The Parallel Repetition of Non-Signaling Games: Counterexamples and Dichotomy

Justin Holmgren[*]
justin.holmgren@princeton.edu

Lisa Yang
lisayang@mit.edu

Abstract

Non-signaling games are an important object of study in the theory of computation, for their role both in quantum information and in (classical) cryptography. In this work, we study the behavior of these games under parallel repetition.

We show that, unlike the situation both for *classical* games and for *two-player non-signaling* games, there are *k-player non-signaling games* (for $k \geq 3$) whose values do not tend to 0 with sufficient parallel repetition. In fact, parallel repetition sometimes does not decrease their value whatsoever.

We show that in general:

1. Every game's non-signaling value under parallel repetition is either lower bounded by a positive constant or decreases exponentially with the number of repetitions.
2. Exponential decrease occurs *if and only if* the game's sub-non-signaling value (Lancien and Winter, CJTCS '16) is less than 1.

[*] Research conducted while this author was a graduate student at MIT, supported in part by the NSF MACS project – CNS-1413920, DARPA IBM – W911NF-15-C-0236, and the SIMONS Investigator Award Agreement Dated 6-5-12

1 Introduction

A multi-player game $\mathcal{G}$ consists of an interaction between a referee and k players $P_1, \ldots, P_k$. The referee samples k questions $q_1, \ldots, q_k$ from some joint distribution π and sends q_i to P_i. The players then, without communicating amongst themelves, respond with answers $a_1, \ldots, a_k$, and are judged to *win* or *lose* the game according to a predicate $W(q_1, \ldots, q_k, a_1, \ldots, a_k)$. The (classical) value of the game, denoted $v(\mathcal{G})$, is the maximum probability with which players can win. The study of multi-player games has been profoundly impactful in diverse areas of theoretical computer science, with foundational applications in complexity theory, hardness of approximation, and cryptography (see e.g., [BGKW88, BFL91, BFLS91, Kil92, FGL$^+$96, ALM$^+$98, Hås01]).

Parallel repetition is a natural and important operation on any game, originally motivated by its potential application to the soundness amplification of interactive proof systems [FRS88, FL92]. In the n-fold parallel repetition of $\mathcal{G}$, denoted $\mathcal{G}^n$, the referee now samples n independent sets of questions $\{(q_1^{(i)}, \ldots, q_k^{(i)})\}_{i=1}^n$ from π, and sends *all* n questions $\{q_j^{(i)}\}_{i=1}^n$ to the j^{th} player at once for each $1 \leq j \leq k$. The players are then required to win all n copies of the game; that is, the j^{th} player must produce $\{a_j^{(i)}\}_{i=1}^n$ so that $W(q_1^{(i)}, \ldots, q_k^{(i)}, a_1^{(i)}, \ldots, a_k^{(i)}) = 1$ for every $1 \leq i \leq n$. Despite the simplicity of this transformation, it is surprisingly tricky to analyze. It was initially claimed that $v(\mathcal{G}^n)$ is always equal to $v(\mathcal{G})^n$ [FRS88]. This was quickly disproved, however, by Fortnow [For89] and Feige [Fei91], who constructed a two-player game $\mathcal{G}$ satisfying $v(\mathcal{G}^2) = v(\mathcal{G}) < 1$. However, Raz's celebrated parallel repetition theorem proves that for any non-trivial two-player game, the value of $\mathcal{G}^n$ decreases exponentially with n [Raz98]. That is, if $v(\mathcal{G}) < 1$, then $v(\mathcal{G}^n) \leq \bar{v}^n$ for some $\bar{v} < 1$ that depends on $\mathcal{G}$ (but may be significantly larger than $v(\mathcal{G})$).

An important variant on the aforementioned classical value of a game is its non-signaling value v_{ns}, originally defined with the intent of modeling players that may share quantum entanglement. This is the players' maximum winning probability with a relaxed non-communication restriction. Specifically, they may communicate freely with one another to generate any distribution of answers $(a_1, \ldots, a_k)$ with a single caveat: for each $S \subseteq [k]$, the *marginal distribution* a_S must be a function only of q_S. While the distinction is perhaps unintuitive, a game's non-signaling value and classical value often differ. The usual example of this is the (two-player) CHSH game [CHSH69], whose classical and non-signaling values are respectively 3/4 and 1.[1]

Specific non-signaling games have also recently found applications in cryptography. In particular, multi-prover interactive proofs (MIPs) with soundness against non-signaling provers [IKM09, KRR13, KRR14] are indispensable in obtaining non-interactive delegation of computation from standard assumptions [KRR13, KRR14, BHK17, BKK$^+$18]. In this work, we focus on the following question:

How does parallel repetition affect a game's non-signaling value?

Prior Work on Non-Signaling Parallel Repetition In a simplification and generalization of Raz's work, Holenstein showed that if a *two*-player game $\mathcal{G}$ has non-signaling value $v < 1$, then the non-signaling value of $\mathcal{G}^n$ is bounded by $\bar{v}^n$ for some $\bar{v} < 1$ that depends only on v [Hol09].

The multi-player case was considered by Buhrman, Fehr, and Schaffner, who proved a similar result for *complete support games* – games in which every combination of queries is asked with positive probability [BFS14]. Their result comes with the caveat that $\bar{v}$ depends not only on v, but also on the minimum probability with which a query is asked. Arnon-Friedman, Renner, and Vidick give an alternative proof of the same result, and they further note that any game $\mathcal{G}$ can be modified (by adding dummy queries) to have complete support without significantly affecting its non-signaling value [FRV16].

At this point, we find it prudent to point out that games without complete support are quite important in "practical" applications. In particular the non-signaling MIPs of [KRR13, KRR14] and of follow-up works heavily rely on games whose query distributions have incomplete support. These MIPs can be viewed as a family of games indexed by binary strings x, where for some time-T decidable language $\mathcal{L}$, the non-signaling

[1] An arbitrarily large gap between the two values can be achieved by parallel repeating the CHSH game.

value of game x indicates whether $x \in \mathcal{L}$. In the MIPs of [KRR13, KRR14] and follow-ups, even the *support* of the query distribution has size that is superpolynomial in T. Thus, the "dummy query" approach of [FRV16] (or any approach that only tweaks the query distribution) *inherently* assigns $T^{-\omega(1)}$ probability to some query. As a result, the aforementioned bounds of [BFS14, FRV16] do not achieve even a constant soundness error unless the number of repetitions is super-polynomial in T. This corresponds to a trivial MIP, as a verifier can check for himself whether $x \in \mathcal{L}$ in time T with no error.

The results of [BFS14, FRV16] were generalized by Lancien and Winter [LW16], who defined a further relaxation of a game's non-signaling value, called its *sub*-non-signaling value (v_{sns}). They show that this value does behave nicely under parallel repetition: for any k-player game $\mathcal{G}$ with $v_{\mathsf{sns}}(\mathcal{G}) < 1$, it holds for all n that $v_{\mathsf{sns}}(\mathcal{G}^n) \leq \bar{v}^n$ for a value $\bar{v} < 1$ that depends only on v and k.

Still, the question of how parallel repetition can affect a general game's non-signaling value has remained open.

1.1 Our Results

Perhaps our most surprising result is the following counterexample to parallel repetition. In multi-player classical games [Ver96], two-player entangled games [Yue16], and two-player non-signaling games [Hol09], it is known that the value of any non-trivial repeated game decreases to 0 after sufficiently many parallel repetitions. For multi-player non-signaling games, we show that this does not hold in general.

Theorem 1 (Parallel Repetition Counterexample). *For every $k \geq 3$, there exists a k-player game $\mathcal{G}$ such that $v_{\mathsf{ns}}(\mathcal{G}) < 1$ and $v_{\mathsf{ns}}(\mathcal{G}^n) \geq \Omega(1)$ for all $n \geq 1$.*

We also prove that if the non-signaling value of the repeated game decays to 0, then the decay must be exponential. This establishes a converse to the main result of [LW16], and is not at all obvious. In fact, there are several contexts (e.g., classical multi-player games and entangled two-player games) where it is known that the value of a game decays to 0 with sufficient parallel repetition, but it is not known (and is a major open question) whether this decay is always exponential.

Theorem 2 (Parallel Repetition Dichotomy). *For every game $\mathcal{G}$, either $v_{\mathsf{ns}}(\mathcal{G}^n) \geq \Omega(1)$ or $v_{\mathsf{ns}}(\mathcal{G}^n) \leq \exp(-\Omega(n))$. The former occurs when $v_{\mathsf{sns}}(\mathcal{G}) = 1$ and the latter occurs when $v_{\mathsf{sns}}(\mathcal{G}) < 1$.*

Thus our counterexample is general: for *any* game $\mathcal{G}$ with sub-non-signaling value 1, it holds that $v_{\mathsf{ns}}(\mathcal{G}^n) \geq \Omega(1)$. Finally, we show that if a game has sufficiently small *non-signaling* value, then the non-signaling value of the repeated game decays exponentially.

Theorem 3 (Parallel Repetition Magic Value). *For every k, there exists a constant $\alpha_k > 0$ such that if $v_{\mathsf{ns}}(\mathcal{G}) < \alpha_k$, then $v_{\mathsf{ns}}(\mathcal{G}^n) \leq \exp(-\Omega(n))$.*

Theorems 1 to 3 each directly follow from our main lemma, which is simple both to state and to use, and may be of independent interest.

Lemma 1 (Non-signaling Value Lower Bound). *For every k, there exists a constant $\alpha_k > 0$ such that for any k-player game $\mathcal{G}$ with $v_{\mathsf{sns}}(\mathcal{G}) = 1$, it holds that $v_{\mathsf{ns}}(\mathcal{G}) \geq \alpha_k$. Moreover, the maximum value of α_k for which this holds is bounded by $2^{-O(k^2)} \leq \alpha_k \leq 2^{O(k)}$.*

Beyond our general results, we also show that there are *specific* games whose non-signaling values do not decrease *at all* under parallel repetition, subsuming Theorem 1.

Theorem 4 (Strong Parallel Repetition Counterexample). *For every $k \geq 3$, there exists a k-player game $\mathcal{G}$ such that $v_{\mathsf{ns}}(\mathcal{G}) < 1$ and $v_{\mathsf{ns}}(\mathcal{G}^n) = v_{\mathsf{ns}}(\mathcal{G})$ for any $n \geq 1$.*

	Two-player games	Multi-player games
Classical	$\exp\left(-\Omega\left(\frac{\epsilon^3 n}{\log\lvert\mathcal{A}\rvert}\right)\right)$ [Raz98, Hol09]	$O\left(\frac{1}{\mathsf{Ackermann}^{-1}(n)}\right)$ [Ver96]
Non-Signaling	$\exp\left(-\Omega\left(\epsilon^2 n\right)\right)$ [Hol09]	$\geq \Omega(1)$ [**This Work**]

Table 1: Known bounds on the worst-case (slowest) decay for $v(\mathcal{G}^n)$ or $v_{\mathsf{ns}}(\mathcal{G}^n)$ for a game $\mathcal{G}$ with $v(\mathcal{G}) = 1-\epsilon$ or $v_{\mathsf{ns}}(\mathcal{G}) = 1-\epsilon$ respectively. $\mathcal{A}$ denotes the set of possible player answers in $\mathcal{G}$. $\mathsf{Ackermann}^{-1}$ denotes the inverse Ackermann function.

1.2 Our Techniques

We give an overview of the proof of our main technical lemma, Lemma 1. We prove Lemma 1 without referencing sub-non-signaling strategies by (1) introducing a condition that is equivalent to a game $\mathcal{G}$ having sub-non-signaling value 1 and (2) proving implications of this condition on the non-signaling value of $\mathcal{G}$.

Suppose that $\mathcal{G}$ is a k-player game in which the joint distribution of questions asked is π (over a set $\mathcal{Q} = \mathcal{Q}_1 \times \cdots \times \mathcal{Q}_k$), and the i^{th} player's answer is expected to lie in a set $\mathcal{A}_i$. Let $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_k$. We will often say "the support of $\mathcal{G}$" to refer to the support of its question distribution π. We say "the S-marginal support of $\mathcal{G}$" to refer to the support of the marginal distribution π_S.

An honest-referee non-signaling strategy for $\mathcal{G}$ is a function $\tilde{P}$ that maps $q = (q_1, \ldots, q_k)$ *in the support* of $\mathcal{G}$ to a distribution on $\mathcal{A}$, satisfying the following non-signaling "consistency" condition: for any q, q' in the support of $\mathcal{G}$, if $S \subseteq [k]$ is a set of players for which $q_S = q'_S$, then it must also hold that $\tilde{P}(q)_S = \tilde{P}(q')_S$. This is in contrast to a (full) non-signaling strategy, which must be defined on the entirety of $\mathcal{Q}$ and must satisfy the non-signaling condition for any $q, q' \in \mathcal{Q}$.

Remark 1.1. *A common reflexive response to the definition of (full) non-signaling is that it somehow is too stringent a requirement. After all, why must an (adversarial) player have a well-defined behavior on queries that are never asked?*

In many settings though, adversaries are *automatically forced to have a well-defined behavior, even on queries that are not typically asked. For instance, if an adversarial player strategy is a collection of physical devices, then each device will do* something *on any query (including possibly refusing to answer or giving an error message, in which case we say the device outputs $\perp$). But this constitutes a response, which must not signal.*

Our conclusion is that an honest-referee non-signaling strategy is, on its own, reasonable only if it extends to a full non-signaling strategy. However, as we demonstrate, the notion of an honest-referee non-signaling strategy remains useful as a tool for the analysis and construction of non-signaling strategies.

It turns out (Proposition 3.6) that a game $\mathcal{G}$ has sub-non-signaling value 1 iff there is an honest-referee non-signaling strategy $\tilde{P}$ that wins $\mathcal{G}$ with probability 1. To prove Lemma 1, we construct a full non-signaling strategy P that has significant "agreement" with $\tilde{P}$ for any honest-referee non-signaling strategy $\tilde{P}$.

A natural first attempt at constructing P might be to somehow extend $\tilde{P}$ into a full non-signaling strategy. However, this approach is doomed to fail in general. A specific counterexample is given by the "anti-correlation" games, first defined in [LW16, FRV16]. We include a formal description of these games for completeness (and with slightly greater generality) in Definition 4.2. These games have honest-referee non-signaling value 1 but non-signaling value less than 1.

Instead, we show that it is always possible to extend $\alpha_k \cdot \tilde{P}$ for some constant $0 < \alpha_k < 1$. Specifically, we construct a non-signaling strategy P such that for every q in the support of $\mathcal{G}$ and every $a \in \mathcal{A}$, it holds that $\Pr_{A \leftarrow P(q)}[A = a] \geq \alpha_k \cdot \Pr_{A \leftarrow \tilde{P}(q)}[A = a]$. This implies that P's winning probability for $\mathcal{G}$ is at least α_k times that of $\tilde{P}$.

1.2.1 Three-Player Games

For simplicity of exposition, we first consider three-player games.

The following somewhat easier task serves both as a building block and as a warm-up in our construction of P. For every $S \subseteq \{1,2,3\}$ with $|S| \leq 2$ and every q in $\mathcal{Q}_S$ (in particular, q may not be in the S-marginal support of $\mathcal{G}$), we construct a distribution $\mu_S(q)$ on $\mathcal{A}_S$ such that $\{\mu_S(q)\}_{S,q}$ satisfies the following consistency conditions.

