol that outputs a $\varepsilon$-approximate correlated equilibrium for multiplayer multi-action games after exchanging $\tilde{O}(m n^4 \varepsilon^{-4})$ bits, saving over the naive $O(mn^4)$-bits protocol when the number of players is large.

For our lower bounds, we exhibit a simple two player game that has a logarithmic information lower bound: for any $\Omega(n^{-1}) < \varepsilon < \frac{1}{10}$, the two players need to communicate $\Omega(\varepsilon^{-1/2} \log n)$-bits of information to compute any $\varepsilon$-correlated equilibrium in the game. For the $m$-players, 2-action setting we show a lower bound of $\Omega(m)$ bits, which matches the upper bound up to polylogarithmic terms and shows that the linear dependence on the number of players is unavoidable.

*Department of Computer Science, Princeton University, yko@cs.princeton.edu.
†Department of Computer Science, Princeton University, acohenca@cs.princeton.edu.
1 Introduction

While Nash equilibria are arguably the most studied notion of equilibrium in strategic games, recent results regarding their communication and computational complexity have undermined their prevalence as a predictable solution concept when agents are computationally bounded. In particular, these results show that two players cannot converge to any approximate Nash equilibria in the limited communication setting where each player only knows its utility function. While there have been multiple attempts to produce procedures that converge to Nash equilibria of general games [HM06, GL07, FY03], it has been shown that at least \( \exp(m) \) bits of communication are required to compute Nash equilibria of \( m \)-player, constant action games [HM07]. For the case of 2-player games, it has been recently shown that even computing approximate Nash equilibria requires \( \text{poly}(n) \) bits of communication [BR16].

In addition, even in the setting where all the payoffs matrices are known, Nash equilibria seem to be unnatural due to their computational hardness. Computing any exact Nash equilibrium is known to be PPAD-complete, making it unlikely to have any polynomial time algorithm [CDT09, DGP09]. Furthermore, it has been shown that under the Exponential Time Hypothesis (ETH) for the class PPAD, \( \varepsilon \)-approximate Nash equilibria cannot be computed in time faster than quasi-polynomial in the number of strategies per player [Rub16]. This almost exactly matches the algorithm of [LMM03]. The picture becomes more bleak when we consider \( m \)-player games. In this case, the problem of even approximating Nash equilibria becomes PPAD-complete [Rub14]. These results suggest that approximate Nash equilibria may not be efficiently computable.

Correlated equilibria arise as an alternative equilibrium concept. This notion, introduced in the seminal work of [Aum74], allows agents to cooperate in order to reach stability. Informally, a strategy profile is a correlated equilibrium when a referee or trusted party can draw strategy samples according to it and recommend them to the players in such a way that they have no incentive to consistently deviate, assuming everyone else plays according to their recommendation. Computationally, correlated equilibria are in sharp contrast to Nash equilibria: there exists an ellipsoid-based algorithm to compute exact correlated equilibrium in polynomial time even for multi-player games [PR08], for a large (but not universal) class of games including graphical games, anonymous games, congestion games and scheduling games. Unfortunately this result is still unsettling: one can imagine many settings where a referee may not have access to all utility functions or where players may not want to share such information with a referee.

This is indeed comparable to many interesting communication or distributed computation problems where if one party knows all parts of the input, it is easy to compute the output (e.g. disjointness [BGPW13], equality, gap Hamming distance [CR11]). In particular, the hardness comes only from the distributional nature of the input not the computational aspect, unlike Nash equilibria.

With this in consideration it becomes more natural to ask whether there is a communication protocol for computing correlated equilibria better than the naive one where each player sends their payoff matrix to a referee who computes the answer. In the case of exact correlated equilibria for 2-player games, a simple reduction from the distributed version of linear programming shows that sending the full payoff matrix is indeed optimal [CSS9]. For the exact or approximate \( m \)-player, constant action case, there are simple procedures that converge quickly and use at most polynomial communication in the natural parameters of the input [HMC00, CL03, CBL06].

In this paper we address the question posed by [BR16] of bounding the communication complexity of approximate correlated equilibrium. We make progress in both providing non-trivial protocols (for the \( n \geq 4 \) case) and deducing non-trivial lower bounds. The arguments used on the lower bound proofs rely on tools from information complexity, which lower bounds communication complexity.
1.1 Our results

Let $G^{m}_{n}$ denote the set of all $m$-players, $n$-action game, described by the payoff tensor of size $n^m$ for each player. We consider three regimes in particular: $G^{m}_{n}$, $G^{m}_{2}$ and $G^{2}_{n}$.

Upper Bounds. Our upper bounds are similar in spirit to those of [HM10, GR16]. The protocol we provide is based on a non-adaptive no-regret learning algorithm by [HMC00]. Unfortunately, this protocol converges to a different notion of approximate correlated equilibrium and assumes that the number of actions per player is constant. We overcome both of these barriers by showing that running their protocol for a longer number of rounds converges to the standard notion of approximate correlated equilibrium. Our result works for general games and has strong implications for the case of $m$-player $O(1)$-action games. Unfortunately, for the 2-player $n$-action case, our protocol has a communication cost higher than the naive protocol where one player just shares their payoff, exchanging $O(n^2)$ bits. But its dependance on the number of players is significantly better than the naive protocol as the number of players increases.

**Theorem 1.** There exists a communication protocol $\Pi$ such that for any $m$-player $n$-action game $G \in G^{m}_{n}$, the players compute an $\varepsilon$-CE after exchanging at most $\tilde{O}(n^{4m\varepsilon^{-4}})$ bits.

Lower Bounds. Our lower bound is similar at a first glance to that of [BR16], but our techniques differ significantly due to the nature of the solution concepts studied. As it is pointed out in [BR16] the hardness of proving lower bounds for equilibria lies in being able to hide the solutions (which, by [Nas51, Aum74], must exist). But unlike computing Nash equilibria, where the strategies are independent, for correlated equilibrium the task of hiding solutions is much harder, in part due to their more general nature as a solution concept (i.e. any lower bounds for $\varepsilon$-correlated equilibria extend to $\varepsilon$-Nash equilibria). In particular, we need to dissuade from arbitrary correlated distributions. This obstacle becomes clearer when we consider the communication complexity of computing correlated equilibrium.

Even in this setting we exhibit a hard game in which $\Omega(\varepsilon^{-1/2} \log n)$ bits of communication must be exchanged for two players to agree on an approximate equilibrium. There is an easy, trivial lower bound of $\Omega(\log n)$ from games where both players have a dominant strategy. Our result provides an explicit dependance on the approximation factor. In the $m$-player 2-action setting, we prove a linear lower bound in the number of players. Note that this proves near-optimality of Theorem 1 and shows that the linear dependence on the number of players is unavoidable.

For the two-player case, the proof consists of two steps. First we construct a game where there is a unique Nash Equilibrium and any $\varepsilon$-CE must have support size $\Omega(\varepsilon^{-1})$. Unfortunately, the Nash Equilibrium is ‘trivial’ in a sense that it requires no communication between the players. We circumvent this by adding a small side game that kills the trivial Nash Equilibrium but makes any $\varepsilon$-correlated equilibria retain $\Omega(\sqrt{\varepsilon})$ weight on the original game.

**Theorem 2.** There exists a 2-player $n$-action game $G \in G^{2}_{n}$ such that $\Omega(\varepsilon^{-1/2} \log n)$ bits of communication are required for the players to agree on a $\varepsilon$-CE for some small $\varepsilon = \Omega(1)$.

