Electronic Colloquium on Computational Complexity, Report No. 141 (2019)

On Rich 2-to-1 Games

Mark Braverman * Subhash Khot † Dor Minzer ‡

Abstract

We propose a variant of the 2-to-1 Games Conjecture that we call the Rich 2-to-1 Games Conjecture and show that it is equivalent to the Unique Games Conjecture. We are motivated by two considerations. Firstly, in light of the recent proof of the 2-to-1 Games Conjecture [16, 6, 5, 17], we hope to understand how one might make further progress towards a proof of the Unique Games Conjecture. Secondly, the new variant along with perfect completeness in addition, might imply hardness of approximation results that necessarily require perfect completeness and (hence) are not implied by the Unique Games Conjecture.

1 Introduction

The Unique Games Conjecture [11] is considered a central question in theoretical computer science. It has many applications to hardness of approximation (e.g. tight results for Max-Cut and Vertex Cover problems [14, 18]) and connections to algorithms, computational complexity, analysis, and geometry (e.g. see the surveys [22, 12, 13]). Recently, a related conjecture called the 2-to-1 Games Conjecture has been proved [16, 6, 5, 17]. This conjecture has many applications of its own, implies the Unique Games Conjecture "halfway" (in the technical sense, with "completeness" $\frac{1}{2}$ instead of $1 - o(1)$), and provides strong evidence in favor of the Unique Games Conjecture.

In light of this development, it is natural to ask whether the proof of the 2-to-1 Games Conjecture can somehow be extended to that of the Unique Games Conjecture. A straightforward extension does not look likely, so we raise the following possibility: perhaps the 2-to-1 Games Conjecture holds with additional structure on its instances, and hardness on such instances is then enough to prove the Unique Games Conjecture? In this paper, we investigate this possibility and make a concrete proposal in this regard. The proposal, that we call the Rich 2-to-1 Games Conjecture, is described next along with the overall context. Our main result is that this variant of the 2-to-1 Games Conjecture turns out to be equivalent to the Unique Games Conjecture.

1.1 The Unique Games Conjecture

The Unique Games and 2-to-1 Games are specialized cases of the more general 2-Prover-1-Round Games.

Definition 1.1. *A 2P1R Games instance $\Psi = (L \cup R, E, \Sigma_L, \Sigma_R, \Phi)$ consists of a regular, bipartite graph $(L \cup R, E)$, the alphabet Σ_L for the vertex set L, the alphabet Σ_R for the vertex set R, and a set of constraints*

*Department of Computer Science, Princeton University. Research supported in part by the NSF Alan T. Waterman Award, Grant No. 1933331, a Packard Fellowship in Science and Engineering, and the Simons Collaboration on Algorithms and Geometry.

†Department of Computer Science, Courant Institute of Mathematical Sciences, New York University. Supported by the NSF Award CCF-1422159, the Simons Collaboration on Algorithms and Geometry, and the Simons Investigator Award.

‡Institute for Advanced Study. Supported partially by NSF grant DMS-1638352 and Rothschild Fellowship.

ISSN 1433-8092

$\Phi = \{\phi_e\}_{e \in E}$*, one for each edge. Each vertex is supposed to receive a label from the respective alphabet. The constraint* ϕ_e *for an edge* $e = (u, v) \in E, u \in L, v \in R$ *is defined by a relation* $\phi_e \subseteq \Sigma_L \times \Sigma_R$*, thought of as the set of label-pairs to the vertices* u *and* v *that satisfy the constraint.*

For $1 \geqslant c > s > 0$, and integers k, n, let Gap-2P1R$_{k,n}[c, s]$ denote the promise problem where given a 2P1R Games instance Ψ as above with $|\Sigma_L| = k, |\Sigma_R| = n$, the problem is to distinguish whether there is a labeling to its vertices that satisfies c fraction of the constraints or whether every labeling satisfies at most s fraction of the constraints (we will often drop the subscripts k, n when clear from context).

2P1R Games are central to the theory of hardness of approximation and Probabilistically Checkable Proofs. These serve as canonical starting point for hardness reductions. The parameters of interest (from the viewpoint of making such reductions "work") are: the alphabet size $\max\{m, k\}$, the "gap" (c, s), and the nature of the relations ϕ_e. Throughout this paper, the parameters k, n, c, s are thought of as constants and the size of the bipartite graph $(L \cup R, E)$ as the instance size.