Condition 1 (Agreement with $\tilde{P}$)**.** *For every q in the support of $\mathcal{G}$ and every $S \subseteq \{1,2,3\}$ with $|S| \leq 2$, the marginal distribution $\tilde{P}(q)_S$ is equal to $\mu_S(q_S)$.*

Condition 2 (Non-Signaling)**.** *The distribution $\mu_S(q)_T$ depends only on T and q_T, and not on S or the rest of q. Formally: for any $S, S' \subseteq [3]$ with $|S|, |S'| \leq 2$, any $T \subseteq S \cap S'$, any $q \in \mathcal{Q}_S$ and $q' \in \mathcal{Q}_{S'}$, if $q_T = q'_T$ then $\mu_S(q)_T = \mu_{S'}(q')_T$.*

We define $\mu_S(q)$ in one of two different ways, depending on whether q is in the S-marginal support of $\mathcal{G}$.

> **If Yes:** Let $\bar{q}$ be in the support of $\mathcal{G}$ such that $\bar{q}_S = q$. We then define $\mu_S(q) \overset{\mathsf{def}}{=} \tilde{P}(\bar{q})_S$. This is well-defined (i.e. does not depend on the choice of $\bar{q}$) because $\tilde{P}$ is honest-referee non-signaling.
>
> This "Yes" case behavior is sufficient to ensure that Condition 1 holds.
>
> **If No:** This means that $|S| = 2$. Say that $S = \{i, j\}$ for $i < j$. We define $\mu_S(q)$ to be the distribution on (a_i, a_j) obtained by independently sampling $a_i \leftarrow \mu_{\{i\}}(q_i)$ and $a_j \leftarrow \mu_{\{j\}}(q_j)$.

Condition 2 holds because if $(S, q) \neq (S', q')$ but $q_T = q'_T$, then $|T| = 1$. Both the "Yes" case and the "No" case define $\mu_S(q)$ so that any 1-marginal $\mu_S(q)_T$ depends only on q_T. In the "Yes" case this follows from the fact that $\tilde{P}$ is honest-referee non-signaling, and in the "No" case this follows from reduction to the "Yes" case.

We now construct the (full) non-signaling strategy P in terms of the distributions $\{\mu_S(q)\}$. We again have two cases, this time corresponding to whether q is in the support of $\mathcal{G}$.

> **If Yes:** We define $P(q)$ to be the distribution on (a_1, a_2, a_3) obtained as follows. With probability $1/3$ sample $a \leftarrow \tilde{P}(q)$. Otherwise (with probability $2/3$), independently sample $a_1 \leftarrow \tilde{P}(q)_1$, $a_2 \leftarrow \tilde{P}(q)_2$, and $a_3 \leftarrow \tilde{P}(q)_3$.
>
> This "Yes" case behavior is sufficient to ensure that P extends $\frac{1}{3} \cdot \tilde{P}$.
>
> **If No:** We define $P(q)$ to be the distribution on (a_1, a_2, a_3) obtained as follows. Pick a uniformly random $i \leftarrow \{1,2,3\}$, define $S = \{1,2,3\} \setminus \{i\}$, and independently sample $a_S \leftarrow \mu_S(q_S)$ and $a_i \leftarrow \mu_{\{i\}}(q_i)$.

It remains to verify that P is non-signaling. For this, we must check that for all $q, q' \in \mathcal{Q}$ and all $S \subseteq \{1,2,3\}$ such that $q_S = q'_S$, it holds that

$$P(q)_S = P(q')_S. \tag{1}$$

If $|S| = 0$ then Eq. (1) is vacuous, and if $|S| = 3$ then Eq. (1) is trivial. We claim that without loss of generality, we can focus on $|S| = 2$ as the remaining case. Indeed, suppose that we have established Eq. (1) for all $|S| = 2$ (and all q, q' for which $q_S = q'_S$). If $q_T = q'_T$ for $|T| = 1$, then we can construct $\hat{q} \in \mathcal{Q}$ and sets $S, S' \supset T$ (with $|S| = |S'| = 2$) such that $q_S = \hat{q}_S$ and $\hat{q}_{S'} = q'_{S'}$. Thus,

$$P(q)_T = \big(P(q)_S\big)_T = \big(P(\hat{q})_S\big)_T = P(\hat{q})_T = \big(P(\hat{q})_{S'}\big)_T = \big(P(q')_{S'}\big)_T = P(q')_T. \tag{2}$$

Now without loss of generality suppose that $S = \{1,2\}$ (the other cases follow analogously due to the symmetry of P). For any $q \in \mathcal{Q}$, if q is in the support of $\mathcal{G}$, then

$$P(q) = \frac{1}{3} \cdot \tilde{P}(q)_{\{1,2,3\}} + \frac{2}{3} \cdot \tilde{P}(q)_{\{1\}} \times \tilde{P}(q)_{\{2\}} \times \tilde{P}(q)_{\{3\}}$$

and so

$$P(q)_S = \frac{1}{3} \cdot \tilde{P}(q)_{\{1,2\}} + \frac{2}{3} \cdot \tilde{P}(q)_{\{1\}} \times \tilde{P}(q)_{\{2\}}. \tag{3}$$

If q is not in the support of $\mathcal{G}$, then

$$P(q) = \frac{1}{3} \cdot \begin{pmatrix} \mu_{\{1,2\}}(q_{\{1,2\}}) \times \mu_{\{3\}}(q_{\{3\}}) \\ +\mu_{\{1,3\}}(q_{\{1,3\}}) \times \mu_{\{2\}}(q_{\{2\}}) \\ +\mu_{\{2,3\}}(q_{\{2,3\}}) \times \mu_{\{1\}}(q_{\{1\}}) \end{pmatrix}$$

and so

$$P(q)_S = \frac{1}{3} \cdot \mu_{\{1,2\}}(q_{\{1,2\}}) + \frac{2}{3} \cdot \mu_{\{1\}}(q_1) \times \mu_{\{2\}}(q_2). \tag{4}$$

We now show that Eq. (1) holds for all $q, q' \in \mathcal{Q}$ with $q_S = q'_S$ for $|S| = 2$.

- If both q and q' are in the support of $\mathcal{G}$, then both $P(q)_S$ and $P(q')_S$ are both given by Eq. (3), and Eq. (1) follows respectively from the fact that $\tilde{P}$ is honest-referee non-signaling.
- If neither q nor q' are in the support of $\mathcal{G}$, then $P(q)_S$ and $P(q')_S$ are given by Eq. (4), and Eq. (1) follows from Condition 2.
- If exactly one of q and q' is in the support of $\mathcal{G}$, then Eq. (1) follows from Condition 1.

1.2.2 Games With More Than Three Players

We now return to the case that $\mathcal{G}$ is a general k-player game. In this case, we construct the non-signaling strategy P from $\tilde{P}$ iteratively, as follows. We will maintain the invariant that at the i^{th} step of our construction, we have two components:

- A collection of distributions $\mu^{(i)} = \{\mu_S^{(i)}\}_{|S| \leq i}$, where for each $S \subseteq [k]$ and $q \in \mathcal{Q}_S$, $\mu_S^{(i)}(q)$ is a distribution on $\mathcal{A}_S$.
- An honest-referee non-signaling strategy $\tilde{P}^{(i)}$.

We will ensure that these components satisfy the following consistency conditions.

Condition 3 (Non-Signaling)**.** *For any S and $q \in \mathcal{Q}_S$, the distribution $\mu_S^{(i)}(q)_T$ depends only on T and q_T, and not on S or the rest of q. Formally: for any $S, S' \subseteq [k]$ with $|S|, |S'| \leq i$, any $T \subseteq S \cap S'$, any $q \in \mathcal{Q}_S$ and $q' \in \mathcal{Q}_{S'}$, if $q_T = q'_T$ then $\mu_S^{(i)}(q)_T = \mu_{S'}^{(i)}(q')_T$.*

Condition 4 (Agreement)**.** *For every q in the support of $\mathcal{G}$ and every $S \subseteq [k]$ with $|S| \leq i$, the marginal distribution $\tilde{P}^{(i)}(q)_S$ is equal to $\mu_S^{(i)}(q_S)$.*

Condition 5. *$\tilde{P}^{(i)}$ is an extension of $\alpha_i \cdot \tilde{P}$ for some $\alpha_i > 0$ – i.e., for each q in the support of $\mathcal{G}$ and each $a \in \mathcal{A}$, we have $\Pr_{A \leftarrow \tilde{P}^{(i)}(q)}[A = a] \geq \alpha_i \cdot \Pr_{A \leftarrow \tilde{P}(q)}[A = a]$.*

To begin, we define $\mu_{\{i\}}^{(1)}(q) \overset{\mathsf{def}}{=} \tilde{P}(q')_{\{i\}}$ for any q' that satisfies $q'_i = q$, and we define $\tilde{P}^{(1)} = \tilde{P}$. After k steps, the resulting $\{\mu_S^{(k)}\}_{|S|=k}$ constitutes our desired non-signaling strategy P.

Symmetric Marginal Compositions of $\tilde{P}^{(i)}$. We will only ever construct $\tilde{P}^{(i+1)}$ and $\mu^{(i+1)}$ by applying a very specific type of transformation, which we call a symmetric marginal composition (SMC), to $\tilde{P}^{(i)}$ and $\mu^{(i)}$. This type of transformation generalizes what we did for three-player games in Section 1.2.1. There, P ($\tilde{P}^{(3)}$ in our current terminology) was constructed as a combination of two transformations applied to $\tilde{P}^{(2)}$ (which happened to be equal to $\tilde{P}$).

- The first transformation, which we denote by [3], generally takes as input some $\tilde{P}^{(i)}$ and produces $\tilde{P}^{(i+1)}$ that on input q, samples and outputs the 3-marginal $\tilde{P}^{(i)}(q)_{\{1,2,3\}}$ – which of course is all of $\tilde{P}^{(i)}(q)$. Thus [3] is the identity transformation.

- The second transformation, which we denote by $[1, 1, 1]$, generally takes as input $\tilde{P}^{(i)}$ and produces $\tilde{P}^{(i+1)}$ that on input q, samples answers from each of the 1-marginals of $\tilde{P}^{(i)}$ – that is, $\tilde{P}^{(i+1)}(q)$ independently samples $a_j \leftarrow \tilde{P}^{(i)}(q)_j$ for each $j \in \{1, 2, 3\}$ and outputs (a_1, a_2, a_3).

In Section 1.2.1, we in fact defined $\tilde{P}^{(3)}$ to select the first transformation of $\tilde{P}^{(2)}$ with probability 1/3, and to otherwise (with probability 2/3) select the second transformation of $\tilde{P}^{(2)}$. We call this a "mixed" transformation (in contrast to [3] and [1, 1, 1], which we call "pure" transformations), and denote it by the linear combination $\frac{1}{3} \cdot [3] + \frac{2}{3} \cdot [1, 1, 1]$.

In general, we represent a pure transformation by $\lambda = [\lambda_1, \ldots, \lambda_m]$ for any $\lambda_1 \geq \cdots \geq \lambda_m \geq 1$ with $\sum_j \lambda_j = k$. This transformation acts on a k-player honest-referee non-signaling strategy $\tilde{P}^{(i)}$ to produce another k-player honest-referee strategy $\tilde{P}^{(i+1)}$. This $\tilde{P}^{(i+1)}$, on any input q in the support of $\mathcal{G}$, acts as follows.

1. Sample m pairwise disjoint subsets $S_1, \ldots, S_m \subseteq [k]$ with $|S_j| = \lambda_j$ for each $j \in [m]$.
2. Independently sample $a_j \leftarrow \tilde{P}^{(i)}(q)_{S_j}$ for each $j \in [m]$.
3. Output the unique $a^* \in \mathcal{A}$ obtained by composing each of the a_j's. That is, the a^* for which $a^*_{S_j} = a_j$ for all $j \in [m]$.

Finally, we say that a linear combination $\mathbf{u} = \sum \mathbf{u}_\lambda \cdot \lambda$ (for $\mathbf{u}_\lambda \in \mathbb{R}_{\geq 0}$) is an SMC if it is a probability distribution over pure transformations – i.e., if $\sum \mathbf{u}_\lambda = 1$. It turns out (Lemma 5.11) that any SMC preserves honest-referee non-signaling. In general, we will define $\tilde{P}^{(i+1)}$ by applying an SMC $\mathbf{u}^{(i)}$ to $\tilde{P}^{(i)}$. For the moment, we leave $\mathbf{u}^{(i)}$ unspecified.

Symmetric Marginal Compositions of $\mu^{(i)}$. We now give a similar parameterization of our construction of $\mu^{(i+1)}$. First, for all S and all q in the S-marginal support of $\mathcal{G}$, $\mu_S^{(i+1)}(q)$ is determined by our choice of $\tilde{P}^{(i+1)}$ and Condition 4. As for defining $\mu_S^{(i+1)}(q)$ when q is *not* in the S-marginal support of $\mathcal{G}$, we first simplify our lives in two ways.

- First, we note that it suffices to define $\mu_S^{(i+1)}(q)$ for $|S| = i + 1$. Indeed, if these distributions satisfy the analogue of Condition 3, then for $|T| \leq i$, we can define $\mu_T^{(i+1)}(q)$ as $\big(\mu_S^{(i+1)}(q')\big)_T$ for any $S \supseteq T$ with $|S| = i + 1$, and any $q' \in \mathcal{Q}_S$ such that $q'_S = q$.
- Second, we simplify Condition 3. Recall this requires that for all S, S', all $T \subseteq S \cap S'$, and all $q \in \mathcal{Q}_S$, $q' \in \mathcal{Q}_{S'}$ such that $q_T = q'_T$, it holds that $\big(\mu_S^{(i+1)}(q)\big)_T = \mu_{S'}^{(i+1)}(q')\big)_T$. In fact this holds as long as it holds *when* $|T| = i$. Indeed, for smaller T we can deduce this via a chain of equalities similar to Eq. (2).

Now fix $|S| = i + 1$, and fix q not in the S-marginal support of $\mathcal{G}$. We will define $\mu_S^{(i+1)}(q)$ in terms of $\mu^{(i)}$ via a similar class of transformations to those that we used to define $\tilde{P}^{(i+1)}$. Specifically, suppose that $i \geq \lambda_1 \geq \cdots \geq \lambda_m \geq 1$ with $\sum_j \lambda_j = i + 1$. We then view $\lambda = [\lambda_1, \ldots, \lambda_m]$ as acting on $\mu^{(i)}$ to produce $\mu_S^{(i+1)}(q)$ as follows. Specifically, $\mu_S^{(i+1)}(q)$ is the distribution obtained from the following sampling process.

1. Sample m pairwise disjoint subsets $S_1, \ldots, S_m \subseteq S$ with $|S_j| = \lambda_j$ for each $j \in [m]$.
2. Independently sample $a_j \leftarrow \mu_{S_j}^{(i)}(q_{S_j})$ for each $j \in [m]$.
3. Output the unique $a^* \in \mathcal{A}_S$ obtained by composing each of the a_j's. That is, the a^* for which $a^*_{S_j} = a_j$ for all $j \in [m]$.

In general, we will consider a mixed SMC of $\mu^{(i)}$, given by a vector $\mathbf{v}^{(i)} = \sum \mathbf{v}_\lambda^{(i)} \cdot \lambda$. We will use the same vector $\mathbf{v}^{(i)}$ for every S and q for which $|S| = i + 1$ and q is not in the S-marginal support of $\mathcal{G}$.