Moreover the game has the following nice property: for any $\varepsilon$-correlated equilibria, the players must know the location of entries whose payoff is non-zero for at least $\Omega(1/\sqrt{\varepsilon})$ number of rows. Since the search space for each entry is exactly $n$, this implies the following corollary:

---

1 Consider the following game. There is a row of 1’s for a row player, and a column of 1’s for column player. Payoff is 0 for any other entries.
**Corollary 1.** There exists a 2-player $n$-action game $G \in G_n^2$ such that the query complexity of computing $\varepsilon$-approximate correlated equilibria is $\Omega(\varepsilon^{-1/2})$ for some small $\varepsilon = \Omega(1)$.

In particular we will provide a game whose payoff matrices are “dominated” by two independent permutation matrices for each player. On such games we show that at least one of the players must learn $\Omega(\log n/\sqrt{\varepsilon})$-bits of information about other player’s permutation.

For the $m$-player, 2-action game we provide a simple matching game. The players are randomly split into two groups and can play one of two actions. The players on the first group only get a payoff if they act according to some random signal provided to them, and the second players must imitate their behavior. In order to achieve an approximate equilibrium, most players on the first group need to share the strategy they are playing in order to be matched by their counterparts.

**Theorem 3.** There exists a $m$-player 2-action game $G \in G_m^2$ such that $\Omega(m)$ bits of communication are required for the players to agree on a $\varepsilon$-CE for $\varepsilon < 1/3$.

### 1.2 Related Work

The communication complexity of predictable solution concepts has gained a lot of attention and by now most problems pertaining exact and approximate Nash equilibria are well understood. It is known that the communication complexity of computing pure Nash equilibria in 2-player $n$-action games is $\text{poly}(n)$ [CS04]. For $m$-player binary action games the complexity escalates to $\exp(m)$, even if we relax the solution concepts from exact pure or mixed Nash equilibria [HM10]. These results were extended to the case of approximate Nash equilibria. In particular, [BR16] showed that the randomized communication complexity of computing $\varepsilon$-Nash equilibria in 2-player $n$-action games and $m$-player binary action games is $\Omega(n^\varepsilon)$ and $2^{\Omega(m^\varepsilon)}$ for some constant $\varepsilon > 0$.

Some results are known for the communication complexity of computing correlated equilibria for the family of $m$-player binary action games with bounded, integer payoffs. There is a protocol for the family that computes correlated equilibria after exchanging polynomially many bits in terms of $n$ and the magnitude of the payoffs [HM10]. The former is based on the polynomial time algorithm for computing correlated equilibria for a large class of games by [PR08] and the later is based on a no-regret learning algorithm by [CBL06]. It is worth noting that in the same paper they exhibit a family of multiplayer games that do not need to communicate at all to find exact correlated equilibria.

**Query Complexity.** Another lens through which to consider the cost of computing equilibria is that of query complexity. In this model, a single agent has black box access to the payoff function and can query it on either pure strategies or mixed strategies. A long line of work [FGGS13, HN13, Bab14, Rub16] has recently established that the query complexity of computing approximate 2-player $n$-action Nash equilibria and approximate $m$-player 2-action Nash equilibria is $\text{poly}(n)$ and $\exp(m)$, respectively, even for randomized algorithms. For approximate $m$-player correlated equilibria, there is an exponential gap between the best randomized algorithms and the deterministic lower bounds [HN13, Bab14].

In the case of correlated equilibria (and coarse correlated equilibria), [GR16] show that for $m$-player binary action games and for any $\varepsilon < 1/2$ the query complexity is $\Theta(\log m)$. They provide an algorithm based on multiplicative weights that uses $\tilde{O}(nm\varepsilon^{-2})$ queries to compute $\varepsilon$-coarse correlated equilibria in $m$-player $n$-action games.

Recently and independently of this work [AG17] have shown a lower bound of $\Omega(n)$ for the randomized communication complexity of approximate correlated equilibria in the domain where $\varepsilon < \frac{1}{\text{poly}(n)}$. Their techniques and ideas are similar to ours, except their reduction is directly from
the disjointness problem whereas our analysis is based on information-theoretic arguments specific to the games we propose.

1.3 Future Directions

Though it closes the gap for $m$-player constant action correlated equilibria, our result leaves open exponential gap for the communication complexity of approximate 2-player $n$-action correlated equilibria, as well as for other small values of $m$. We conjecture that the right bounds are $O(n \text{ poly log } n)$. We share some future directions that might help in settling this question.

- Our argument for the lower bound for the 2-player $n$-action case relies on a claim about the support size of the game we construct. It is known that approximate correlated equilibria with small supports exist [BBP13] with size $O(\log^2 n)$. There is a small gap with the best lower bounds, $\Omega(\log n)$. If there are games for which the upper bounds were tight, our techniques could raise the communication lower by a $\log(n)$ factor.

- There exist algorithms to compute approximate correlated and coarse correlated equilibria [BBP13, HN13]. However they either rely on computing exact correlated equilibria, which is prohibitively expensive in the communication setting, or require polynomially many rounds, which already brings the communication cost above our conjectured answer. Progress in algorithms that are distributed in nature and exploit the structure of the solutions could improve on the cost of the protocol we propose.

- Not much is known about the query complexity of 2-player $n$-action approximate correlated equilibria. The folklore lower bound of $\Omega(n)$ from games with dominant strategies is significantly far from the trivial upper bound of $O(n^2)$. The result from [BR16] relies heavily on having a good understanding of the query complexity of exact, 2-player Nash equilibria and related Fixed Point Problems. Recent connections between lower bounds in query complexity and lower bounds in communication complexity [GPW15, Göö15] suggest that strong query complexity lower bounds could provide better communication complexity lower bounds.

2 Preliminaries

2.1 Game Theoretic definitions

We consider $m$-player $n$-action games where each player has a strategy set $A_i$ and a payoff function $u_i : A \rightarrow [0, 1]$, where $A = \prod_i A_i$. Let $A^{-i} = \prod_{j \neq i} A_j$. In 2-player games we will refer to the first player as Alice and the second player as Bob.

In this paper, we will be interested in studying approximate correlated equilibria (CE) and a different relaxation of exact correlated equilibria due to [HMC00], which we will refer to as approximate Hart-Mas-Colell Correlated Equilibria (HMCE).

A common interpretation of $\varepsilon$-CE is that a referee or trusted third party draws a strategy profile $a \in A$ according to the correlated distribution $x$ and recommends action $a_i$ to player $i$. A distribution $x$ is a $\varepsilon$-CE if any deviation from the recommended action does not yield a benefit greater than $\varepsilon$ for any player. A $\varepsilon$-HMCE only requires that no player benefits more than $\varepsilon$ by changing a single recommendation by any other action. We now formally define them in terms of regret, in accordance to [BBP13] (for more standard, equivalent definitions, see e.g. [HN13]).

\footnote{rephrase this}
**Definition 1.** Let \( R^i_f(a) = u_i(f(a_i), a_{-i}) - u_i(a) \) be the regret of player \( i \) for playing switching rule \( f \) at strategy profile \( a \). A distribution \( x \in \Delta(A) \) is an \( \varepsilon \)-correlated equilibrium if \( \mathbb{E}_{a \sim x}[R^i_f(a)] \leq \varepsilon \) for all players \( i \) and switching rules \( f : A_i \to A_i \).