Thus, we have parameterized our candidate construction of $\tilde{P}^{(i+1)}$ and $\mu^{(i+1)}$ by two SMCs $\mathbf{u}^{(i)}$ and $\mathbf{v}^{(i)}$. In this construction, $\tilde{P}^{(i+1)}$ and $\mu^{(i+1)}$ satisfy the "Agreement" property automatically. There are three remaining properties that must hold.

- $\tilde{P}^{(i+1)}$ must extend a constant multiple of $\tilde{P}^{(i)}$. This corresponds to $\mathbf{u}^{(i)}$ containing $[k]$ with positive probability.
- $\tilde{P}^{(i+1)}$ must be non-signaling. As noted earlier, this follows from the fact that all SMCs preserve non-signaling.
- $\mu^{(i+1)}$ must be non-signaling, despite the fact that $\mu_S^{(i+1)}(q)$ is defined via $\mathbf{u}^{(i)}$ when q is in the S-marginal support of $\mathcal{G}$, and via $\mathbf{v}^{(i)}$ otherwise.

The first and last conditions impose constraints on $\mathbf{u}^{(i)}$ and $\mathbf{v}^{(i)}$ that we show are always satisfiable (Section 5.3).

2 Preliminaries

2.1 Notation and Terminology

A probability distribution on a finite set Ω is a function $P : \Omega \to \mathbb{R}_{\geq 0}$ satisfying $\sum_{\omega \in \Omega} P(\omega) = 1$. We write $\Delta(\Omega)$ to denote the set of all probability distributions on Ω. A sub-probability distribution on Ω is a function $P : \Omega \to \mathbb{R}_{\geq 0}$ satisfying $\sum_{\omega \in \Omega} P(\omega) \leq 1$. We write $\Delta^{\leq 1}(\Omega)$ to denote the set of all sub-probability distributions on Ω.

- For any (sub-)probability distribution P, let $\mathrm{Supp}\,(P)$ denote the set of all $\omega \in \Omega$ for which $P(\omega) > 0$.
- For any (sub-)probability distribution P and $n \in \mathbb{N}$, let P^n denote the product (sub-)probability distribution $P^n : \Omega^n \to \mathbb{R}_{\geq 0}$ defined by $P^n(\omega^{(1)}, \ldots, \omega^{(n)}) = \prod_i P(\omega^{(i)})$.
- We will denote the total variational distance between two (sub-)probability distributions P and Q by $d_{\mathsf{TV}}(P, Q)$, which is defined as half of their ℓ_1-distance:

$$d_{\mathsf{TV}}(X, Y) \overset{\mathsf{def}}{=} \frac{1}{2} \sum_{z \in \mathrm{Supp}(X) \cup \mathrm{Supp}(Y)} \big| \Pr[X = z] - \Pr[Y = z] \big|.$$

  We write $P \approx_\epsilon Q$ if $d_{\mathsf{TV}}(P, Q) \leq \epsilon$.

- We now consider the case when Ω is a product of sets $\Omega_1 \times \cdots \times \Omega_k$ and S is a subset of $[k]$.
  - The restriction of Ω to S is $\Omega_S = \prod_{i \in S} \Omega_i$. For any $\omega \in \Omega$, the restriction of ω to S is $\omega_S = (\omega_i)_{i \in S}$.
  - For any (sub-)probability distribution P, the marginal (sub-)probability distribution $P_S : \Omega_S \to \mathbb{R}_{\geq 0}$ is defined by $P_S(\omega_S) = \sum_{\omega' \in \Omega : \omega'_S = \omega_S} P(\omega')$.

2.2 Multiplayer Games

Definition 2.1. *A k-player game is a tuple $(\mathcal{Q}, \mathcal{A}, \pi, W)$, where $\mathcal{Q} = \mathcal{Q}_1 \times \cdots \times \mathcal{Q}_k$ and $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_k$ are finite sets, $\pi : \mathcal{Q} \to \mathbb{R}_{\geq 0}$ is a probability distribution, and $W : \mathcal{Q} \times \mathcal{A} \to [0, 1]$ is a "winning probability" function.*

Definition 2.2. *Given a k-player game $\mathcal{G} = (\mathcal{Q}, \mathcal{A}, \pi, W)$ where $\mathcal{Q} = \mathcal{Q}_1 \times \cdots \times \mathcal{Q}_k$ and $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_k$, its n-fold parallel repetition is defined as the k-player game $\mathcal{G}^n \overset{\mathsf{def}}{=} (\mathcal{Q}', \mathcal{A}', \pi', W')$ where $\mathcal{Q}' \overset{\mathsf{def}}{=} \mathcal{Q}_1^n \times \cdots \times \mathcal{Q}_k^n$, $\mathcal{A}' \overset{\mathsf{def}}{=} \mathcal{A}_1^n \times \cdots \times \mathcal{A}_k^n$, $\pi'(q) \overset{\mathsf{def}}{=} \prod_{i=1}^n \pi(q^{(i)})$, and $W'(q, a) \overset{\mathsf{def}}{=} \prod_{i=1}^n W(q^{(i)}, a^{(i)})$.*

In the above we write elements $q \in \mathcal{Q}'$ as $\left(\{q_1^{(i)}\}_{i \in [n]}, \ldots, \{q_k^{(i)}\}_{i \in [n]}\right)$, we write q_j to denote $(q_j^{(1)}, \ldots, q_j^{(n)})$, and we write $q^{(i)}$ to denote $(q_1^{(i)}, \ldots, q_k^{(i)})$. Our notation for components of elements of $\mathcal{A}'$ is analogous.

Definition 2.3. *A strategy for a game $\mathcal{G} = (\mathcal{Q}, \mathcal{A}, \pi, W)$ is a function $P : \mathcal{Q} \to \Delta(\mathcal{A})$.*

Definition 2.4. *A strategy P for a k-player game $\mathcal{G} = (\mathcal{Q}, \mathcal{A}, \pi, W)$ is* local *if there are functions $\{P_i : \mathcal{Q}_i \to \mathcal{A}_i\}_{i \in [k]}$ such that for every q, $P(q)$ is the distribution outputting $(P_1(q_1), \dots, P_k(q_k))$ with probability 1.*

Definition 2.5. *A strategy P for a k-player game $\mathcal{G} = (\mathcal{Q}, \mathcal{A}, \pi, W)$ is* non-signaling *if for every subset $S \subseteq [k]$, there exists a function $\mathsf{Sim}_S : \mathcal{Q}_S \to \Delta(\mathcal{A}_S)$ where for all $q \in \mathcal{Q}$, it holds that $P(q)_S = \mathsf{Sim}_S(q_S)$.*

Definition 2.6 (Values of a Game)**.** *The* value *of a game $\mathcal{G} = (\mathcal{Q}, \mathcal{A}, \pi, W)$ with respect to a strategy P for $\mathcal{G}$, denoted by $v[P](\mathcal{G})$, is the expected value of $W(Q, A)$ in the probability space obtained by sampling $Q \leftarrow \pi$ and $A \leftarrow P(Q)$. The* classical value *of $\mathcal{G}$, denoted by $v(\mathcal{G})$ is the maximum value of $\mathcal{G}$ with respect to any local strategy. The* non-signaling value *of $\mathcal{G}$, denoted by $v_{\mathsf{ns}}(\mathcal{G})$, is the maximum value of $\mathcal{G}$ with respect to any non-signaling strategy.*

3 Generalizations of Non-signaling Strategies

[LW16] introduces the following generalization of non-signaling strategies, called sub-non-signaling strategies, where $P(q)$ is only required to be a sub-probability distribution.

Definition 3.1 ([LW16])**.** *A* sub-non-signaling *strategy P for a k-player game $\mathcal{G} = (\mathcal{Q}, \mathcal{A}, \pi, W)$ is a function $P : \mathcal{Q} \to \Delta^{\leq 1}(\mathcal{A})$ satisfying the following property: for every $S \subseteq [k]$, there exists a function $\mathsf{Sim}_S : \mathcal{Q}_S \to \Delta(\mathcal{A}_S)$ such that for every $q \in \mathcal{Q}$ and $a \in \mathcal{A}_S$, it holds that $P(q)_S(a) \leq \mathsf{Sim}_S(q_S)(a)$.*

Definition 3.2. *The* sub-non-signaling value *of $\mathcal{G}$, denoted by $v_{\mathsf{sns}}(\mathcal{G})$, is the maximum value of*

$$\sum_{q \in \mathcal{Q}} \pi(q) \sum_{a \in \mathcal{A}} P(q)(a) W(q, a)$$

attained by any sub-non-signaling strategy P.

In our work, it will be useful to consider the following generalization of non-signaling strategies where the non-signaling condition is only required to hold over queries in $\mathrm{Supp}\,(\pi)$. Such strategies appear to be non-signaling to an honest referee asking queries according to π.

Definition 3.3. *An* honest-referee non-signaling *strategy P for a k-player game $\mathcal{G} = (\mathcal{Q}, \mathcal{A}, \pi, W)$ is a function $P : \mathrm{Supp}\,(\pi) \to \Delta(\mathcal{A})$ satisfying the following property: for every $S \subseteq [k]$, there exists a function $\mathsf{Sim}_S : \mathrm{Supp}\,(\pi_S) \to \Delta(\mathcal{A}_S)$ such that for every $q \in \mathrm{Supp}\,(\pi)$, it holds that $P(q)_S = \mathsf{Sim}_S(q_S)$.*

Remark 3.4. *We can view any non-signaling strategy P as a complete set of marginals $\{P_S : \mathcal{Q}_S \to \Delta(\mathcal{A}_S)\}_{S \subseteq [k]}$: for any $|S| \subset [k]$ and $q \in \mathcal{Q}_S$, $P_S(q) \overset{\mathsf{def}}{=} P(q')_S$ for any $q' \in \mathcal{Q}$ such that $q'_S = q$. This is well-defined (i.e., does not depend on the choice of q') because P is non-signaling, so we have $P_S = \mathsf{Sim}_S$. Similarly, we can view any honest-referee non-signaling strategy P as a complete set of marginals $\{P_S : \mathrm{Supp}\,(\pi_S) \to \Delta(\mathcal{A}_S)\}_{S \subseteq [k]}$ where $P_S = \mathsf{Sim}_S$.*

The following proposition is analogous to the fact that $v_{\mathsf{ns}}(\mathcal{G}^n) \geq v_{\mathsf{ns}}(\mathcal{G})^n$ and implies that if $v_{\mathsf{hr\text{-}ns}}(\mathcal{G}) = 1$, then also $v_{\mathsf{hr\text{-}ns}}(\mathcal{G}^n) = 1$ for every n.

Proposition 3.5. *For any game $\mathcal{G}$ and any number of repetitions n, it holds that $v_{\mathsf{hr\text{-}ns}}(\mathcal{G}^n) \geq v_{\mathsf{hr\text{-}ns}}(\mathcal{G})^n$.*

Proof. Let $\mathcal{G} = (\mathcal{Q}, \mathcal{A}, \pi, W)$ be a k-player game, and let $P : \mathrm{Supp}\,(\pi) \to \Delta(\mathcal{A})$ be an honest-referee non-signaling strategy for $\mathcal{G}$. Let $\mathcal{G}^n = (\mathcal{Q}^n, \mathcal{A}^n, \pi^n, W^n)$ denote the n-fold parallel repetition of $\mathcal{G}$.

Define

$$P^n : \mathrm{Supp}\,(\pi^n) \to \Delta(\mathcal{A}^n)$$

$$P^n(q^{(1)}, \dots, q^{(n)})(a^{(1)}, \dots, a^{(n)}) = \prod_{i=1}^{n} P(q_1^{(i)}, \dots, q_k^{(i)})(a_1^{(i)}, \dots, a_k^{(i)}),$$

which is well-defined because $(q^{(1)}, \dots, q^{(n)}) \in \mathrm{Supp}\,(\pi^n)$ iff for all $i \in [n]$, $(q_1^{(i)}, \dots, q_k^{(i)}) \in \mathrm{Supp}\,(\pi)$. It is easy to verify that P^n is an honest-referee non-signaling strategy for $\mathcal{G}^n$ since for any $(q^{(1)}, \dots, q^{(n)}) \in \mathrm{Supp}\,(\pi^n)$, we can define $\mathsf{Sim}_S(q_S^{(1)}, \dots, q_S^{(n)}) = \prod_i \mathsf{Sim}_S(q_S^{(i)})$. Also, it is clear that $v[P^n](\mathcal{G}^n) = v[P](\mathcal{G})^n$. □

For several of our corollaries, it will be useful to relate the sub-non-signaling value and honest-referee non-signaling value of a game $\mathcal{G}$. For this, we have the following proposition.

Proposition 3.6. *For any game $\mathcal{G}$, $v_{\mathsf{sns}}(\mathcal{G}) = 1$ iff $v_{\mathsf{hr\text{-}ns}}(\mathcal{G}) = 1$.*

Proof. Let $\mathcal{G} = (\mathcal{Q}, \mathcal{A}, \pi, W)$ be any k-player game.

> **Claim 3.7.** *If $v_{\mathsf{sns}}(\mathcal{G}) = 1$ then $v_{\mathsf{hr\text{-}ns}}(\mathcal{G}) = 1$.*
>
> *Proof.* Suppose that $P : \mathcal{Q} \to \Delta^{\leq 1}(\mathcal{A})$ is a sub-non-signaling strategy for $\mathcal{G}$ such that $v[P](\mathcal{G}) = 1$. Since $v[P](\mathcal{G}) = 1$, it must hold for all $q \in \mathrm{Supp}\,(\pi)$ that $P(q)$ is a probability distribution, i.e. $\sum_{a \in \mathcal{A}} P(q)(a) = 1$. Therefore for each $S \subseteq [k]$, $P(q)_S$ is a probability distribution.
>
> Since P is sub-non-signaling, there exists $\mathsf{Sim}_S : \mathcal{Q}_S \to \Delta(\mathcal{A}_S)$ such that for all $q \in \mathcal{Q}$ and all $a \in \mathcal{A}_S$, it holds that $P(q)_S(a) \leq \mathsf{Sim}_S(q_S)(a)$. Since $P(q)_S$ is a probability distribution, $\mathsf{Sim}_S(q_S)$ must also be a probability distribution. Therefore, for any $q \in \mathrm{Supp}\,(\pi)$ and any $S \subseteq [k]$, it holds that $P(q)_S = \mathsf{Sim}_S(q_S)$ so P is in fact honest-referee non-signaling. □
>
> **Claim 3.8.** *If $v_{\mathsf{hr\text{-}ns}}(\mathcal{G}) = 1$ then $v_{\mathsf{sns}}(\mathcal{G}) = 1$.*
>
> *Proof.* Suppose that $\tilde{P} : \mathrm{Supp}\,(\pi) \to \Delta(\mathcal{A})$ is an honest-referee non-signaling strategy with $v[\tilde{P}](\mathcal{G}) = 1$. Extend $\tilde{P}$ to $P : \mathcal{Q} \to \Delta^{\leq 1}(\mathcal{A})$ by defining $P(q)$ to be the zero sub-probability distribution for all $q \notin \mathrm{Supp}\,(\pi)$. Clearly $v[P](\mathcal{G}) = v[\tilde{P}](\mathcal{G}) = 1$.
>
> It remains to establish that P is sub-non-signaling: For each $S \subseteq [k]$, define Sim_S so that for any $q \in \mathrm{Supp}\,(\pi)$, $\mathsf{Sim}_S(q_S) = \tilde{P}(q)_S$ – this is possible because $\tilde{P}$ is honest-referee non-signaling. For $q' \notin \mathrm{Supp}\,(\pi_S)$, define $\mathsf{Sim}_S(q')$ arbitrarily. Now clearly for any $q \in \mathcal{Q}$ and any $a \in \mathcal{A}_S$, it holds that $P(q)_S(a) \leq \mathsf{Sim}_S(q_S)(a)$. □

Proposition 3.6 follows directly from Claims 3.7 and 3.8. □

4 Main Lemma and Implications

We prove the theorems outlined in Section 1.1 assuming our main lemma, whose proof is deferred to Section 5. The main theorems are stated here with greater generality in the context of honest-referee non-signaling strategies.

Lemma 1 (Non-signaling Value Lower Bound)**.** *For every k, there exists a constant $\alpha_k > 0$ such that for any k-player game $\mathcal{G}$, it holds that $v_{\mathsf{ns}}(\mathcal{G}) \geq \alpha_k \cdot v_{\mathsf{hr\text{-}ns}}(\mathcal{G})$.*

From here on, we write α_k to refer to the maximum value for which Lemma 1 holds. We have not attempted to optimize α_k, but a loose accounting of our proof shows that $\alpha_k \geq 2^{-O(k^2)}$. Conversely, Theorem 4 below shows that $\alpha_k \leq 2^{-\Omega(k)}$.

To show that there are counterexamples to non-signaling parallel repetition, we show that there exist games whose sub-non-signaling values are 1 but whose non-signaling values are less than 1.

Proposition 4.1. *For every $k \geq 3$, there exists a k-player game $\mathcal{G}$ for which $v_{\mathsf{hr\text{-}ns}}(\mathcal{G}) = 1$ (i.e. $v_{\mathsf{sns}}(\mathcal{G}) = 1$) and $v_{\mathsf{ns}}(\mathcal{G}) < 1$.*

One example of such a game is the "anti-correlation game". The three-player version of this game was described in [FRV16, LW16] as an example of a game whose non-signaling value is less than 1, but whose sub-non-signaling value is 1.

Definition 4.2 (Anti-Correlation Game)**.** *The k-*player anti-correlation game *$\tilde{\mathcal{G}}_k$ is the game $(\tilde{\mathcal{Q}}_k, \tilde{\mathcal{A}}_k, \tilde{\pi}_k, \tilde{W}_k)$ where $\tilde{\mathcal{Q}}_k = \{0,1\}^k$, $\tilde{\mathcal{A}}_k = [k-1]^k$, $\tilde{\pi}_k$ is the uniform distribution on k-bit strings of Hamming weight $k-1$, and*

$$\tilde{W}_k(q,a) = \begin{cases} 1 & \textit{if for all } i \neq j \in [k] \textit{ such that } q_i = q_j = 1, \textit{ it holds that } a_i \neq a_j \\ 0 & \textit{otherwise.} \end{cases}$$

Proof of Proposition 4.1. Let $\mathcal{G}$ be the k-player anti-correlation game $\tilde{\mathcal{G}}_k$. For $1 \leq \ell \leq k$, let $h_\ell \in \{0,1\}^k$ denote the string of Hamming weight $k-1$ with 0 in the ℓth coordinate.

Define $P : \mathrm{Supp}\,(\tilde{\pi}_k) \to \Delta(\tilde{\mathcal{A}}_k)$ as follows: $P(h_\ell)$ chooses a random element of $[k-1]$ for a_ℓ and a random permutation of $[k-1]$ for $(a_i)_{i \neq \ell}$. This is an honest-referee non-signaling strategy: for any $h_\ell, h_m \in \mathrm{Supp}\,(\tilde{\pi}_k)$, the marginals $P(h_\ell)_S$ and $P(h_m)_S$, where $S = [k] \setminus \{\ell, m\}$, are both equivalent to choosing a $(k-2)$-sized subset of $[k-1]$ and a random permutation of these elements. For distinct $i, j \neq \ell$, we have $a_i \neq a_j$ so $v_{\textsf{hr-ns}}[P](\tilde{\mathcal{G}}_k) = 1$. Thus $v_{\textsf{hr-ns}}(\tilde{\mathcal{G}}_k) = 1$ (by Proposition 3.6, this is equivalent to $v_{\textsf{sns}}(\tilde{\mathcal{G}}_k) = 1$).

For any non-signaling $P' : \tilde{\mathcal{Q}}_k \to \Delta(\tilde{\mathcal{A}}_k)$ and any $h_\ell \in \mathrm{Supp}\,(\tilde{\pi}_k)$, $P'(h_\ell)_{[k]\setminus\{\ell\}} = P'(1^k)_{[k]\setminus\{\ell\}}$ by non-signaling. Thus $\Pr_{\ell, P'(h_\ell)}[\tilde{W}_k(h_\ell, P'(h_\ell)) = 1] = \Pr_{\ell, P'(1^k)}[\tilde{W}_k(h_\ell, P'(1^k)) = 1]$. This equals the probability that $P'(1^k)_{[k]\setminus\{\ell\}}$ is a permutation of $[k-1]$. By the Pigeonhole principle, every element in $\mathrm{Supp}\,\big(P'(1^k)\big)$ has an element of $[k-1]$ appearing in at least two indices so this probability is at most $2/k$. □

Theorem 1 (Parallel Repetition Counterexample)**.** *For every $k \geq 3$, there exists a k-player game $\mathcal{G}$ such that $v_{\textsf{ns}}(\mathcal{G}) < 1$ and $v_{\textsf{ns}}(\mathcal{G}^n) \geq \alpha_k$ for all $n \geq 1$, where α_k is the constant in Lemma 1.*

Proof. This follows immediately from Proposition 4.1 and Lemma 1: for the anti-correlation game $\tilde{\mathcal{G}}_k$, it holds that $v_{\textsf{hr-ns}}(\tilde{\mathcal{G}}_k) = 1 \implies v_{\textsf{hr-ns}}(\tilde{\mathcal{G}}_k^n) = 1 \implies v_{\textsf{ns}}(\tilde{\mathcal{G}}_k^n) \geq \alpha_k$. □

Theorem 2 (Parallel Repetition Dichotomy)**.** *For every game $\mathcal{G}$, either $v_{\textsf{ns}}(\mathcal{G}^n) \geq \Omega(1)$ or $v_{\textsf{ns}}(\mathcal{G}^n) \leq \exp(-\Omega(n))$. The former occurs when $v_{\textsf{sns}}(\mathcal{G}) = 1$ and the latter occurs when $v_{\textsf{sns}}(\mathcal{G}) < 1$.*

Proof. If $v_{\textsf{sns}}(\mathcal{G}) < 1$, then by [LW16, Theorem 4], we have $v_{\textsf{ns}}(\mathcal{G}^n) \leq v_{\textsf{sns}}(\mathcal{G}^n) \leq e^{-\Omega(n)}$.

If $v_{\textsf{sns}}(\mathcal{G}) = 1$, then by Proposition 3.6 $v_{\textsf{hr-ns}}(\mathcal{G}) = 1$. Then by Proposition 3.5, $v_{\textsf{hr-ns}}(\mathcal{G}^n) \geq v_{\textsf{hr-ns}}(\mathcal{G})^n = 1$, so $v_{\textsf{hr-ns}}(\mathcal{G}^n) = 1$ for all $n \geq 1$. By Lemma 1, this implies that $v_{\textsf{ns}}(\mathcal{G}^n) \geq \alpha_k = \Omega(1)$. □

Theorem 3 (Parallel Repetition Magic Value)**.** *For every k-player game $\mathcal{G}$, if $v_{\textsf{ns}}(\mathcal{G}) < \alpha_k$, then $v_{\textsf{ns}}(\mathcal{G}^n) \leq \exp(-\Omega(n))$ where α_k is the constant in Lemma 1.*

Proof. If $v_{\textsf{ns}}(\mathcal{G}) < \alpha_k$, then by Lemma 1 and Proposition 3.6, we have $v_{\textsf{sns}}(\mathcal{G}) < 1$. Then by [LW16], it holds that $v_{\textsf{ns}}(\mathcal{G}^n) \leq v_{\textsf{sns}}(\mathcal{G}^n) \leq \exp(-\Omega(n))$. □

Upper Bounding the Magic Value Given Theorem 3, it is natural to ask whether α_k is optimal: For each k, what is the largest value β_k such that for every k-player game $\mathcal{G}$, $v_{\textsf{ns}}(\mathcal{G}) < \beta_k$ implies $v_{\textsf{ns}}(\mathcal{G}^n) \leq \exp(-\Omega(n))$? Theorem 1 shows that $\beta_k < 1$, and Theorem 3 shows that $\beta_k \geq \alpha_k$, which our proof of Lemma 1 shows to be at least 2^{-k^2}. This is a large gap, but Theorem 4 shows that $\beta_k \leq \exp(-\Omega(k))$. The proof is in Section 6: the main tool is a "distributed" form of repetition that does exponentially decrease the non-signaling value of a game, but linearly increases the number of players.

Theorem 4 (Strong Parallel Repetition Counterexample)**.** *There is a sequence of games $\{\mathcal{G}_k\}_{k\geq 3}$, with $\mathcal{G}_k$ a k-player game, such that $v_{\textsf{ns}}(\mathcal{G}_k) \leq \exp(-\Omega(k))$ and $v_{\textsf{ns}}(\mathcal{G}_k^n) \geq \alpha_k$ for all $n \geq 1$.*

5 Proof of Main Lemma

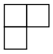
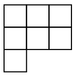
Overview. For an arbitrary game $\mathcal{G}$, we show how any honest-referee non-signaling strategy $\tilde{P}$ can be scaled down and extended to a full non-signaling strategy. The strategies we construct have the following format. On any given query $q = (q_1, \ldots, q_k)$, we first randomly partition $[k] = \cup_i S_i$, according to a carefully chosen distribution of partitions, while ensuring that q_{S_i} is in the S_i-marginal support of $\mathcal{G}$. For each i, we then independently sample $a_{S_i} \leftarrow \tilde{P}(q')$ for some q' with $q'_{S_i} = q_{S_i}$.[2]

In Section 5.1, we present the basic notation and formalism that we use to refer to partitions and distributions of partitions. In Section 5.2, we introduce a notion of a "d-ary marginal strategy" that is incomparable to the notion of an honest-referee non-signaling strategy, but coincides with the notion of a full strategy for $d = k$. We give a construction of a $(d+1)$-ary marginal strategy that is parameterized by a distribution of partitions, as well as a d-ary marginal strategy and an honest-referee strategy. In Section 5.3, we concoct nice distributions of partitions that, when applied iteratively to the constructions of Section 5.2, yield a non-signaling strategy that extends the original honest-referee non-signaling strategy. We put it all together in Section 5.4.

5.1 Technical Tools

We begin with some necessary preliminaries. Throughout this section, let $\mathcal{G} = (\mathcal{Q}, \mathcal{A}, \pi, W)$ be any fixed k-player game.

Definition 5.1 (Partitions of Natural Numbers)**.** *For $n \in \mathbb{N}$, a* partition *λ* of *n is a weakly decreasing sequence $[\lambda_1, \ldots, \lambda_m]$ of non-negative integers such that $\sum_i \lambda_i = n$. Occasionally it is convenient to think of λ as an* infinite *sequence, in which case we set $\lambda_i = 0$ for all $i > m$. We denote by Λ_n the set of all partitions of n and we denote by $\Lambda_n^{\leq d}$ the set of all partitions of n whose largest component is at most d.*

Young diagrams graphically represent partitions – a partition $\lambda = [\lambda_1, \ldots, \lambda_m]$ of n is represented by an array of m left-justified rows of unit squares, where the i^{th} row from top to bottom contains λ_i squares. For example, the partition $[2, 1]$ of 3 is represented by  and the partition $[3, 3, 1]$ of 7 is represented by . Many natural combinatorial operations on partitions have simple interpretations as operations on the corresponding Young diagrams.

Definition 5.2 (Vector Space of Partitions)**.** *We denote by $\mathbb{R}(\Lambda_n)$ the space of all formal linear combinations of partitions of n with coefficients in $\mathbb{R}$. We view $\mathbb{R}(\Lambda_n)$ as a $|\Lambda_n|$-dimensional vector space over $\mathbb{R}$ in the natural way.*

We denote by $\mathbb{R}_{\geq 0}(\Lambda_n)$ the vectors in $\mathbb{R}(\Lambda_n)$ with nonnegative coefficients. We denote by $\Delta(\Lambda_n)$ the vectors in $\mathbb{R}(\Lambda_n)$ with nonnegative coefficients summing to 1. These vectors can be thought of as probability distributions over partitions of n.

Definition 5.3 (Reduction Operator)**.** *For each $n \in \mathbb{N}$, we define the* reduction operator

$$\mathcal{R} : \Lambda_n \to \Delta(\Lambda_{n-1}) \subset \mathbb{R}(\Lambda_{n-1}),$$

which maps a partition λ of n to a distribution $\mathcal{R}(\lambda)$ on partitions of $n-1$ as follows. For $\lambda \in \Lambda_n$ and $\lambda' \in \Lambda_{n-1}$, if for some $i \in \mathbb{N}$ it holds that

$$\lambda' = [\lambda_1, \ldots, \lambda_{i-1}, \lambda_i - 1, \lambda_{i+1}, \ldots], \tag{5}$$

then the coefficient of λ' in $\mathcal{R}(\lambda)$ is $\mathcal{R}(\lambda)_{\lambda'} \overset{\mathsf{def}}{=} \frac{1}{n} \sum_{j : \lambda_j = \lambda_i} \lambda_j$. If Eq. (5) does not hold for any i, then $\mathcal{R}(\lambda)_{\lambda'} \overset{\mathsf{def}}{=} 0$. We extend the reduction operator to $\mathcal{R} : \mathbb{R}(\Lambda_n) \to \mathbb{R}(\Lambda_{n-1})$ by linearity, i.e. for $\mathbf{v} \in \mathbb{R}(\Lambda_n)$, we define $\mathcal{R}\big(\sum_{\lambda \in \Lambda_n} \mathbf{v}_\lambda \cdot \lambda\big) \overset{\mathsf{def}}{=} \sum_{\lambda \in \Lambda_n} \mathbf{v}_\lambda \cdot \mathcal{R}(\lambda)$.

When λ is a partition of n, the reduction operator has a simple interpretation as an operation on the Young diagram for λ. Sample a random square of the Young diagram for λ and delete it. Then rearrange the rows to get a valid Young diagram, i.e. rearrange the rows so they are weakly decreasing in length. As defined, $\mathcal{R}(\lambda)_{\lambda'}$ equals the probability of obtaining λ' after this operation, so $\mathcal{R}(\lambda)$ is indeed a distribution on partitions of $n-1$.

This operation is illustrated in Figs. 1 and 2.

[2] This is well-defined because q_{S_i} is in the S_i-marginal support of $\mathcal{G}$, and because $\tilde{P}$ is honest-referee non-signaling.

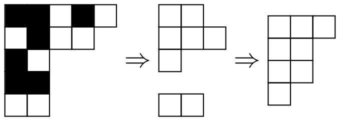


Figure 1: To remove specified squares from a Young diagram (here, black squares), first left-justify all remaining squares, then re-arrange rows in order of decreasing width.