**Definition 2.** A distribution \( x \in \Delta(A) \) is an \( \varepsilon \)-Hart-Mas-Colell correlated equilibrium if for every player \( i \), every recommendation \( a_i \in A_i \) and every action \( j \in A_i \), \( \sum_{a_{-i} \in A_{-i}} |u_i(j, a_{-i}) - u_i(a_i, a_{-i})| x(a_i, a_{-i}) \leq \varepsilon \).

The definitions are relaxations of the definition of exact (\( \varepsilon = 0 \)) correlated equilibria. However, as noted in [BBP13], approximate HMCE are uninteresting to study from a communication perspective. For any game there exists a 0-communication protocol that produces \( \frac{1}{k} \)-HMCE: independently of the payoff functions the players can agree on a set of \( k \) strategies in \( \Delta(A) \) and directly output a uniform distribution over them, where \( \frac{1}{k} \leq k \leq n \). It is not hard to see that this is indeed a \( \frac{1}{k} \)-HMCE. The advantage of working with this definition is that there exists a non-adaptive no-regret learning algorithms to compute such \( \varepsilon \)-equilibria for \( m \)-player games with a constant number of actions in a number of rounds polynomial in \( 1/\varepsilon \) [HMC00]. We adapt the algorithm into a communication protocol and revisit their analysis with the consideration that the number of actions per player is part of the input.

### 2.2 Communication Complexity definitions

In the classical communication problems there are \( m \) parties each of which are given inputs \( x_i \in \{0,1\}^n \) and who are interested in computing a joint function of their inputs, \( f(x) \), where \( x = (x_1, x_2, ..., x_m) \). The (randomized) communication complexity of a protocol \( \Pi \) for computing the function \( f(x) \) is the (expected) number of bits the two parties need to exchange to compute \( f(x) \) by following \( \Pi \) (with high probability). This quantity will be referred to as \( \text{CC}(\Pi, f, x) \). The communication complexity of protocol \( \Pi \) for computing \( f \) is the worst-case communication complexity for any pair of inputs, i.e. \( \text{CC}(\Pi, f) = \max_x \text{CC}(\Pi, f, x) \). The communication complexity of a function \( f \) is the minimum communication complexity over all protocols that compute \( f \), \( \text{CC}(f) = \min_{\Pi} \text{CC}(\Pi, f) \).

We will be interested in computing \( \varepsilon \)-CE, \( \varepsilon \)-HMCE of general games \( G = (A, u) \) belonging to the family of \( m \)-player \( n \)-action games \( \mathcal{G}^m_n \) with bounded payoff functions. We assume each player only has access to their payoff function \( u_i \). We consider protocols where for every round \( t \), each player broadcasts as many bits as it wants to the other players. We say that \( \Pi \) is a protocol for computing \( \varepsilon \)-CE of the game \( G \) if there exists a number of rounds \( T \) after which one of the players outputs a distributions \( x \in \Delta(A) \) that forms a \( \varepsilon \)-correlated equilibria with high probability. We let \( \text{CC}(\varepsilon \text{-CE}, \mathcal{G}^m_n) = \min_{\Pi} \text{CC}(\Pi, \varepsilon \text{-CE}, \mathcal{G}^m_n) = \min_{\Pi} \max_{G \in \mathcal{G}^m_n} \text{CC}(\Pi, \varepsilon \text{-CE}, G) \). We can analogously define the communication complexity of computing \( \varepsilon \)-HMCE.

Our lower bound proofs also use tools from information theory. We defer them to Appendix A as they are not fundamental to understanding the high level view of the high level view of the results.

### 3 Lower Bounds

#### 3.1 Warm up: A candidate game

In this subsection we present the first component of our lower bound for the two-player \( n \)-action case. The non-degenerate game we present, which we will refer to as the chasing game, has a simple structure in terms of equilibria. It has a unique exact Nash equilibrium which corresponds to the uniform strategy over all actions and multiple exact correlated equilibria that must be supported on
a large number of actions. Moreover, we can show that any \( \varepsilon \)-correlated Nash equilibrium requires strategies of support size at least \( \Omega(\varepsilon^{-1}) \). We defer any missing proofs to Appendix B.

**The Chasing Game** \( CG_n \). Take two permutations \( \sigma_A, \sigma_B \) from the set of \([n]\) elements such that \( \sigma_B \) is an \( n \)-cycle (\( \sigma_A \) is unconstrained). Then Alice gets payoff 1 whenever \((i, \sigma_A(i))\) is played and 0 otherwise. Bob gets payoff 1 whenever \((i, \sigma_A(j))\) is played, where \( j \) is such that \( \sigma_B(i) = j \), and 0 otherwise. By our choice of \( \sigma_B \), it is never the case that \( \sigma_B(i) = i \). Even though we use the permutations for the construction, we do not give the players access to them. Alice implicitly knows hers, but it provides her no significant information about Bob’s payoff matrix (since there are still \((n - 1)!\) of them). Bob doesn’t know either and learns nothing about Alice’s payoff from observing his own. An example for \( n = 4 \) is shown in Figure 1. The intuition is that Alice’s payoff is a random permutation matrix and for each of her actions \( \sigma_B \) points to Bob’s best response. Due to the cyclical nature of \( \sigma_B \), if we look at Alice’s best response to Bob’s action we will come across a different action for Alice, eventually spanning all \([n]\) actions.

**Claim 1.** The unique exact Nash equilibrium of \( CG_n \) is the uniform strategy. Any exact correlated equilibria must be supported on all the non-zero entries.

We now show something more subtle about the game: any \( \varepsilon \)-correlated equilibrium must be supported on at least \( \Omega(\varepsilon^{-1}) \) entries.

**Lemma 1.** Any \( \varepsilon \)-CE must have support \( \Omega(\varepsilon^{-1}) \).

**Proof.** Let \((i, j)\) be the entry with the largest total probability, \( p \). If \( p \leq 2\varepsilon \) then we are done, since this would require at least \( \varepsilon^{-1} \) strategies. If \( p > 2\varepsilon \) and \((i, j)\) is a \((0, 0)\)-entry, then the entry \((i, \sigma_A(i))\) must have total probability \( p' \geq \varepsilon \), since otherwise Alice would simply deviate to \( \sigma_B^{-1}(j) \) and gain more than \( \varepsilon \).

So there must be an entry \((i, \sigma_A(i))\) with probability \( p \geq \varepsilon \). Then the probability Bob gets recommended \( \sigma_A(i), p_B(i) \), must be non-zero. In order for this to be a correlated equilibrium, \((\sigma_B^{-1}(i), \sigma_A(i))\) must have total probability at least \( p - p_B(i)\varepsilon \), which is positive since \( p \geq \varepsilon \). This means that Alice must get recommended \( \sigma_B^{-1}(i) \) with non-zero probability \( p_A(\sigma_B^{-1}(i)) \). But then, in turn, for this to be a \( \varepsilon \)-correlated equilibrium we need total probability at least \( p - p_B(i)\varepsilon - p_A(\sigma_B^{-1}(i))\varepsilon \) on \((\sigma_B^{-1}(i), i)\). This chaining reasoning goes on until \( p - \varepsilon \sum_{i, i' \in \text{Supp}(p_B(i) + p_A(i'))} < \varepsilon \). But we know that the sum of the supported strategies is at most 2 so we get that \( p < 3\varepsilon \). Since \( p \) is the largest total mass, we must have at least \( \frac{1}{3\varepsilon} \) strategies in the support of our \( \varepsilon \)-correlated equilibrium. \( \blacksquare \)
3.2 Unrestricted lower bound for $G_n^2$

Even though we have a good understanding of the approximate equilibria of the chasing game, we still cannot show lower bounds for its communication complexity since there is a trivial solution, namely the uniform distribution over all strategies. It turns out that a simple modification of the game suffices to get a $\Omega(\varepsilon^{-1/2} \log n)$ lower bound. We add a small game on the side which dissuades from largely supported or 0-communication strategies. In particular, we show that for $\varepsilon < \frac{1}{10}$ any $\varepsilon$-CE must allocate most of its mass on the original chasing game. This allows us to use the results from the previous section to bound the cost of communicating equilibria for this game.