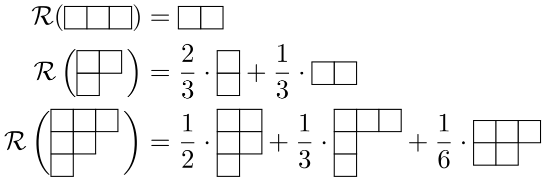


Figure 2: Examples of how $\mathcal{R}$ operates on Young diagrams

Definition 5.4 (Partitions of Sets). *We say that a tuple of pairwise disjoint sets $(S_i)_i = (S_1, \dots, S_m)$ is a* partition of S *if $S = \cup_i S_i$ and $|S_1| \geq \dots \geq |S_m|$. We denote by $\Lambda(S)$ the set of partitions of S, and we denote by $\Lambda^{\leq d}(S)$ the set of partitions of S where the largest component has size at most d.*

We say that $(S_1, \dots, S_m)$ is a partition of S with shape λ *if $\lambda = [|S_1|, \dots, |S_m|]$. For any partition λ and any set S with $|S| = |\lambda|$, we write $\lambda(S)$ to denote the set of partitions of S with shape λ.*

If $\mathbf{u} \in \Delta(\Lambda_d)$ is a distribution of partitions of d, then we write $\mathbf{u}(S) \in \Delta(\Lambda(S))$ to denote the distribution of partitions of S obtained by first sampling a partition λ from $\mathbf{u}$, and then choosing a uniformly random partition of S from $\lambda(S)$.

5.2 Marginal Strategies and Symmetric Marginal Compositions (SMCs)

We now describe a notion of a "marginal" non-signaling strategy for a game that will be useful as an intermediate step in constructing (full) non-signaling strategies.

Definition 5.5 (Marginal Non-Signaling Strategies). *A d-ary* marginal non-signaling strategy *for $\mathcal{G}$ is a collection of functions $\mu = \big\{\mu_S : \mathcal{Q}_S \to \Delta(\mathcal{A}_S)\big\}_{S \subseteq [k], |S| = d}$ such that for all $S, S' \subseteq [k]$ with $|S| = |S'| = d$ and for all $q \in \mathcal{Q}_S, q' \in \mathcal{Q}_{S'}$, it holds that $\mu_S(q)_T = \mu_{S'}(q')_T$ where $T \subseteq S \cap S'$ is the largest set such that $q_T = q'_T$.*

This notion coincides with the notion of non-signaling strategies when $d = k$.

Remark 5.6. *As with strategies and honest-referee strategies, any d-ary marginal non-signaling strategy $\mu = \{\mu_S : \mathcal{Q}_S \to \Delta(\mathcal{A}_S)\}_{S \subseteq [k], |S| = d}$ can be extended to a larger set of marginals $\{\mu_S : \mathcal{Q}_S \to \Delta(\mathcal{A}_S)\}_{S \subseteq [k], |S| \leq d}$ such that for all $S, S' \subseteq [k]$ with $|S|, |S'| \leq d$ and for all $q \in \mathcal{Q}_S, q' \in \mathcal{Q}_{S'}$, it holds that $\mu_S(q)_T = \mu_{S'}(q')_T$ where $T \subseteq S \cap S'$ is the largest set for which $q_T = q'_T$.*

Specifically, for any $|S| < d$ and $q \in \mathcal{Q}_S$, one defines $\mu_S(q) \overset{\mathsf{def}}{=} \mu_T(q')_S$ for any $T \supset S$ and $q' \in \mathcal{Q}_T$ such that $|T| = d$ and $q'_S = q$. This is well-defined (i.e., does not depend on the choice of T and q') because μ is (marginal) non-signaling.

In the strategies we construct, every answer distribution is obtained by composing, or "gluing together" the marginal answer distributions of another strategy. The precise way in which the original strategy's joint distribution is viewed as a collection of different marginal distributions to be glued together is specified by a partition of the relevant players.

Definition 5.7 (SMCs of Marginal Strategies)**.** *For any $S \subseteq [k]$, any d-ary marginal non-signaling strategy $\mu = \{\mu_{S'} : \mathcal{Q}_{S'} \to \Delta(\mathcal{A}_{S'})\}_{S' \subseteq [k], |S'|=d}$, and any partition $(S_1, \dots, S_m) \in \Lambda^{\leq d}(S)$, we define the $(S_1, \dots, S_m)$-*symmetric marginal composition *of μ for S as follows:*

$$\mu_{(S_1,\dots,S_m)} : \mathcal{Q}_S \to \Delta(\mathcal{A}_S)$$

$$\mu_{(S_1,\dots,S_m)}(q)(a) \overset{\mathsf{def}}{=} \prod_{i=1}^{m} \mu_{S_i}(q_{S_i})(a_{S_i}).$$

*More generally, for any $\mathbf{u} \in \Delta(\Lambda^{\leq d}(S))$, we define the $\mathbf{u}$-*symmetric marginal composition *of μ for S as follows:*

$$\mu_{\mathbf{u}} \overset{\mathsf{def}}{=} \sum_{(S_1,\dots,S_m) \in \mathrm{Supp}(\mathbf{u})} \mathbf{u}_{(S_1,\dots,S_m)} \cdot \mu_{(S_1,\dots,S_m)}.$$

Note that $\sum_{a \in \mathcal{A}_S} \mu_{(S_1,\dots,S_m)}(q)(a) = \prod_{i=1}^{m} \left(\sum_{a_{S_i} \in \mathcal{A}_{S_i}} \mu_{S_i}(q_{S_i})(a_{S_i}) \right) = 1$ so $\mu_{(S_1,\dots,S_m)}(q)$ is indeed a distribution over $\mathcal{A}_S$.

We also define an analogous operation for honest-referee non-signaling strategies.

Definition 5.8 (SMCs of Honest-referee Strategies)**.** *For any $S \subseteq [k]$, any honest-referee non-signaling strategy $P : \mathrm{Supp}\,(\pi) \to \Delta(\mathcal{A})$, and any partition $(S_1, \dots, S_m) \in \Lambda(S)$, we define the $(S_1, \dots, S_m)$-*symmetric marginal composition *of P for S as follows:*

$$P_{(S_1,\dots,S_m)} : \mathrm{Supp}\,(\pi_S) \to \Delta(\mathcal{A}_S)$$

$$P_{(S_1,\dots,S_m)}(q)(a) \overset{\mathsf{def}}{=} \prod_{i=1}^{m} P_{S_i}(q_{S_i})(a_{S_i}).$$

*More generally, for any $\mathbf{u} \in \Delta(\Lambda(S))$, we define the $\mathbf{u}$-*symmetric marginal composition *of P for S as follows:*

$$P_{\mathbf{u}} \overset{\mathsf{def}}{=} \sum_{(S_1,\dots,S_m) \in \mathrm{Supp}(\mathbf{u})} \mathbf{u}_{(S_1,\dots,S_m)} \cdot P_{(S_1,\dots,S_m)}.$$

We now characterize the marginal distributions of symmetric marginal compositions, in particular relating them to the reduction operator described in Definition 5.3.

Lemma 5.9. *For any d-ary marginal non-signaling strategy μ, any set $S \subseteq [k]$, any $\mathbf{u} \in \Delta(\Lambda^{\leq d}_{|S|})$, any $q \in \mathcal{Q}_S$, and any $T \subseteq S$ it holds that $\left(\mu_{\mathbf{u}(S)}(q)\right)_T = \mu_{(\mathcal{R}^{|S|-|T|}(\mathbf{u}))(T)}(q_T)$.*

Proof. By linearity, we may assume without loss of generality that $\mathbf{u}$ is just a single partition $\lambda \in \Lambda^{\leq d}_{|S|}$. Then, $\mu_{\lambda(S)}(q)$ is by definition the following distribution over $a' \in \mathcal{A}_S$:

$$\mu_{\lambda(S)}(q)(a') = \frac{1}{|\lambda(S)|} \cdot \sum_{(S_i)_i \in \lambda(S)} \prod_i \mu_{S_i}(q_{S_i})(a'_{S_i}).$$

Because the sets in $(S_i)_i$ are pairwise disjoint, the T-marginal distribution $\mu_{\lambda(S)}(q)_T$ over $a \in \mathcal{A}_T$ is:

$$\mu_{\lambda(S)}(q)_T(a) = \frac{1}{|\lambda(S)|} \cdot \sum_{(S_i)_i \in \lambda(S)} \prod_i \left(\mu_{S_i}(q_{S_i})\right)_T (a_{S_i \cap T}).$$

Since μ is d-ary marginal non-signaling, this can be written as

$$\mu_{\lambda(S)}(q)_T(a) = \frac{1}{|\lambda(S)|} \cdot \sum_{(S_i)_i \in \lambda(S)} \prod_i \mu_{S_i \cap T}(q_{S_i \cap T})(a_{S_i \cap T}).$$

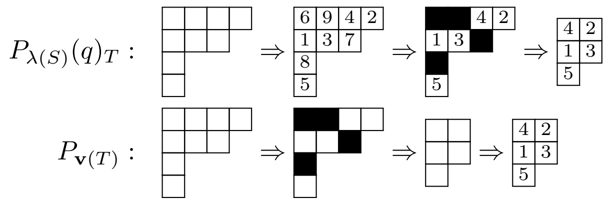


Figure 3: Illustration of Claim 5.10 in the case $S = [9]$, $\lambda = [4, 3, 1, 1]$ and $T = [5]$

Writing $\mathbf{v} = \sum_{\lambda' \in \Lambda_{|T|}} \mathbf{v}_{\lambda'} \cdot \lambda'$ as shorthand for $\mathcal{R}^{|S|-|T|}(\lambda)$, by definition $\mu_{\mathbf{v}(T)}(q_T)$ is the following distribution over $a \in \mathcal{A}_T$:

$$\mu_{\mathbf{v}(T)}(q_T)(a) = \sum_{\lambda' \in \Lambda_{|T|}} \mathbf{v}_{\lambda'} \cdot \frac{1}{|\lambda'(T)|} \cdot \sum_{(T_i)_i \in \lambda'(T)} \prod_i \mu_{T_i}(q_{T_i})(a_{T_i}).$$

Therefore, $\mu_{\lambda(S)}(q)_T$ and $\mu_{\mathbf{v}(T)}(q_T)$ can each be viewed as distributions on $\mathcal{A}_T$ that take the following form:

1. Sample a partition $(T_i)_i$ of T as follows:
   (a) In the case of $\mu_{\lambda(S)}(q)_T$, we first sample $(S_i)_i \leftarrow \lambda(S)$, and then set $T_i = S_i \cap T$.
   (b) In the case of $\mu_{\mathbf{v}(T)}(q_T)$, we directly sample $(T_i)_i \leftarrow \mathcal{R}^{|S|-|T|}(\lambda)(T)$.
2. Independently sample $a_{T_i} \leftarrow \mu_{T_i}(q_{T_i})$ for each i.
3. Output the unique $a \in \mathcal{A}_T$ that is consistent with each of the chosen values a_{T_i}.

It thus suffices to prove the following claim.

Claim 5.10. *The distributions on $(T_i)_i$ induced by steps 1a and 1b are identical.*

Proof. The easiest way to see this is through an equivalent reformulation of step 1a. Namely, we think of obtaining $(S_i)_i \leftarrow \lambda(S)$ by filling in the squares of the Young diagram for λ with the elements of S in a uniformly random order, and defining S_i to be the set of elements in the i^{th} row. Then, it is easy to see that (1) the elements of $S \setminus T$ are uniformly distributed (without replacement) in the resulting Young tableau, and (2) conditioned on their positions, the elements of T are uniformly distributed in the remaining positions.

Thus, we can equivalently think of step 1a as first obtaining a partition of $|T|$ by choosing and removing $|S| - |T|$ random squares from the Young diagram for λ, and then obtaining a partition of T by filling in the squares of the resulting Young diagram with the elements of T. But this is exactly equivalent to step 1b. The equivalence is illustrated in Fig. 3. □

This concludes the proof of Lemma 5.9. □

Lemma 5.11. *For any honest-referee non-signaling strategy $P : \mathrm{Supp}(\pi) \to \Delta(\mathcal{A})$, any set $S \subseteq [k]$, any $\mathbf{u} \in \Delta(\Lambda_{|S|})$, any $q \in \mathrm{Supp}(\pi_S)$, and any $T \subseteq S$, it holds that $\left(P_{\mathbf{u}(S)}(q)\right)_T = P_{(\mathcal{R}^{|S|-|T|}(\mathbf{u}))(T)}(q_T)$.*

Proof. The proof of Lemma 5.9 applies to any $q \in \mathcal{Q}$ where $\mu_T(q_T)$ is defined for all $T \subseteq S$. Although P is only defined over $\mathrm{Supp}(\pi)$ and hence is not a d-ary marginal non-signaling strategy, the marginal distribution $P_T(q_T)$ is defined for all $q \in \mathrm{Supp}(\pi_S)$ and $T \subseteq S$. Thus the proof of Lemma 5.11 follows identically to that of Lemma 5.9. □

5.3 $\mathcal{R}$ Identities

In this section we establish identities involving partitions and the $\mathcal{R}$ operator that will prove useful in our proof of the main lemma. In particular, in Section 5.4 these identities will be crucial for obtaining a player strategy that is both non-signaling and also non-trivially related to an arbitrary specified *honest-referee* non-signaling strategy.

For any partition $\lambda = [\lambda_1, \ldots, \lambda_m]$ of d and $k \geq d$, let $\Gamma_k(\lambda)$ denote the partition λ' obtained by extending λ to a partition of k by adding $k-d$ single 1's; that is,

$$\lambda'_i = \begin{cases} \lambda_i & \text{if } 1 \leq i \leq m \\ 1 & \text{if } m < i \leq m+k-d \\ 0 & \text{otherwise.} \end{cases}$$

We will find it convenient to impose the standard lexicographical ordering on partitions, also known as the dictionary ordering.

Definition 5.12 (Lexicographical Ordering)**.** *Under the* lexicographical ordering*, also known as the* dictionary ordering*, we say for two partitions λ and λ' that λ* is larger than *λ', denoted $\lambda \succ \lambda'$, if there exists $j \in \mathbb{N}$ such that $\lambda_i = \lambda'_i$ for all $i \in \{1, \ldots, j-1\}$, and $\lambda_j > \lambda'_j$. We equivalently say that λ'* is smaller than *λ and write $\lambda' \prec \lambda$.*

We recall that Definition 5.12 defines a total ordering on Λ_n for every n. That is, (1) for any $\lambda, \lambda' \in \Lambda_n$, exactly one of $\lambda \succ \lambda'$, $\lambda \prec \lambda'$, and $\lambda = \lambda'$ is true, and (2) the relation $\succ$ (and therefore also $\prec$) is transitive: if $\lambda \succ \lambda'$ and $\lambda' \succ \lambda''$, then $\lambda \succ \lambda''$. In particular, any finite set of partitions has a maximum element under $\succ$.

Definition 5.13. *For any non-zero vector $\mathbf{w} \in \mathbb{R}(\Lambda_d)$, if λ is the lexicographically largest $\lambda \in \Lambda_d$ whose coefficient in $\mathbf{w}$ is non-zero, and α is the coefficient of λ in $\mathbf{w}$, then we say that λ is the* leading partition of $\mathbf{w}$*, α is the* leading coefficient of $\mathbf{w}$*, and $\alpha \cdot \lambda$ is the* leading term of $\mathbf{w}$*.*

Lemma 5.14. *For any $d \in \mathbb{N}$, there exist vectors $\mathbf{v}_0, \mathbf{v}_1 \in \mathbb{R}_{\geq 0}(\Lambda_d)$ such that the coefficient of $[d]$ in $\mathbf{v}_0$ is 1, the coefficient of $[d]$ in $\mathbf{v}_1$ is 0, and $\mathcal{R}(\mathbf{v}_0) = \mathcal{R}(\mathbf{v}_1)$.*

Proof. Let $\mathbf{v}_0$ and $\mathbf{v}_1$ be defined as follows.

$$\mathbf{v}_0 = \sum_{i=0}^{\lfloor d/2 \rfloor} \binom{d}{2i} \cdot \Gamma_d([d-2i])$$

$$\mathbf{v}_1 = \sum_{i=0}^{\lfloor (d-1)/2 \rfloor} \binom{d}{2i+1} \cdot \Gamma_d([d-2i-1])$$

For each $0 \leq j \leq d$, it holds that

$$\mathcal{R}\big(\Gamma_d([j])\big) = \begin{cases} \Gamma_{d-1}([0]) & \text{if } j = 0 \\ \frac{j}{d} \cdot \Gamma_{d-1}\big([j-1]\big) + \big(1 - \frac{j}{d}\big) \cdot \Gamma_{d-1}\big([j]\big) & \text{if } 0 < j < d \\ \Gamma_{d-1}([d-1]) & \text{if } j = d. \end{cases}$$

Thus, $\mathcal{R}(\mathbf{v}_0 - \mathbf{v}_1)$ is a linear combination of $\big\{\Gamma_{d-1}([j])\big\}_{j=0}^{d-1}$. For each $0 \leq j \leq d-1$, $\Gamma_{d-1}([j])$ appears in the reduction of $\Gamma_d([j])$ and $\Gamma_d([j+1])$ so the coefficient of $\Gamma_{d-1}([j])$ in $\mathcal{R}(\mathbf{v}_0 - \mathbf{v}_1)$ is

$$(-1)^{d-j} \cdot \binom{d}{j} \cdot \left(1 - \frac{j}{d}\right) + (-1)^{d-(j+1)} \cdot \binom{d}{j+1} \cdot \frac{j+1}{d} = (-1)^{d-j} \left(\binom{d-1}{j} - \binom{d-1}{j}\right) = 0. \qquad \square$$

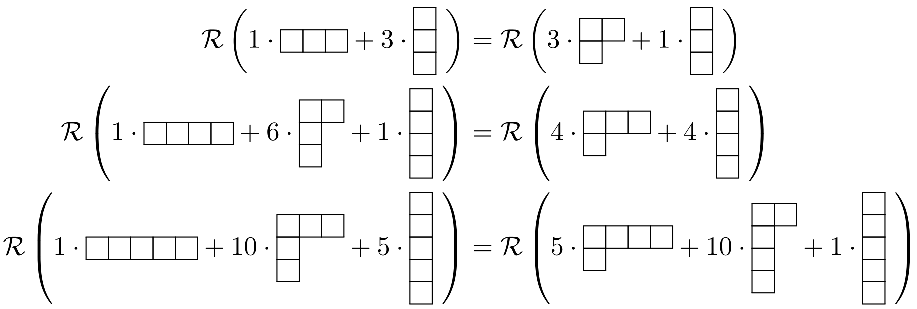


Figure 4: The statement of Lemma 5.14 in the cases $d = 3$, $d = 4$, and $d = 5$.

Lemma 5.15. *For any $d, k \in \mathbb{N}$ such that $d < k$ and any vector $\mathbf{v} \in \mathbb{R}_{\geq 0}(\Lambda_d)$, there exist vectors $\mathbf{u} \in \mathbb{R}_{\geq 0}(\Lambda_k)$ and $\delta \in \mathbb{R}_{\geq 0}(\Lambda_d)$ such that*

1. *$\mathcal{R}^{k-d}(\mathbf{u}) = \mathbf{v} + \delta$, and*
2. *The coefficient of $[k]$ in $\mathbf{u}$ is equal to the coefficient of $[d]$ in $\mathbf{v}$.*

Proof. We construct $\mathbf{u}$ and δ via an iterative process. That is, we construct sequences $\left\{\mathbf{u}^{(i)} \in \mathbb{R}_{\geq 0}(\Lambda_k)\right\}_{i \geq 0}$ and $\left\{\delta^{(i)} \in \mathbb{R}_{\geq 0}(\Lambda_d)\right\}_{i \geq 0}$ satisfying the invariant that for each i, the coefficient of $[k]$ in $\mathbf{u}^{(i)}$ is equal to the coefficient of $[d]$ in $\mathbf{v}$.

We also guarantee the following "progress" condition: The leading partition of $\mathbf{w}^{(i)} \stackrel{\mathsf{def}}{=} \mathbf{v} + \delta^{(i)} - \mathcal{R}^{k-d}(\mathbf{u}^{(i)})$ is lexicographically strictly decreasing with i up until some $i^* \leq |\Lambda_d|$, at which point $\mathbf{w}^{(i^*)} = 0$. Together with the above invariant, this implies that $(\mathbf{u}^{(i^*)}, \delta^{(i^*)})$ are the desired vectors, proving Lemma 5.15. It thus remains only to construct $\{(\mathbf{u}^{(i)}, \delta^{(i)})\}_{i \geq 0}$.

Let $\mathbf{v}_{[d]}$ denote the coefficient of $[d]$ in $\mathbf{v}$. We define $\mathbf{u}^{(0)} = \mathbf{v}_{[d]} \cdot [k]$ and $\delta^{(0)} = 0$. For $i > 0$, if $\mathbf{w}^{(i-1)} = 0$, then we are done. Otherwise, let $\alpha_i \cdot \lambda^{(i)}$ denote the leading term of $\mathbf{w}^{(i-1)}$. Then:

- If $\alpha_i > 0$, let β_i denote the coefficient of $\lambda^{(i)}$ in $\mathcal{R}^{k-d}\big(\Gamma_k(\lambda^{(i)})\big)$. Define $\mathbf{u}^{(i)} = \mathbf{u}^{(i-1)} + \frac{\alpha_i}{\beta_i} \cdot \Gamma_k(\lambda^{(i)})$, and define $\delta^{(i)} = \delta^{(i-1)}$.
- If $\alpha_i < 0$, define $\mathbf{u}^{(i)} = \mathbf{u}^{(i-1)}$, and define $\delta^{(i)} = \delta^{(i-1)} - \alpha_i \cdot \lambda^{(i)}$.

Then it suffices to establish the following claims.

Claim 5.16. *The coefficient of $[k]$ in $\mathbf{u}^{(i)}$ is equal to the coefficient of $[d]$ in $\mathbf{v}$.*

Proof. The proof is by induction. For $i = 0$, this is true by construction. For $i > 0$, it holds because $\mathbf{u}^{(i-1)}$ and $\mathbf{u}^{(i)}$ differ at most in the coefficient of $\Gamma_k(\lambda^{(i-1)})$. But $\Gamma_k(\lambda^{(i-1)}) \neq [k]$ because $\lambda^{(i-1)} \in \Lambda_d$ for $d < k$. □

Claim 5.17. *The vectors $\mathbf{u}^{(i)}$ and $\delta^{(i)}$ have non-negative coefficients.*

Proof. The proof follows immediately by induction. □

Claim 5.18. *For each i, if $\mathbf{w}^{(i)} \neq 0$, then either $\mathbf{w}^{(i+1)} = 0$ or the leading term of $\mathbf{w}^{(i+1)}$ is lexicographically smaller than the leading term of $\mathbf{w}^{(i)}$.*

Proof. The proof is by induction. Suppose that $\mathbf{w}^{(i-1)} \neq 0$ and that its leading term is $\lambda^{(i)}$ with coefficient $\alpha_i \neq 0$. If $\alpha_i < 0$, then the claim follows from the fact that $\mathbf{w}^{(i)} - \mathbf{w}^{(i-1)} = -\alpha_i \cdot \lambda^{(i)}$. If $\alpha_i > 0$, then the claim follows from the fact that $\mathbf{w}^{(i)} - \mathbf{w}^{(i-1)} = -\frac{\alpha_i}{\beta_i} \cdot \Gamma_k(\lambda^{(i)})$, and the fact that the *leading term* of $\Gamma_k(\lambda^{(i)})$ is $\beta_i \cdot \lambda^{(i)}$. □

Since Λ_d is a finite set, the leading term of $\mathbf{w}^{(i)}$ can strictly decrease for at most $|\Lambda_d|$ iterations, resulting in some $i^* \leq |\Lambda_d|$ for which $\mathbf{w}^{(i^*)} = 0$. Then setting $\mathbf{u} = \mathbf{u}^{(i^*)}$ and $\delta = \delta^{(i^*)}$ completes the proof of Lemma 5.15. □

5.4 Proof of Lemma 1

Proof of Lemma 1. Let $\mathcal{G}$ be any game $(\mathcal{Q}, \mathcal{A}, \pi, W)$. We generically transform any *honest-referee* non-signaling strategy $\tilde{P} : \mathrm{Supp}\,(\pi) \to \Delta(\mathcal{A})$ for $\mathcal{G}$ into a (fully) non-signaling strategy $P_{\mathsf{ns}} : \mathcal{Q} \to \Delta(\mathcal{A})$ such that $v[P_{\mathsf{ns}}](\mathcal{G}) \geq \alpha_k \cdot v[\tilde{P}](\mathcal{G})$ for some constant $\alpha_k > 0$ (that does not depend on $\mathcal{G}$).

We will show by induction that for each $d \in [k]$, there is a d-ary marginal non-signaling strategy $\mu^{(d)} = \{\mu_S^{(d)} : \mathcal{Q}_S \to \Delta(\mathcal{A}_S)\}_{S \subseteq [k], 1 \leq |S| \leq d}$ for $\mathcal{G}$ and an honest-referee non-signaling strategy $\tilde{P}^{(d)}$ for $\mathcal{G}$ satisfying:

1. Consistency on d-marginals: For any $q \in \mathrm{Supp}\,(\pi)$ and $S \subseteq [k]$ with $|S| = d$, we have $\left(\tilde{P}^{(d)}(q)\right)_S = \mu_S^{(d)}(q_S)$.
2. Closeness to $\tilde{P}$: For each $q \in \mathrm{Supp}\,(\pi)$, it holds that $\tilde{P}^{(d)}(q)$ is a convex combination
$$\alpha_{k,d} \cdot \tilde{P}(q) + (1 - \alpha_{k,d}) \cdot E^{(d)}(q)$$
for some absolute constant $\alpha_{k,d} > 0$ and a (possibly $\tilde{P}$-dependent) distribution $E^{(d)}(q)$.

We then conclude by defining $P_{\mathsf{ns}} \overset{\mathsf{def}}{=} \mu_{[k]}^{(k)}$.

When $d = 1$, we define $\tilde{P}^{(1)} \overset{\mathsf{def}}{=} \tilde{P}$. For each $i \in [k]$ and $q \in \mathcal{Q}_i$, let $q' \in \mathrm{Supp}\,(\pi)$ be such that $q'_i = q$ and define $\mu_{\{i\}}^{(1)}(q) \overset{\mathsf{def}}{=} \tilde{P}(q')_{\{i\}}$. Such a q' always exists, as without loss of generality $\mathrm{Supp}\,(\pi)_{\{i\}} = \mathcal{Q}_i$. Moreover, because $\tilde{P}$ is (honest-referee) non-signaling, the distribution $\tilde{P}(q')_{\{i\}}$ does not depend on the choice of q'. Thus $\mu^{(1)}$ is a well-defined 1-ary marginal non-signaling strategy.

Assume that properties 1 and 2 hold for an honest-referee non-signaling strategy $\tilde{P}^{(d-1)}$ and a $(d-1)$-ary marginal non-signaling strategy $\mu^{(d-1)} = \{\mu_S^{(d-1)} : \mathcal{Q}_S \to \Delta(\mathcal{A}_S)\}_{S \subseteq [k], |S| = d-1}$. Then, for a choice of $\mathbf{u}_0^* \in \Delta(\Lambda_k)$ and $\mathbf{v}_0^*, \mathbf{v}_1^* \in \Delta(\Lambda_d)$ to be specified later, we define

$$\tilde{P}^{(d)}(q) \overset{\mathsf{def}}{=} \tilde{P}^{(d-1)}_{\mathbf{u}_0^*([k])}(q)$$

and for all $|S| = d$ (extending to $|S| < d$ as described in Remark 5.6), we define

$$\mu_S^{(d)}(q) \overset{\mathsf{def}}{=} \begin{cases} \tilde{P}^{(d-1)}_{\mathbf{v}_0^*(S)}(q) & \text{if } q \in \mathrm{Supp}\,(\pi_S) \\ \\ \mu^{(d-1)}_{\mathbf{v}_1^*(S)}(q) & \text{otherwise.} \end{cases}$$

For this definition to make sense, we must ensure that $\mathbf{v}_1^* \in \Delta(\Lambda_d^{\leq d-1})$ (i.e., the coefficient of $[d]$ is 0).

Let $\mathbf{v}_0, \mathbf{v}_1 \in \mathbb{R}_{\geq 0}(\Lambda_d)$ be vectors, whose existence is guaranteed by Lemma 5.14, such that the coefficient of $[d]$ in $\mathbf{v}_0$ is 1, the coefficient of $[d]$ in $\mathbf{v}_1$ is 0, and $\mathcal{R}(\mathbf{v}_0) = \mathcal{R}(\mathbf{v}_1)$.

Now we apply Lemma 5.15 to $\mathbf{v}_0$ to obtain vectors $\mathbf{u}_0 \in \mathbb{R}_{\geq 0}(\Lambda_k)$ and $\delta \in \mathbb{R}_{\geq 0}(\Lambda_d)$ such that $\mathcal{R}^{k-d}(\mathbf{u}_0) = \mathbf{v}_0 + \delta$ and the coefficient of $[k]$ in $\mathbf{u}_0$, like the coefficient of $[d]$ in $\mathbf{v}_0$, is equal to 1. Then clearly $\mathbf{u}_0, \mathbf{v}_0 + \delta$, and $\mathbf{v}_1 + \delta$ all have nonnegative coefficients, and moreover the sum of their coefficients is the same. The latter follows from the fact that $\mathcal{R}$ preserves coefficient sums, along with the facts that $\mathcal{R}(\mathbf{v}_0) = \mathcal{R}(\mathbf{v}_1)$ and $\mathcal{R}^{k-d}(\mathbf{u}_0) = \mathbf{v}_0 + \delta$. We scale all three vectors so that their coefficients sum to 1, and let $\mathbf{u}_0^* \in \Delta(\Lambda_k)$ and $\mathbf{v}_0^*, \mathbf{v}_1^* \in \Delta(\Lambda_d)$ denote the resulting vectors. Since each vector is scaled by the same factor, it holds that $\mathcal{R}^{k-d}(\mathbf{u}_0^*) = \mathbf{v}_0^*$ and $\mathcal{R}(\mathbf{v}_0^*) = \mathcal{R}(\mathbf{v}_1^*)$.