Construction. Consider the chasing game $CG_n$ from the previous section with a slight adjustment: give each player an additional action $n+1$. Choose $j_a, j_b \in [n]$ independently at random. We refer to $G_P$ as the main part of the game and the remainder as the auxiliary part of the game.

For $i \neq j_a$ make $u_A(n+1, i) = c_1$ and 0 otherwise. For $j \neq j_b$, make $u_B(j, n+1) = c_1$ and 0 otherwise. Make $u_A(j_b, n+1) = u_B(n+1, j_a) = c_1$, $u_A(i, n+1) = 0$ for all $i \in [n]\{j_b\}$, and $u_B(n, j) = 0$ for all $j \in [n+1]\{j_a\}$ and $u_A(n+1, n+1) = u_B(n+1, n+1) = c_2$ (see Figure 2 for an example).

As a simple exercise note that after this amendment the uniform distribution is no longer a $\varepsilon$-correlated equilibrium for $3\varepsilon \leq c_1$, given that $\varepsilon > \frac{1}{n}$. Any player can unilaterally switch to the new strategy and gain $\frac{n-1}{n}c_1 - \frac{1}{n}$ from deviating. We fix the values of these variables in the appendix, but roughly speaking $c_1, c_3$ are $O(\sqrt{\varepsilon})$ and $c_2$ is $O(\varepsilon)$. We can still show that $\Omega(\sqrt{\varepsilon})$ of the mass of the game remains on the non-zero entries of the main game, as stated in the following Lemma whose proof is on Appendix B.

Lemma 2. For any $\varepsilon$-CE, there is at least $\Omega(\sqrt{\varepsilon})$ fraction of the mass on the non-zero entries of the main part of the game and the support on these strategies must be $\Omega(\varepsilon^{-1})$. In particular, no $\varepsilon$-CE can have $1 - o(\sqrt{\varepsilon})$ fraction of the mass on the auxiliary part of the game or on the $(0,0)$ entries of the main part of the game.

Now we are ready to argue that computing $\varepsilon$-CE in the main part of the game requires $\Omega(\varepsilon^{-1/2} \log n)$ communication.

Lemma 3. Any protocol $\Pi$ that computes $\varepsilon$-CE of this game requires $\Omega\left(\frac{\log n}{\sqrt{\varepsilon}}\right)$ information cost for $\varepsilon > \Omega(1/n)$.

This Lemma directly implies Theorem 2. We suspect that this addition of auxiliary row and column to rule out equilibria with large support will be useful elsewhere. Interestingly, this is in direct contrast to Althofer games [?] which is used to rule out equilibria with small support.
3.3 Unrestricted lower bound for $G^m_2$

In this section we exhibit a game $G \in G^m_2$ whose $\varepsilon$-correlated equilibrium communication complexity is $\Omega(m)$. We defer the proof of Theorem 3 to Appendix B.

**Construction**  Without loss of generality suppose $m$ is an even number. Each player is equipped with two actions: 0 and 1. We will refer to the first $m/2$ as ‘state setters’ and define their payoffs as follows: let $\vec{R} \in \{0, 1\}^{m/2}$ be a string of random boolean variables where each coordinate is set independently at random at with probability 1/2. Then

$$u_i(a_i, \vec{a}_{-i}) = \begin{cases} 1 & \text{if } a_i = r_i \\ 0 & \text{otherwise} \end{cases}.$$  

We refer to the last $m/2$ players as ‘imitators’, and define their payoffs as follows:

$$u_i(a_i, \vec{a}_{-i}) = \begin{cases} 1 & \text{if } \vec{a}_{i-m/2} = a_i \\ 0 & \text{otherwise} \end{cases}.$$  

We state the following basic claims which implies the desired result and defer their proofs to Appendix B.

**Claim 2.** For any $\varepsilon$-CE, state setter $i$ must have $> 1 - \varepsilon$ mass on the recommended action $r_i$.

**Claim 3.** For any $\varepsilon$-CE, imitator $i$ must have mass $> 1 - \varepsilon$ on its setter’s action, $r_{i-m/2}$.

4 Upper Bounds for General Games

In this section we present a high level view of the communication protocol for computing $\varepsilon$-CE based on an adaptive algorithm of [HMC00] that converges to $\varepsilon$-HMCE. The overall communication cost of the protocol is $\tilde{O}(mn^4\varepsilon^{-4})$. The number of rounds is $\tilde{O}(n^4\varepsilon^{-4})$ in order to guarantee convergence to approximate correlated equilibria. The protocol is extremely simple: on each round a player picks a strategy, based on the previous history of actions played, and writes it on a blackboard shared by all the players. Thus for each round, there are $m \log n$ bits written on the blackboard. Note that the naive upper bound is $O(mn^m)$ where every player simply shares their utility functions.

More specifically, at round $t$ each player looks at the history of strategies $h(t) = \{s'_t : s \in A, (t' < t)\}$ played up until then and computes a matrix that measures the average regret of not having played action $k$ whenever action $j$ was played, for all actions $k, j \in A^i$ from round 1 up to round $t - 1$. They then compute the stationary distribution of this matrix (which [HMC00] shows that always exists), pick an action according to the stationary distribution and announce it to everyone else. After $T$ rounds, the first player (or any player) outputs the empirical distribution of actions played $z_T$, where for a given $s \in A$ the probability it gets played is $z_T(s) = \frac{1}{T} |\{t \leq T : s = s_t\}|$. The beauty of the procedure is that players only need to communicate the index of the action they perform at the current time period, using at most $O(\log n)$ bits of communication per player.

Intuitively, the matrix $A_t$ simply counts the regret of not having played action $k$ at time $t$ when action $j$ was played. The matrices $D_t, R_t$ average the regret and ignore actions with negative regret up until time $t$, respectively. It is clear that the communication cost of the protocol will be $Tm \log n$ where $T$ is the number of rounds.

Blackwell’s Approachability Theorem (with an appropriate martingale inequality) then guarantees that the $\ell_2$-norm of the matrix $A_t$ converges at the rate of $1/\varepsilon^4$ w.h.p. It is straightforward to
show that if $\ell_2$ norm of each row of $A_t$ is less than $\varepsilon$, then we have $\sqrt{n}\varepsilon$-CE. Combining these two observations, we obtain the bound of $T = \tilde{O}(n^4\varepsilon^{-4})$ on the number of rounds to guarantee convergence w.h.p.
References


A Omitted definitions

A.1 Information Theoretic definitions

Our communication lower bounds are actually based on information theoretic results, so here we provide the tools that will be used in Section 3. Throughout the paper log is the logarithm in base 2 and \(\ln\) is the natural logarithm. For further references, we refer the reader to [CT12].

**Definition 3 (Entropy).** The entropy of a random variable \(A\), denoted by \(H(A)\) is defined as

\[
H(A) = \sum_{a \in \text{Supp}(A)} \Pr[A = a] \log \frac{1}{\Pr[A = a]}.
\]

Intuitively this quantifies how much uncertainty we have about variable \(A\). This can be extended to define the relation between various variables. For instance suppose we have possibly correlated random variables \(A\) and \(B\). Then we can define conditional entropy of \(A\) given \(B\) as \(H(A|B) := H(AB) - H(B)\). Note that if \(A = B\), the conditional entropy is 0. We formalize this dependency as mutual information.

**Definition 4 (Mutual Information).** The mutual information between two random variables \(A\) and \(B\), denoted by \(I(A; B)\) is defined as

\[
I(A; B) := H(A) - H(A|B) = H(B) - H(B|A).
\]

The conditional mutual information between \(A\) and \(B\) given \(C\), denoted by \(I(A; B|C)\), is defined as

\[
I(A; B|C) := H(A|C) - H(A|BC) = H(B|C) - H(B|AC).
\]

This quantity measures how much information the random variable \(B\) reveals about \(A\) and vice-versa (even conditioned on the value of \(C\)).

We now provide useful properties that will be relevant to our proofs.

**Fact 1 (Chain Rule for Mutual Information).**

\[
I(X_1, \ldots, X_n; Y|Z) = I(X_1; Y|Z) + I(X_2; Y|Z, X_1) + \ldots + I(X_n; Y|Z, X_{n-1}, \ldots, X_1)
\]

**Definition 5 (Information Complexity).** The Information Cost of a 2-party protocol \(\Pi\) that computes \(f\) is defined as

\[
IC(\Pi) = I(\Pi; A|B) + I(\Pi; B|A),
\]

where \(A\) is the input to the first party and \(B\) is the to the second party. The information cost of \(f\) is simply the minimum information cost over all protocols that compute \(f\).

It is easy to show that for any protocol \(\Pi\) computing a function \(f\), \(CC(f, \Pi) \geq IC(f, \Pi)\), since 1-bit can carry at most 1-bit of information. Namely refer to [?]

B Proofs omitted from Section 3

B.1 Lower bounds for 2-player, \(n\) action games

*Proof of Claim 1.* We will first show that for any action for any player there is a unique best response. If Bob plays action \(i\) then Alice should simply respond with the unique action \(j = \sigma_A^{-1}(i)\)
to get a payoff of 1. This \( j \) may be the same as \( i \) but will not be the same as \( \sigma_B^{-1}(\sigma_A^{-1}(i)) \), the action that gives Bob a payoff of 1 (because \( \sigma_B \) is a cycle, and so is \( \sigma_B^{-1} \)). Moreover, for any two distinct actions by Bob, the best responses are distinct as well since the inverse of the permutation is well-defined. Therefore any strategy played by Alice is a best response to a strategy played by Bob.

Likewise, if Alice plays action \( i \), it is in Bob’s best interest to play \( j = \sigma_A(\sigma_B(i)) \) to get payoff 1, which is different from \( \sigma_A(i) \) which gives Alice a payoff of 1. Similarly, for any two distinct actions played by Alice, Bob’s response must be distinct. Therefore, any strategy played by Bob is a best response to a strategy played by Alice.

It is known that in 2 player games, an action is played on a Nash equilibrium if and only if it is a best response to an action by the other player. We argued that every action is a best response, so both players must play fully-supported strategies in equilibrium. We also need it is convenient to work out the results in general terms. Note that a non-zero payoff must have positive probability in an exact correlated equilibrium. Since \( \sigma_B \) is an \( n \)–cycle, chasing this argument will show that Alice gets recommended every action with some probability. Moreover, on any action that she is recommended we know her distribution must assign some probability on the unique strategy that gives Bob a non-zero payoff, otherwise he wouldn’t comply. Therefore we get that all \( 2n \) strategies where a player gets non-zero payoff must have positive probability in an exact correlated equilibrium.

Proof. Let \( a_i \) be Alice’s distribution conditioned on receiving recommendation \( i \). It must be the case that \( \sigma_A(i) \in \text{Supp}(a_i) \), since otherwise Alice would deviate to \( \sigma_A^{-1}(\arg \max_j a_{ij}) \). This also means that Bob must be recommended to play \( \sigma_A(i) \) with non-zero probability. When he gets that recommendation, by the same argument as before, there must be non-zero mass on Alice to play \( \sigma_B^{-1}(i) \), which is different from \( i \). Otherwise, Bob would disregard the recommendation \( \sigma_A(i) \).

So now we also know that Alice gets recommended \( j = \sigma_B^{-1}(i) \) with some probability and can use the same argument as before to show that \( \sigma_B^{-1}(\sigma_B^{-1}(i)) \) must also be recommended with some probability. Since \( \sigma_B \) is an \( n \)–cycle, chasing this argument will show that Alice gets recommended every action with some probability. Moreover, on any action that she is recommended we know that her distribution must assign some probability on the unique strategy that gives Bob a non-zero payoff, otherwise he wouldn’t comply. Therefore we get that all \( 2n \) strategies where a player gets non-zero payoff must have positive probability in an exact correlated equilibrium.

Proof. We will show that for any player, conditioned on being recommended action \( n + 1 \), the probability the other player assigns on action on \( n + 1 \) is at most \( 1/2 \). Suppose Alice is recommended action \( n + 1 \). Let \( p_{n+1,j} = \sum_{i \neq j} p_{n+1,i} \). Alice’s current payoff is \( p_{n+1,j} c_1 + p_{n+1,n+1} c_2 \). Since this is an approximate equilibrium we get the following inequality for deviating to \( j_a \),

\[
p_{n+1,j_a} c_1 + p_{n+1,n+1} c_2 + \varepsilon p_{n+1,*} \geq p_{n+1,n+1} c_3.
\]

Now since \( p_{n+1,*} - p_{n+1,n+1} \geq p_{n+1,*} - p_{n+1,n+1} = p_{n+1,j_a} \), we get

\[
c_1 p_{n+1,*} + \varepsilon p_{n+1,*} \geq p_{n+1,n+1}(c_3 - c_2 + c_1),
\]
and the rest follows from our choice of $c_1, c_2, c_3$ and our assumption on $\varepsilon$.

For the second statement, recall that Alice’s current payoff is exactly $c_1(p_{n+1, *} - p_{n+1, i*} - p_{n+1, n+1}) + c_2p_{n+1, n+1}$. Now if we consider a deviation to $\sigma_A^{-1}(j_a)$:

$$c_1(p_{n+1, *} - p_{n+1, ja} - p_{n+1, n+1}) + c_2p_{n+1, n+1} + \varepsilon p_{n+1, *} \geq p_{n+1, ja}.$$ 

Again rearranging we get,

$$p_{n+1, ja}(1 + c_1) \leq c_1 p_{n+1, *} + p_{n+1, n+1}(c_2 - c_1) + \varepsilon p_{n+1, *} \leq p_{n+1, *} + p_{n+1, n+1}(c_2 - c_1) + \varepsilon p_{n+1, *} \leq p_{n+1, *} + \varepsilon p_{n+1, n+1}.$$ 

where the last inequality uses the fact that $c_2 < c_1$. Then we get

$$\frac{c_1 + \varepsilon}{1 + c_1} p_{n+1, *} \geq p_{n+1, ja}.$$ 

The argument is symmetric for $p_{j_b, n+1}$. 

\[\blacksquare\]

**Claim 5 (Row Bound).** Consider an action $i \neq j_a$ for Alice. Then, $\frac{1 + \varepsilon}{1 + c_3} p_{i, *} \geq p_{i, n+1}$. Similarly for any action $j \neq j_b$ for Bob, $\frac{1 + \varepsilon}{1 + c_3} p_{*, j} \geq p_{n+1, j}$.