Claim 5.19. *$\tilde{P}^{(d)}$ is a honest-referee non-signaling strategy.*

Proof. For any $q, q' \in \mathrm{Supp}\,(\pi)\,, T \subseteq [k]$ such that $q_T = q'_T$, we have

$$\left(\tilde{P}^{(d-1)}_{\mathbf{u}_0^*([k])}(q)\right)_T = \tilde{P}^{(d-1)}_{\left(\mathcal{R}^{k-|T|}(\mathbf{u}_0^*)\right)(T)}(q_T) = \tilde{P}^{(d-1)}_{\left(\mathcal{R}^{k-|T|}(\mathbf{u}_0^*)\right)(T)}(q'_T) = \left(\tilde{P}^{(d-1)}_{\mathbf{u}_0^*([k])}(q')\right)_T.$$

The first and third equalities follow from Lemma 5.11, and the second follows from the assumption that $q_T = q'_T$. But the left-hand side and right-hand sides are respectively equal to $\tilde{P}^{(d)}(q)_T$ and $\tilde{P}^{(d)}(q')_T$, which we have therefore shown to be equal. □

Claim 5.20. *$\mu^{(d)}$ is a d-ary marginal non-signaling strategy.*

Proof. For any $S, S' \subseteq [k]$ with $|S| = |S'| = d$ and any $q \in \mathcal{Q}_S, q' \in \mathcal{Q}_{S'}$, let $T \subseteq S \cap S'$ be the largest set such that $q_T = q'_T$. We want to establish that $\mu_S^{(d)}(q)_T = \mu_{S'}^{(d)}(q')_T$. We have the following cases:

- If $q \in \mathrm{Supp}\,(\pi)_S$ and $q' \in \mathrm{Supp}\,(\pi)_{S'}$, then we have

  $$\left(\tilde{P}^{(d-1)}_{\mathbf{v}_0^*(S)}(q)\right)_T = \tilde{P}^{(d-1)}_{\left(\mathcal{R}^{d-|T|}(\mathbf{v}_0^*)\right)(T)}(q_T) = \tilde{P}^{(d-1)}_{\left(\mathcal{R}^{d-|T|}(\mathbf{v}_0^*)\right)(T)}(q'_T) = \left(\tilde{P}^{(d-1)}_{\mathbf{v}_0^*(S')}(q')\right)_T.$$

  The first and third equalities follow from Lemma 5.11, the middle equality from the assumption that $q_T = q'_T$, and the left and right-hand sides are respectively equal to $\mu_S^{(d)}(q)_T$ and $\mu_{S'}^{(d)}(q')_T$ by definition.

- If $q \notin \mathrm{Supp}\,(\pi)_S$ and $q' \notin \mathrm{Supp}\,(\pi)_{S'}$, then we have

  $$\left(\mu^{(d-1)}_{\mathbf{v}_1^*(S)}(q)\right)_T = \mu^{(d-1)}_{\left(\mathcal{R}^{d-|T|}(\mathbf{v}_1^*)\right)(T)}(q_T) = \mu^{(d-1)}_{\left(\mathcal{R}^{d-|T|}(\mathbf{v}_1^*)\right)(T)}(q'_T) = \left(\mu^{(d-1)}_{\mathbf{v}_1^*(S')}(q')\right)_T.$$

  The first and third equalities follow from Lemma 5.9, the middle equality from the assumption that $q_T = q'_T$, and the left and right-hand sides are respectively equal to $\mu_S^{(d)}(q)_T$ and $\mu_{S'}^{(d)}(q')_T$ by definition.

- Otherwise, without loss of generality suppose that $q \in \mathrm{Supp}\,(\pi)_S$ and $q' \notin \mathrm{Supp}\,(\pi)_{S'}$. Then $|T| \leq d-1$ and we have

  $$\left(\tilde{P}^{(d-1)}_{\mathbf{v}_0^*(S)}(q)\right)_T = \tilde{P}^{(d-1)}_{\left(\mathcal{R}^{d-|T|}(\mathbf{v}_0^*)\right)(T)}(q_T) = \mu^{(d-1)}_{\left(\mathcal{R}^{d-|T|}(\mathbf{v}_1^*)\right)(T)}(q'_T) = \left(\mu^{(d-1)}_{\mathbf{v}_1^*(S')}(q')\right)_T.$$

  The first and third equalities follow from Lemmas 5.9 and 5.11. The second equality follows from the identity $\mathcal{R}(\mathbf{v}_0^*) = \mathcal{R}(\mathbf{v}_1^*)$ (implying $\mathcal{R}^n(\mathbf{v}_0^*) = \mathcal{R}^n(\mathbf{v}_1^*)$ for any $n \in \mathbb{N}$) and by the guarantee that $\tilde{P}^{(d-1)}$ and $\mu^{(d-1)}$ agree on marginals of arity at most $d-1$ over $\mathrm{Supp}\,(\pi)_T$ which $q_T = q'_T$ lies in.

By the definition of $\mu^{(d)}$, this shows that $\left(\mu_S^{(d)}(q)\right)_T = \left(\mu_{S'}^{(d)}(q')\right)_T$ in any of the cases. □

Finally, we show that $\mu^{(d)}$ and $\tilde{P}^{(d)}$ satisfy the consistency and closeness conditions defined above. This will complete the inductive step.

Claim 5.21 (Consistency)**.** *For any $q \in \mathrm{Supp}\,(\pi)$ and $S \subseteq [k]$ with $|S| = d$, we have $\left(\tilde{P}^{(d)}(q)\right)_S = \mu_S^{(d)}(q_S)$.*

Proof. Since $\mathcal{R}^{k-d}(\mathbf{u}_0^*) = \mathbf{v}_0^*$, we have

$$\big(\tilde{P}^{(d-1)}_{\mathbf{u}_0^*([k])}(q)\big)_S = \tilde{P}^{(d-1)}_{\left(\mathcal{R}^{k-|S|}(\mathbf{u}_0^*)\right)(S)}(q_S) = \tilde{P}^{(d-1)}_{\mathbf{v}_0^*(S)}(q_S).$$

The first equality is by Lemma 5.11, and the second is by Lemma 5.9. So by the definition of $\tilde{P}^{(d)}$ and $\mu^{(d)}$, we have $\left(\tilde{P}^{(d)}(q)\right)_S = \mu_S^{(d)}(q_S)$. □

Claim 5.22 (Closeness)**.** *For each $q \in \operatorname{Supp}(\pi)$, it holds that $\tilde{P}^{(d)}(q)$ is a convex combination*

$$\alpha_{k,d} \cdot \tilde{P}(q) + (1-\alpha_{k,d}) \cdot E^{(d)}(q)$$

for some distribution $E^{(d)}(q)$, which could be arbitrary, and a constant $\alpha_{k,d} > 0$ that depends only on k and d.

Proof. By definition, $\tilde{P}^{(d)}(q)$ is a convex combination of all "marginal compositions" $\tilde{P}^{(d-1)}_{(S_i)_i}(q)$, where $(S_i)_i$ is a partition of $S = \{1,\ldots,k\}$ with shape λ for some $\lambda \in \Lambda_k$ whose coefficients in $\mathbf{v}_0^*$ are non-zero. In particular, this includes the singleton partition (S) of S – recall that by construction the coefficient of $[k]$ in $\mathbf{u}_0^*$ is non-zero.

Thus there is some positive α such that $\tilde{P}^{(d)}(q)$ can be written as $\alpha \cdot \tilde{P}^{(d-1)}(q) + (1-\alpha) \cdot E'(q)$ for some constant $\alpha > 0$ and some distribution $E'(q) \in \Delta(\mathcal{A})$.

By the inductive hypothesis, $\tilde{P}^{(d-1)}(q) = \alpha_{k,d-1} \cdot \tilde{P}(q) + (1-\alpha_{k,d-1}) \cdot E^{(d-1)}(q)$. Therefore, $\tilde{P}^{(d)} = \alpha_{k,d-1} \cdot \alpha \cdot \tilde{P}(q) + (1-\alpha_{k,d-1} \cdot \alpha) \cdot E^{(d)}(q)$ for some distribution $E^{(d)}(q) \in \Delta(\mathcal{A})$.

The claim follows by setting $\alpha_{k,d} \overset{\mathsf{def}}{=} \alpha_{k,d-1} \cdot \alpha$. □

We conclude the proof of Lemma 1 by setting $P_{\mathsf{ns}} \overset{\mathsf{def}}{=} \mu^{(k)}_{[k]}$. Letting $\alpha_k = \alpha_{k,k}$, we have $v_{\mathsf{ns}}(\mathcal{G}) \geq v[P_{\mathsf{ns}}](\mathcal{G}) \geq \alpha_k \cdot v[\tilde{P}](\mathcal{G})$ for any honest-referee non-signaling $\tilde{P}$, so $v_{\mathsf{ns}}(\mathcal{G}) \geq \alpha_k \cdot v_{\mathsf{hr\text{-}ns}}(\mathcal{G})$. □

6 Strong Parallel Repetition Counterexample

Theorem 4. *For every $k \geq 3$, there exists a k-player game $\mathcal{G}$ such that $v_{\mathsf{ns}}(\mathcal{G}^n) = (2/3)^{\lfloor k/3 \rfloor}$ for any $n \geq 1$.*

We first prove the theorem in the case $k = 3$ in Section 6.1. In this case, the game is the three-player anti-correlation game $\tilde{\mathcal{G}}_3 = (\{0,1\}^3, \{0,1\}^3, \tilde{\pi}_3, \tilde{W}_3)$ defined in Definition 4.2. Then we show via *distributed repetition* that for any $k \geq 3$, this implies a k-player game satisfying Theorem 4.

6.1 A Non-signaling Strategy for $\tilde{\mathcal{G}}_3^n$ with Value 2/3

Proposition 4.1 shows that the non-signaling value of $\tilde{\mathcal{G}}_3$ (and hence of $\tilde{\mathcal{G}}_3^n$) is at most 2/3. Now we show that the hardness of this game is *completely non-amplifying*. That is, we construct a non-signaling strategy for $\tilde{\mathcal{G}}_3^n$ that achieves value 2/3.

Proposition 6.1. *For every $n \geq 1$, $v_{\mathsf{ns}}(\tilde{\mathcal{G}}_3^n) = 2/3$.*

Proof. We give a non-signaling strategy P for $\tilde{\mathcal{G}}_3^n$ and show that $v[P](\tilde{\mathcal{G}}_3^n) = 2/3$.

Construction 6.2. *Given $q = (q^{(1)},\ldots,q^{(n)})$, P samples answers $a = (a^{(1)},\ldots,a^{(n)})$ as follows.*

- *If no $q^{(i)}$ is equal to 111, then:*
  1. *Sample $b \leftarrow \mathrm{Ber}(1/3)$, i.e. $b = 1$ with probability 2/3 and $b = 0$ otherwise.*
  2. *Sample each $a^{(i)}$ independently and uniformly at random, subject to the constraint that if $q_j^{(i)} = q_k^{(i)} = 1$ for some $j \neq k$, then $a_j^{(i)} \oplus a_k^{(i)} = b$.*
- *If some $q^{(i)}$ is equal to 111, then:*
  1. *Sample $t \leftarrow \{1,2,3\}$ uniformly at random.*
  2. *Sample each $a^{(i)}$ independently and uniformly at random, subject to the constraint that if $q_j^{(i)} = q_k^{(i)} = 1$ for some $j \neq k$, then:*
     - *If $j = t$ or $k = t$, then $a_j^{(i)} \neq a_k^{(i)}$.*

– *Otherwise,* $a_j^{(i)} = a_k^{(i)}$.

Loosely speaking, in the first case, P randomly decides with probability 1/3 to lose all instances of $\mathcal{G}_3$. In the second case, P randomly chooses a designated player t to disagree with all other players receiving 1. It may seem strange that in the first case, P artificially chooses to lose all the games. However, this is necessary for the existence of a consistent answer distribution when all players receive 1 queries.

We claim that the value of $\tilde{\mathcal{G}}_3^n$ with respect to P is 2/3. This is solely determined by P's behavior in the first case, because for honestly generated queries, no $q^{(i)}$ is ever equal to 111. In the first case, with probability 2/3 (whenever $b = 1$), the answers $a_j^{(i)}$ and $a_k^{(i)}$ corresponding to the "1" queries in the i^{th} game satisfy $a_j^{(i)} \neq a_k^{(i)}$ for every i, so P wins $\tilde{\mathcal{G}}_3^n$ with probability 2/3.

It remains to verify that P is non-signaling, i.e. that for all sets $S \subseteq \{1, 2, 3\}$, the distribution $P(q)_S$ depends only on q_S. This is trivially true when $|S| = 0$ or $|S| = 3$. The remaining cases are $|S| = 1$ and $|S| = 2$.

These cases are easier to verify when keeping in mind the structure of P: based on q, P probabilistically chooses a set of constraints on $a^{(1)}, \dots, a^{(n)}$. Each constraint specifies the equality or inequality of different components of each $a^{(i)}$. P then independently chooses $a^{(i)}$ satisfying the constraints. Thus, to demonstrate that the distribution of a_S depends only on q_S, it suffices to show that the distribution of the constraints on a_S depends only on q_S.

Case 1: $|S| = 1$**.** For any q, we claim that the distribution $P(q)_S$ is uniformly random on $\{0, 1\}^n$, and thus depends only on q_S (in fact, on nothing) as required.

This holds because all constraints chosen by P satisfy

- Symmetry: The constraints only enforce equality or inequality of specific bits of a. Thus, when a is chosen uniformly at random to satisfy these constraints, each individual bit of a is equally likely to be 0 or 1.
- Independence: Each constraint only relates the bits of a single $a^{(i)}$. Thus, $a_S^{(1)}, \dots, a_S^{(n)}$ are independent as random variables.

Case 2: $|S| = 2$**.** For any q, we claim that $(a_S^{(1)}, \dots, a_S^{(n)}) = P(q)_S$ is distributed as follows. For concreteness say that $S = \{j, k\}$. For any q, we have:

- With probability 2/3, the constraints generated by P on $a_S^{(i)}$ are that $a_S^{(i)} \in \{01, 10\}$ for all i for which $q_S^{(i)} = 11$. In particular, P generates these constraints if $b = 1$ (when no $q^{(i)}$ is 111), and when $t \in S$ (when some $q^{(i)}$ is 111).
- Otherwise the constraints generated by P on $a_S^{(i)}$ are that $a_S^{(i)} \in \{00, 11\}$ for all i for which $q_S^{(i)} = 11$.

We note that P may also generate constraints on $a^{(i)}$ beyond those explicitly mentioned above, specifically when $q_j^{(i)} = 1$ for some $j \notin S$. However, inspection of P reveals that these constraints do not affect the distribution of $a_S^{(i)}$. For example, suppose that $S = \{1, 2\}$, $q^{(i)} = 111$, and $t = 2$. Then the constraints generated by P require not only that $a_1^{(i)} \neq a_2^{(i)}$, but also that $a_1^{(i)} = a_3^{(i)}$ and $a_2^{(i)} \neq a_3^{(i)}$. In this case, the latter two constraints are *redundant*: whenever $a_1^{(i)} \neq a_2^{(i)}$, they are satisfiable for a unique choice of $a_3^{(i)}$. Thus, the redundant constraints do not affect the distribution of $a_S^{(i)}$. □

6.2 Proof of Theorem 4

Definition 6.3 (Distributed Repetition)**.** *Given a k-player game* $\mathcal{G} = (\mathcal{Q}, \mathcal{A}, \pi, W)$ *where* $\mathcal{Q} = \mathcal{Q}_1 \times \cdots \times \mathcal{Q}_k$ *and* $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_k$*, its* ℓ-fold distributed repetition *is defined as the* k'*-player game* $\mathcal{G}^{\|\ell} = (\mathcal{Q}^{\|\ell}, \mathcal{A}^{\|\ell}, \pi^{\|\ell}, W^{\|\ell})$ *for* $k' = k \cdot \ell$ *where:*

- $\mathcal{Q}^{\|\ell} \overset{\text{def}}{=} \mathcal{Q}'_1 \times \cdots \times \mathcal{Q}'_{k'}$, *where* $\mathcal{Q}'_i = \mathcal{Q}_j$ *for the unique* $j \in [k]$ *such that* $j \equiv i \pmod{k}$*,*
- $\mathcal{A}^{\|\ell} \overset{\text{def}}{=} \mathcal{A}'_1 \times \cdots \times \mathcal{A}'_{k'}$, *where* $\mathcal{A}'_i = \mathcal{A}_j$ *for the unique* $j \in [k]$ *such that* $j \equiv i \pmod{k}$*,*
- $\pi^{\|\ell}(q_1, \ldots, q_{k'}) \overset{\text{def}}{=} \prod_{i=0}^{\ell-1} \pi(q_{i\cdot k+1}, \ldots, q_{i\cdot k+k})$*, and*
- $W^{\|\ell}\big((q_1, \ldots, q_{k'}), (a_1, \ldots, a_{k'})\big) \overset{\text{def}}{=} \prod_{i=0}^{\ell-1} W\big((q_{i\cdot k+1}, \ldots, q_{i\cdot k+k}), (a_{i\cdot k+1}, \ldots, a_{i\cdot k+k})\big)$.