**Proof.** On recommendation $i$, Alice’s payoff is at most $p_{i, *} - p_{i, n+1}$. Now from deviating to $j_a$, we get

$$(p_{i, *} - p_{i, n+1}) + \varepsilon p_{i, *} \geq c_3 p_{i, n+1}.$$ 

Rearranging, we get the desired claim:

$$\frac{1 + \varepsilon}{1 + c_3} p_{i, *} \geq p_{i, n+1}.$$ 

Applying a symmetric argument for Bob, we get $\frac{1 + \varepsilon}{1 + c_3} p_{*, j} \geq p_{n+1, j}$ as well. 

\[\blacksquare\]

**Claim 6 (Main Part Bound).** Let $M_1 := \sum_{i=1}^{n} p_{i, *}$ and $M_2 := \sum_{j=1}^{n} p_{*, j}$. If $c_1 = c_3 = \sqrt{\varepsilon}$, $c_2 = \varepsilon$ and $\varepsilon < 1/100$, then at least one of these two inequalities must hold

$$M_1 \geq 1/4,$$

$$M_2 \geq 1/4.$$ 

**Proof.** First, observe that

$$M_1 = 1 - p_{n+1, n+1} - p_{j_a, n+1} - \sum_{i=1}^{n} p_{i, n+1}. \quad (1)$$ 

Recall that from Claim 4 and Claim 5, we get

$$p_{n+1, n+1} \leq \frac{c_1 + \varepsilon}{c_1 + c_3 - c_2} p_{n+1, *}, \quad (2)$$

$$p_{j_a, n+1} \leq \frac{c_1 + \varepsilon}{1 + c_1} p_{n+1, *}, \quad (3)$$

$$\sum_{i=1}^{n} p_{i, n} \leq \frac{1 + \varepsilon}{1 + c_3} \sum_{i=1}^{n} p_{i, *} \leq \frac{1 + \varepsilon}{1 + c_3} M_2. \quad (4)$$

14
Combining (2), (3) and (4), along with $M_1 = 1 - p_{n+1,*}$ we have

$$M_1 \geq 1 - \frac{c_1 + \varepsilon}{c_1 + c_3 - c_2} (1 - M_1) - \frac{c_1 + \varepsilon}{1 + c_1} (1 - M_1) - \frac{1 + \varepsilon}{1 + c_3} M_2. \quad (5)$$

By applying symmetric argument, we also get

$$M_2 \geq 1 - \frac{c_1 + \varepsilon}{c_1 + c_3 - c_2} (1 - M_2) - \frac{c_1 + \varepsilon}{1 + c_1} (1 - M_2) - \frac{1 + \varepsilon}{1 + c_3} M_1. \quad (6)$$

Let $M_0 := \frac{M_1 + M_2}{2}$. Then combining (5) and (6) then rearranging we get

$$\left(1 + \frac{1 + \varepsilon}{1 + c_3} - \frac{c_1 + \varepsilon}{c_1 + c_3 - c_2} - \frac{c_1 + \varepsilon}{1 + c_1}\right) M_0 \geq 1 - \frac{c_1 + \varepsilon}{1 + c_3} - \frac{c_1 + \varepsilon}{c_1 + c_3 - c_2}. \quad (7)$$

Thus for $\varepsilon < \frac{1}{100}$ and by our choice of $c_1, c_2, c_3$, we have $M_0 \geq 1/4$. Since $M_0$ is just the average of $M_1, M_2$ then at least one of the two must be greater than $1/4$.

**Claim 7** (Chasing Points). For all $i \in [n]$, $p_{i,\sigma_A(i)} + p_{i,\sigma_A(\sigma_B(i))} < 5\varepsilon$.

**Proof.** Suppose Alice is recommended with row $i$. If $i \neq j_a$, then from the payoff bound we get for any $j \in [n],

$$p_{i,\sigma_A(i)} + \varepsilon p_{i,*} > p_{i,j}. \quad (8)$$

In particular, this implies that $p_{i,\sigma_A(i)} + \varepsilon p_{i,*} > p_{i,\sigma_A(\sigma_B(i))}$. Otherwise, if $i = j_a$, then the payoff bound similarly gives for any $j \in [n],

$$p_{i,\sigma_A(i)} + c_3 p_{i,n+1} + \varepsilon p_{i,*} > p_{i,j}. \quad (9)$$

Via a symmetric argument for Bob we also get

$$p_{\sigma_A^{-1}(j_b),j_b} + c_3 p_{n+1,j_b} + \varepsilon p_{*,j_b} > p_{i,j_b}. \quad (10)$$

$$p_{\sigma_A^{-1}(j),j} + \varepsilon p_{*,j} > p_{i,j}. \quad (11)$$

Combining (8), (9), (10), (11) via applying them recursively, consider $R \subset [n] \times [n]$ where $(i,j) \in R$ if $j = \sigma_A(i)$ or $i = \sigma_A^{-1}(j)$. Then for any $(i,j),(k,l) \in R$, we have

$$p_{i,j} < p_{k,l} + \varepsilon \left( \sum_j p_{k,j} + \sum_i p_{i,*} \right) + c_3 p_{n+1,j_b} + c_3 p_{j_a,n+1}$$

$$< p_{k,l} + 2\varepsilon + c_3 p_{n+1,j_b} + c_3 p_{j_a,n+1}. \quad (12)$$

Note that (12) and our choice of $c_3$ implies that if there exists $(i,j) \in R$ such that $p_{i,j} < \varepsilon$ then $\forall (k,l) \in R, p_{k,l} < 5\varepsilon$. Suppose no such $(i,j)$ exists in $R$. Then

$$\sum_{(i,j) \in R} p_{i,j} > 2n\varepsilon.$$  

Since $\varepsilon > 1/n$, this is a contradiction.

**Claim 8** (Non-zero mass). Consider $N := \sum_{i=1}^n p_{i,\sigma_A(i)} + p_{i,\sigma_A(\sigma_B(i))}$, that is the total mass on the non-zero entries in the main game. Then $N \geq \frac{\sqrt{\varepsilon}}{20}$.
Proof. Suppose not. Recall that either \(M_1\) or \(M_2\) has a mass of 1/4. Without loss of generality, suppose \(M_1\) has at least mass of 1/4. If \(N < \frac{\sqrt{\varepsilon}}{20}\), note that we can rewrite \(N\) as

\[ N = M_1 \cdot \mathbb{E}_{i \sim \mathcal{P}} \left[ \frac{p_{i,A(i)} + p_{i,A(\sigma_B(i))}}{p_{i,\ast}} \right] < \sqrt{\varepsilon}, \]

where \(\mathcal{P}\) is defined as picking \(i\) with probability \(p_{i,\ast} / M_1\) for \(i \in [n]\). Since \(M_1 \geq 1/4\), There exists \(i \in [n]\) such that \(\frac{p_{i,A(i)} + p_{i,A(\sigma_B(i))}}{p_{i,\ast}} < \frac{\sqrt{\varepsilon}}{20}\).

Suppose that \(i \neq j_a\). If \(\frac{p_{j_a,b}}{p_{i,\ast}} > 1/10\), then consider a deviation to \(\sigma_A^{-1}(j_b)\). Then the new payoff is at least 1/10, this is a contradiction. If \(\frac{p_{j_{a,n+1}}}{p_{i,\ast}} > 1/10\), then this is again a contradiction by considering a deviation to \(j_a\), which guarantees a payoff of \(c_1/10 \gg \frac{\sqrt{\varepsilon}}{20} + \varepsilon\). Otherwise, consider deviating to \((n + 1)\)-row. Then the payoff is at least \(0.8c_1 \gg \frac{\sqrt{\varepsilon}}{20} + \varepsilon\), again a contradiction.

If \(i = j_a\), we divide into two cases depending on \(\frac{p_{j_{a,n+1}}}{p_{j_{a,\ast}}}\). Again note that the same argument shows that \(\frac{p_{j_{a,b}}}{p_{j_{a,\ast}}} < 1/10\). If \(\frac{p_{j_{a,n+1}}}{p_{j_{a,\ast}}} < 1/10\), then by deviating to action \((n + 1)\) guarantees payoff of 0.8\(c_1\). However, the current payoff is at most \(\frac{\sqrt{\varepsilon}}{20} + c_1/10\), which is a contradiction. Otherwise, it must be the case that \(\frac{p_{j_{a,n+1}}}{p_{j_{a,\ast}}} > 1/10\). Claim 4 shows that \(p_{j_{a,n+1}} < 2c_1\). Thus \(p_{j_{a,\ast}} < 20c_1 = O(\sqrt{\varepsilon})\). This implies that

\[ \mathbb{E}_{i \sim \mathcal{P}} \left[ \frac{p_{i,A(i)} + p_{i,A(\sigma_B(i))}}{p_{i,\ast}} \mid i \neq j_a \right] < \sqrt{\varepsilon}, \]

reducing to the case to \(i \neq j_a\). \(\square\)

Note that Claim 7 and Claim 8 imply that for any \(\varepsilon\)-CE at least \(\Omega(1/\sqrt{\varepsilon})\) of the entries in the permutation matrices are in the support of the \(\varepsilon\)-CE. Intuitively, this means that either Alice or Bob learns about \(\Omega(1/\sqrt{\varepsilon})\)-bits of information about the random permutation. We make this observation concrete in the following claim and lemma.

Claim 9. Consider row \(i \neq j_a\). Then \(p_{i,A(i)}/p_{i,\ast} > \Omega(\sqrt{\varepsilon})\).

Proof. Denote \(p := p_{i,b}/p_{i,\ast}\) and \(q := p_{i,n+1}/p_{i,\ast}\). Then from deviations to \(\sigma_A^{-1}(j_b)\), \(j_a\) and \((n + 1)\) we get the following set of inequalities:

\[ p_{i,A(i)}/p_{i,\ast} + \varepsilon > p, \]
\[ p_{i,A(i)}/p_{i,\ast} + \varepsilon > c_1q, \]
\[ p_{i,A(i)}/p_{i,\ast} + \varepsilon > c_1(1 - p - q) + c_2q. \]

To satisfy all these inequalities it is necessary that \(p_{i,A(i)}/p_{i,\ast} > \Omega(c_1)\). Since \(c_1 = \sqrt{\varepsilon}\), this proves the claim. \(\square\)

Lemma 4. Any protocol \(\Pi\) that computes \(\varepsilon\)-CE requires \(\Omega\left(\frac{\log n}{\sqrt{\varepsilon}}\right)\) information cost for \(\varepsilon > \Omega(1/n)\).

Proof. We show that we can extract \(\Omega\left(\frac{\log n}{\sqrt{\varepsilon}}\right)\)-bits of information about the random permutation from marginal distribution induced by \(\varepsilon\)-CE. Let \(R_1 := \{(i, \sigma_A(i)) \mid i \in [n]\}\) and w.l.o.g. suppose at least \(\Omega(1/\sqrt{\varepsilon})\) elements in \(R_1\) is contained in the support of \(\varepsilon\)-CE as from Claim 7 and Claim 8. Let \(R_{\text{supp}}\) denote the set of first index in \(R_1\) that is included in the support of \(\varepsilon\)-CE. Then by Fact 2, we can write \(I(X; \Pi|Y)\) (i.e information learned by Bob about Alice’s input) as

\[ I(X; \Pi|Y) = \sum_{i \in R_{\text{supp}}} I(X_i; \Pi|Y, \{X_r\}_{r \in R_{\text{supp}}}). \]
Now for each term we can rewrite as
\[ I(X; \Pi | Y, \{X_r\}_{r \in R_{\text{supp}}}) = H(X_i | Y, \{X_r\}_{r \in R_{\text{supp}}}) - H(X_i | \Pi, Y, \{X_r\}_{r \in R_{\text{supp}}}) \]
by the definition of mutual information. Note that
\[ H(X_i | Y, \{X_r\}_{r \in R_{\text{supp}}}) = H(X_i | \{X_r\}_{r \in R_{\text{supp}}}) \geq \Omega(\log(n - O(1/\sqrt{\varepsilon}))) \geq \Omega(\log n). \]
While after running the protocol, the mass on \((i, \sigma_A(i))\)-entry is at least \(\Omega(\sqrt{\varepsilon})\). In other words, the support size for the possible \(\sigma_A(i)\) is at most \(O(1/\sqrt{\varepsilon})\). Then we have
\[ H(X_i | \Pi, Y, X_{i-1}, \ldots, X_1) \leq O\left(\log \frac{1}{\sqrt{\varepsilon}}\right). \]
Then combining two bounds we get
\[ I(X; \Pi | Y) = \sum_{i \in R_{\text{supp}}} \Omega(\log \sqrt{\varepsilon}n) \geq \Omega\left(\frac{\log n}{\sqrt{\varepsilon}}\right). \]
since \(\varepsilon = \Omega(1/n)\). Thus the information cost of \(\Pi\) is at least \(\Omega\left(\frac{\log n}{\sqrt{\varepsilon}}\right)\).

C Proofs omitted from Section 4

First, we formally state the protocol in Algorithm 1.

**Algorithm 1** Protocol \(\Pi\) to compute \(\varepsilon\)-CE

- At time \(t = 1\), each player plays according to some arbitrary initial distribution \(p^i_1\).
- From \(t = 2\) to \(t = T\):
  - Each player \(i\) computes the matrices \(A^i_t, D^i_t, R^i_t\) where
    \[ A^i_t(j, k) = 1_{s^i_t=j}|u^i(k, s^i_{t-1}) - u^i(j, s^i_{t-1})|, \]
    \[ D^i_t(j, k) = \frac{1}{t} \sum_{\tau=1}^t A^i_t(j, k), \]
    \[ R^i_t(j, k) = \max\{0, D^i_t(j, k)\}. \]
  - Each player computes the stationary distribution of \(R^i_t, p^i_t\), plays according to it and announces his move to everyone else.
- At the end of the protocol, each player computes the empirical distribution \(z_T\) of all strategies played.

It suffices to bound \(T\) such that guarantees \(\varepsilon\)-CE as an output. To bound \(T\), we use the \(\ell_2\) norm of the regret matrix as a “potential” function. This is indeed a natural candidate for the potential function due to the following lemmas.

We connect \(\max_j \{\sum_k R^i_t(j, k)\}\) to \(\varepsilon\)-CE through following observation.

**Lemma 5.** Consider any sequence of plays and let \(\varepsilon \geq 0\). If \(\limsup_{t \to \infty} \max_j \{\sum_k R^i_t(j, k)\} \leq \varepsilon\) for all players \(i\) and all actions \(j \in A^i\), the sequence of empirical distributions converges to the set of \(\varepsilon\)-CE.
Then the empirical distribution of plays $z_T$ converges as $T \to \infty$ to the stationary distribution of Theorem 4 ([HMC00]).