Theorem 4 follows from Proposition 6.4 below. The proposition is folklore, but we include a proof for completeness, as we could not find one in the literature.

Proposition 6.4. *For any game* $\mathcal{G}$*, its* ℓ*-fold distributed repetition* $\mathcal{G}^{\|\ell}$ *satisfies* $v_{\mathsf{ns}}(\mathcal{G}^{\|\ell}) = v_{\mathsf{ns}}(\mathcal{G})^{\ell}$*.*

Proof. First we show that $v_{\mathsf{ns}}(\mathcal{G}^{\|\ell}) \geq v_{\mathsf{ns}}(\mathcal{G})^{\ell}$. Suppose that $\mathcal{G} = (\mathcal{Q}, \mathcal{A}, \pi, W)$ is a k-player game, and let P be a non-signaling strategy for $\mathcal{G}$ such that $v[P](\mathcal{G}) = v_{\mathsf{ns}}(\mathcal{G})$. Define the following strategy for $\mathcal{G}^{\|\ell}$:

$$P^{\|\ell}(q_1, \ldots, q_{\ell\cdot k}) \overset{\text{def}}{=} \prod_{i=0}^{\ell-1} P\left(q_{i\cdot k+1}, \ldots, q_{i\cdot k+k}\right).$$

That is, $P^{\|\ell}(q_1, \ldots, q_{\ell\cdot k})$ is the distribution on $(a_1, \ldots, a_{\ell\cdot k})$ obtained by independently sampling

$$(a_{i\cdot k+1}, \ldots, a_{i\cdot k+k}) \leftarrow P(q_{i\cdot k+1}, \ldots, q_{i\cdot k+k})$$

for each $i \in \{0, \ldots, \ell-1\}$.

Claim 6.5. $v[P^{\|\ell}](\mathcal{G}^{\|\ell}) = \big(v[P](\mathcal{G})\big)^{\ell} = v_{\mathsf{ns}}(\mathcal{G})^{\ell}$.

Proof. This follows directly from the definition of $\mathcal{G}^{\|\ell}$ and $P^{\|\ell}$. □

Claim 6.6. $P^{\|\ell}$ *is non-signaling.*

Proof. For any $q, q' \in \mathcal{Q}^{\|\ell}$, we need to show that if $q_S = q'_S$, then a_S and a'_S are identically distributed in the probability space defined by sampling

$$(a_1, \ldots, a_{\ell\cdot k}) \leftarrow P^{\|\ell}(q)$$
$$(a'_1, \ldots, a'_{\ell\cdot k}) \leftarrow P^{\|\ell}(q').$$

Let S_i denote the set $S \cap \{i\cdot k+1, \ldots, i\cdot k+k\}$. By definition of $P^{\|\ell}$, the tuples $\big\{(a_{i\cdot k+1}, \ldots, a_{i\cdot k+k})\big\}_{i=0}^{\ell-1}$ are mutually independent, as are the tuples $\big\{(a'_{i\cdot k+1}, \ldots, a'_{i\cdot k+k})\big\}_{i=0}^{\ell-1}$. Thus, it suffices for us to show that for each $i \in \{0, \ldots, \ell-1\}$, a_{S_i} and a'_{S_i} are identically distributed. This in turn follows from the definition of $P^{\|\ell}$ and the fact that P is non-signaling. □

Now we show that $v_{\mathsf{ns}}(\mathcal{G}^{\|\ell}) \leq v_{\mathsf{ns}}(\mathcal{G})^{\ell}$. The bound holds trivially when $\ell = 1$. For $\ell > 1$, we proceed by induction. Suppose we have established that $v_{\mathsf{ns}}(\mathcal{G}^{\|\ell-1}) \leq v_{\mathsf{ns}}(\mathcal{G})^{\ell-1}$.

Suppose for the sake of contradiction that there exists a non-signaling strategy P for $\mathcal{G}^{\|\ell}$ for which $v[P](\mathcal{G}^{\|\ell}) > v_{\mathsf{ns}}(\mathcal{G})^{\ell}$. We will construct a non-signaling strategy P' for $\mathcal{G}$ for which $v[P'](\mathcal{G}) > v_{\mathsf{ns}}(\mathcal{G})$, which is impossible.

We define $P'(q_1, \ldots, q_k)$ to be the distribution on $(a_1, \ldots, a_k)$ obtained by the following process.

1. Sample $(q_{i\cdot k+1}, \ldots, q_{i\cdot k+k}) \leftarrow \pi$ for each $i \in \{1, \ldots, \ell-1\}$.
2. Sample $(a_1, \ldots, a_{\ell\cdot k}) \leftarrow P(q_1, \ldots, q_{\ell\cdot k})$.

3. If $W\big(q_{i\cdot k+1},\dots,q_{i\cdot k+k},a_{i\cdot k+1},\dots,a_{i\cdot k+k}\big) = 1$ for each $i \in \{1,\dots,\ell-1\}$, then output $(a_1,\dots,a_k)$. Otherwise, try again (from step 1).

For this to be a well-defined distribution, there must be a finite number of iterations with probability 1. This holds because $v[P](\mathcal{G}^{\|\ell}) > 0$, so the process halts on Step 3 with constant probability in each iteration.

Claim 6.7. *P' is non-signaling.*

Proof. Consider any $(q_1,\dots,q_k),(q'_1,\dots,q'_k) \in \mathcal{Q}$ and $S \subseteq [k]$ such that $q_S = q'_S$, and consider the probability space defined by sampling

$$\begin{aligned}
&(q_{i\cdot k+1},\dots,q_{i\cdot k+k}) \leftarrow \pi \quad \text{for each } i \in \{1,\dots,\ell-1\}\\
&(a_1,\dots,a_{\ell\cdot k}) \leftarrow P(q_1,\dots,q_k,q_{k+1},\dots,q_{\ell\cdot k})\\
&(a'_1,\dots,a'_{\ell\cdot k}) \leftarrow P(q'_1,\dots,q'_k,q_{k+1},\dots,q_{\ell\cdot k}).
\end{aligned}$$

Since P is non-signaling,

$$\big((q_{i\cdot k+q},\dots,q_{i\cdot k+k}),a_S\big) \tag{6}$$

and

$$\big((q_{i\cdot k+q},\dots,q_{i\cdot k+k}),a'_S\big) \tag{7}$$

are identically distributed. $P'(q_1,\dots,q_k)_S$ is obtained as a deterministic function of Eq. (6), and $P'(q'_1,\dots,q'_k)_S$ is obtained as the same deterministic function of Eq. (7). Hence, $P'(q_1,\dots,q_k)_S$ and $P'(q'_1,\dots,q'_k)_S$ are equal as distributions. □

To relate the values $v[P](\mathcal{G}^{\|\ell})$ and $v[P'](\mathcal{G})$, we define $P''(q_1,\dots,q_{(\ell-1)\cdot k})$ to be the distribution on $(a_1,\dots,a_{(\ell-1)\cdot k})$ obtained by sampling $(q_{(\ell-1)\cdot k+1},\dots,q_{\ell\cdot k}) \leftarrow \pi$ and $(a_1,\dots,a_{\ell\cdot k}) \leftarrow P(q_1,\dots,q_{\ell\cdot k})$.

Claim 6.8. *P'' is non-signaling.*

Proof. Follows from the definition of P'' and the fact that P is non-signaling. □

Consider the probability space defined by sampling

$$\begin{aligned}
&(q_{i\cdot k+1},\dots,q_{i\cdot k+k}) \leftarrow \pi \quad \text{for each } i \in \{0,\dots,\ell-1\}\\
&(a_1,\dots,a_{\ell\cdot k}) \leftarrow P(q_1,\dots,q_{\ell\cdot k})
\end{aligned}$$

For $i \in \{0,\dots,\ell-1\}$, let W_i denote the event[3] that $W\big((q_{i\cdot k+1},\dots,q_{i\cdot k+k}),(a_{i\cdot k+1},\dots,a_{i\cdot k+k})\big) = 1$. We then have

$$\begin{aligned}
v[P](\mathcal{G}^{\|\ell}) &= \Pr[W_1 \wedge \ldots \wedge W_{\ell-1}] \cdot \Pr[W_0 | W_1 \wedge \ldots \wedge W_{\ell-1}]\\
&= v[P''](\mathcal{G}^{\|\ell-1}) \cdot \mathbb{E}_{(q_1,\dots,q_k)}\left[\Pr\left[W_0 | (q_1,\dots,q_k), W_1 \wedge \cdots \wedge W_{\ell-1}\right]\right]\\
&= v[P''](\mathcal{G}^{\|\ell-1}) \cdot v[P'](\mathcal{G})\\
&\le v_{\mathsf{ns}}(\mathcal{G}^{\|\ell-1}) \cdot v_{\mathsf{ns}}(\mathcal{G})\\
&\le v_{\mathsf{ns}}(\mathcal{G})^{\ell},
\end{aligned}$$

where the last inequality is by the inductive hypothesis. This establishes that $v_{\mathsf{ns}}(\mathcal{G}^{\|\ell}) \le v_{\mathsf{ns}}(\mathcal{G})^{\ell}$, completing the proof of Proposition 6.4. □

[3] Here we abuse notation and write $W(q,a)$ to denote a Bernoulli random variable with parameter $W(q,a)$.

Proof of Theorem 4. We prove Theorem 4 when k is a multiple of 3; the more general case follows by adding one or two dummy players. For $k = 3$, the three-player anti-correlation game $\tilde{\mathcal{G}}_3$ has $v_{\mathsf{ns}}(\tilde{\mathcal{G}}_3^n) = 2/3$ for any $n \geq 1$ by Proposition 6.1.

If $k = 3\ell$ for $\ell > 1$, the k-player game in question is $\mathcal{G} \overset{\mathsf{def}}{=} (\tilde{\mathcal{G}}_3)^{\|\ell}$. To analyze the non-signaling value of its n-fold parallel repetition, we first observe that game $\mathcal{G}^n$ is equivalent to $(\tilde{\mathcal{G}}_3^n)^{\|\ell}$. Thus, applying Proposition 6.4 to the game $\tilde{\mathcal{G}}_3^n$, we obtain

$$v_{\mathsf{ns}}(\mathcal{G}^n) = v_{\mathsf{ns}}\left((\tilde{\mathcal{G}}_3^n)^{\|\ell}\right) = v_{\mathsf{ns}}(\tilde{\mathcal{G}}_3^n)^\ell = (2/3)^\ell. \qquad \square$$

Acknowledgments

We thank Ron Rothblum for his generous help in reading and commenting on an early version of this manuscript.

References

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, 1998.

[BFL91] László Babai, Lance Fortnow, and Carsten Lund. Non-deterministic exponential time has two-prover interactive protocols. *Computational Complexity*, 1:3–40, 1991.

[BFLS91] László Babai, Lance Fortnow, Leonid A. Levin, and Mario Szegedy. Checking computations in polylogarithmic time. In *STOC*, pages 21–31. ACM, 1991.

[BFS14] Harry Buhrman, Serge Fehr, and Christian Schaffner. On the parallel repetition of multi-player games: The no-signaling case. In *TQC*, volume 27 of *LIPIcs*, pages 24–35. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2014.

[BGKW88] Michael Ben-Or, Shafi Goldwasser, Joe Kilian, and Avi Wigderson. Multi-prover interactive proofs: How to remove intractability assumptions. In *STOC*, pages 113–131. ACM, 1988.

[BHK17] Zvika Brakerski, Justin Holmgren, and Yael Tauman Kalai. Non-interactive delegation and batch NP verification from standard computational assumptions. In *STOC*, pages 474–482. ACM, 2017.

[BKK+18] Saikrishna Badrinarayanan, Yael Tauman Kalai, Dakshita Khurana, Amit Sahai, and Daniel Wichs. Non-interactive delegation for low-space non-deterministic computation. In *STOC*, page to appear. ACM, 2018.

[CHSH69] John F Clauser, Michael A Horne, Abner Shimony, and Richard A Holt. Proposed experiment to test local hidden-variable theories. *Physical review letters*, 23(15):880, 1969.

[Fei91] Uriel Feige. On the success probability of the two provers in one-round proof systems. In *Structure in Complexity Theory Conference*, pages 116–123. IEEE Computer Society, 1991.

[FGL+96] Uriel Feige, Shafi Goldwasser, László Lovász, Shmuel Safra, and Mario Szegedy. Interactive proofs and the hardness of approximating cliques. *J. ACM*, 43(2):268–292, 1996.

[FL92] Uriel Feige and László Lovász. Two-prover one-round proof systems: Their power and their problems (extended abstract). In *STOC*, pages 733–744. ACM, 1992.

[For89] Lance Jeremy Fortnow. *Complexity-theoretic aspects of interactive proof systems.* PhD thesis, 1989.

[FRS88] Lance Fortnow, John Rompel, and Michael Sipser. On the power of multi-prover interactive protocols. In *Structure in Complexity Theory Conference*, pages 156–161. IEEE Computer Society, 1988.

[FRV16] Rotem Arnon Friedman, Renato Renner, and Thomas Vidick. Non-signaling parallel repetition using de finetti reductions. *IEEE Trans. Information Theory*, 62(3):1440–1457, 2016.

[Hås01] Johan Håstad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001.

[Hol09] Thomas Holenstein. Parallel repetition: Simplification and the no-signaling case. *Theory of Computing*, 5(1):141–172, 2009.

[IKM09] Tsuyoshi Ito, Hirotada Kobayashi, and Keiji Matsumoto. Oracularization and two-prover one-round interactive proofs against nonlocal strategies. In *IEEE Conference on Computational Complexity*, pages 217–228. IEEE Computer Society, 2009.

[Kil92] Joe Kilian. A note on efficient zero-knowledge proofs and arguments (extended abstract). In *STOC*, pages 723–732. ACM, 1992.

[KRR13] Yael Tauman Kalai, Ran Raz, and Ron D. Rothblum. Delegation for bounded space. In *STOC*, pages 565–574. ACM, 2013.

[KRR14] Yael Tauman Kalai, Ran Raz, and Ron D. Rothblum. How to delegate computations: the power of no-signaling proofs. In *STOC*, pages 485–494. ACM, 2014.

[LW16] Cécilia Lancien and Andreas Winter. Parallel repetition and concentration for (sub-)no-signalling games via a flexible constrained de finetti reduction. *Chicago J. Theor. Comput. Sci.*, 2016, 2016.

[Raz98] Ran Raz. A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803, 1998.

[Ver96] Oleg Verbitsky. Towards the parallel repetition conjecture. *Theor. Comput. Sci.*, 157(2):277–282, 1996.

[Yue16] Henry Yuen. A parallel repetition theorem for all entangled games. In *ICALP*, volume 55 of *LIPIcs*, pages 77:1–77:13. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2016.