The standard setup considers an agent $i$ who plays actions $a_i \in A_i$ and gets payoff vectors in $\mathbb{R}^L$, that depends on his action and another action $a^{-i} \in A^{-i}$ chosen by an opponent, possibly adversarially. In other words, agent $i$’s payoff is of the form $v_i : A^i \times A^{-i} \to \mathbb{R}^L$. The game is played for many rounds and the agents goal is to make her average payoff vector, $D_T = \frac{1}{T} \sum_{t=1}^T v_i(a^i_t, a^{-i}_t)$, approach some given set $C \in \mathbb{R}^L$. We say a convex, closed set $C$ is approachable if there is a procedure that almost surely guarantees that the $\ell_2$ distance between $D_T, C$ approaches 0 as $T$ goes to $\infty$, irrespective of the opponents actions. Blackwell’s Approachability Theorem (see Appendix C) states necessary and sufficient conditions under which this can be done. More precisely, the probability that the proposed strategy is far from the set decays with the following martingale bound.

\[ \text{Proof.} \] For each player $i$ and every $j \in A_i$ we have
\[
\sum_{k \neq j} D_T^i(j, k) = \frac{1}{T} \sum_{t=1}^T \sum_{k \neq j} A_T(j, k) = \sum_{s \in A : s_i = j} z_T(s) \sum_{k \neq j} (u^i(k, s^{-i}) - u^i(j, s^{-i})) = \mathbb{E}_{u \sim z_T} [R_T^i(s)] \leq \varepsilon,
\]
for any switching rule $f$ where the last equality follows from the fact that a switching rule is just a linear combination of single deviations. Since the sum of the regret of all individual actions is bounded by $\varepsilon$ so is any convex combination of them. \hfill \square

Unfortunately, we do not immediately get $\lim \sup_{T \to \infty} \max_j \{\sum_k R_T^i(j, k)\} \leq \varepsilon$ in terms of the convergence rate, rather the guarantee bounds are on $||R_T^i||_2$. However, it is not hard to show that one implies the other with a loss in the approximation parameter.

**Lemma 6** ($\ell_2$ bound translation). *If the regret matrix for each player $i$ satisfies $\sqrt{\sum_k R_T^i(j, k)^2} \leq \varepsilon$ for all $j$, then we obtain a $\sqrt{n\varepsilon}$-CE.*

**Proof.** By Lemma 5 it suffices to show that $\max_j \{\sum_k R_T^i(j, k)\} \leq \sqrt{n\varepsilon}$.
\[
\max_j \left\{ \sum_k R_T^i(j, k) \right\} \leq \sqrt{n} \max_j \sqrt{\sum_k R_T^i(j, k)^2} \leq \sqrt{n} \sqrt{\sum_{j,k} R_T^i(j, k)^2} \leq \sqrt{n}\varepsilon,
\]
where the first inequality follows from the relationship between the $\ell_1$ and $\ell_2$ norms, and the second follows from the relationship between $\ell_\infty$ and $\ell_1$. \hfill \square

These two lemmas imply that if $\forall j, k \sum_k R_T^i(j, k)^2 < \varepsilon / \sqrt{n}$, then the empirical distribution indeed forms an $\varepsilon$-CE. The proof in [HMC00] shows the following theorem in a restricted setting.\footnote{Though it is not mentioned explicitly, the analysis assumes that the number of actions per player is $O(1)$ in [HMC00].}

**Theorem 4 ([HMC00]).** *Suppose that at each period $t + 1$ every player $i$ chooses strategies according to the stationary distribution of $R_T^i$. Furthermore, suppose the number of actions per player is $O(1)$. Then the empirical distribution of plays $z_T$ converges as $||R_T^i||_2 < \varepsilon$ if $T \geq 1/\varepsilon^2$.***
Lemma 7 (Section 4 of [FV99]).

\[ \Pr [d(D_T, C) \geq \delta] < e^{-\delta^4 T/98R^4} \]

where \( R \) is the largest distance between any two points in the set of possible payoffs.

From a high level view, [HMC00] uses the result directly with \( L = n^2 \) and the individual vector payoffs \( v_i \) as the regret matrix \( R_i \). They show that the stationary distribution of the regret matrices is exactly the vector \( q_\lambda \) that guarantees convergence in Blackwell’s Theorem. Unfortunately the dimension of the vectors plays a role in the convergence rate, and we need to take a close look into the proof of the Theorem.

Let \( \rho_t \) be the distance between the average payoff vector of agent \( i \) and the convex set \( C \). The analysis of Blackwell’s Theorem relies on recursively bounding \( \rho_{t+1}^2 \) as a function of \( \rho_t^2 \). Reorganizing the terms yields the following recursion:

\[ (t+1)^2 \rho_{t+1}^2 \leq t^2 \rho_t^2 + ||v_i(a_i, a^{-i}) - \pi_C(D_t)||^2, \]

where \( \pi_C(D_t) \) is the projection of the average regret vector at time \( t \) onto set \( C \). In many applications of Blackwell’s Theorem, \( C \) is contained in bounded region as well as the payoff vector \( v_i \). Thus, a standard analysis would bound the rightmost term by a constant (arguing that all points belong to some ball of bounded radius) and, with the use of a telescoping argument, show that the distance converges at a rate of \( O(\frac{1}{\sqrt{T}}) \).

However, in our case, if the dimension is a parameter we care about, then an appropriate upper bound in the worst case on the rightmost term would be \( O(L) \) (this is because the vectors lie on \( L \)-dimensional space and are entry-wise bounded due to the nature of the utility functions.). Carrying the standard analysis as is would then give a convergence rate of \( O(\frac{n}{\sqrt{T}}) \), which would in turn significantly blow up the cost of our communication protocol. In order to obtain \( \varepsilon \)-CE, we need \( \frac{n}{\sqrt{T}} < \frac{\varepsilon}{\sqrt{m}} \), and thus \( T > n^3/\varepsilon \), which is in fact worse than a naive protocol for 2-player setting.

Even worsening the problem, the rate of the convergence is in expectation, while we must argue that the bound holds with high probability. The Martingale bound guaranteed by Lemma 7 heavily depends on the dimension due to the \( R \) factor, which is \( n \) in our setting, since the payoff vectors are only bounded by 1 in \( \ell_{\infty} \) norm.

To fully exploit Lemma 6, instead of bounding the \( \ell_2 \) norm of the whole matrix, we bound \( \sqrt{\sum_k R_i^k(j, k)^2} \), that is the \( \ell_2 \) norm of each row via the same adaptive process.

Lemma 8. Protocol \( \Pi \) with \( T = O\left(\frac{n^4 \log(mn)}{\varepsilon^4}\right) \) rounds produces a \( \varepsilon \)-CE.

Proof. With \( T = O\left(\frac{n^4 \log(mn)}{\varepsilon^4}\right) \) rounds in Theorem 7 guarantee that

\[ \Pr \left[ \sqrt{\sum_k R_i^k(j, k)^2} \geq \frac{\varepsilon}{\sqrt{n}} \right] < 1/(mn)^3 \]

By applying the union bound, this implies that with high probability for all players \( i \), \( \sqrt{\sum_k R_i^k(j, k)^2} \leq \varepsilon \) for all \( j \).

Proof of Theorem 1. Combining Lemma 8 with Lemma 6 finishes the proof of Theorem 1: the protocol runs for \( \tilde{O}(n^4 \varepsilon^{-4}) \) rounds and each round requires \( O(m \log n) \) bits of communication. \[\square\]