The 2P1R Games studied in applications are almost exclusively "Projection Games", i.e. instances in which $|\Sigma_L| \geqslant |\Sigma_R|$ and the constraint on each edge $e = (u, v)$ is defined by a mapping $\pi_e \colon \Sigma_L \to \Sigma_R$; the relation ϕ_e is then

$$\phi_e = \{\, (\sigma, \pi_e(\sigma)) \mid \sigma \in \Sigma_L \},$$

so that for every label to vertex u, there is a unique label to the vertex v that satisfies the constraint. We will restriction to Projection Games henceforth and denote the corresponding gap problem as Gap-Projection.

In the language of 2P1R Games, the celebrated PCP Theorem [8, 2, 1] states that Gap-Projection$_{7,2}[1, s]$ is NP-hard for some absolute constant $s < 1$. Combining the PCP Theorem and Raz's Parallel Repetition Theorem [20] gives the very important theorem that Gap-Projection$_{k,n}[1, s]$ is NP-hard for every constant $s > 0$ and with the alphabet size at most polynomial in $\frac{1}{s}$.

For an integer d (thought of as a small constant, say $d = 2$), a d-to-1 Games instance is a Projection Games instance in which $|\Sigma_L| = d \cdot |\Sigma_R|$ and the projection map $\pi_e \colon \Sigma_L \to \Sigma_R$ defining the constraint is a d-to-1 map. The 1-to-1 Games are more commonly called the Unique Games and were studied by Feige and Lovasz [9] (in a different context). The corresponding gap versions are denoted as Gap-d-to-1 and Gap-Unique and the alphabet sizes are identified by one paramter n such that $|\Sigma_L| = d \cdot n$ and $|\Sigma_R| = n$. The conjectures made in [11] are stated below (we take liberty to modify statements slightly regarding the issue of perfect versus imperfect completeness):

Conjecture 1.2. Unique Games Conjecture
For every constant $\varepsilon > 0$*, there is a sufficiently large integer* n *such that* Gap-Unique$_n[1-\varepsilon, \varepsilon]$ *is NP-hard.*

Conjecture 1.3. d**-to-**1 **Games Conjecture**
For every constant $\varepsilon > 0$*, there is a sufficiently large integer* n *such that* Gap-d-to-$1_n[1 - \varepsilon, \varepsilon]$ *is NP-hard.*

Conjecture 1.4. d**-to-**1 **Games Conjecture with Perfect Completeness**
For every constant $\varepsilon > 0$*, there is a sufficiently large integer* n *such that* Gap-d-to-$1_n[1, \varepsilon]$ *is NP-hard.*

In a recent development, the 2-to-1 Games Conjecture is proved in a sequence of papers [16, 6, 5, 17] (with additional contributions from [3, 15]), also proving as a simple corollary that Gap-Unique$[\frac{1}{2}, \varepsilon]$ is NP-hard (for every $\varepsilon > 0$ and for sufficiently large alphabet size). This gives a strong evidence towards correctness of the Unique Games Conjecture (which prior to this development was viewed skeptically by most researchers).

1.2 The Rich 2-to-1 Games

One naturally asks whether the proof of the 2-to-1 Games Conjecture extends, without substantial effort, to that of the Unique Games Conjecture. We do not believe this to be the case and instead make the following proposal and conjecture. We conjecture that the 2-to-1 Games Conjecture holds with additional structure on its instances (referred to as "richness") and is then enough to prove the Unique Games Conjecture (in fact is equivalent to it). The new conjecture and the notion of richness are well-motivated as explained later on.

Let $\Psi = (L \cup R, E, \Sigma_L, \Sigma_R, \Phi)$ be a 2-to-1-Game, with $|\Sigma_L| = 2n$ and $|\Sigma_R| = n$. Fix a vertex $u \in L$. Let $e = (u, v) \in E$ be an edge incident on u and let π_e be the 2-to-1 projection defining that constraint. The map defines a partition of Σ_L as

$$\Sigma_L = \bigcup_{\rho \in \Sigma_R} \pi_e^{-1}(\rho)$$

into disjoint sets of size 2. Let us denote by $\mathcal{P}(u)$ the distribution over partitions of Σ_L into sets of size 2, given by first sampling a uniformly random edge $e = (u, v)$ incident on u and then outputting the partition of Σ_L as above.

Definition 1.5. *An instance of Rich 2-to-1 Games is an instance of 2-to-1 Games with the additional property that for every vertex $u \in L$, the distribution $\mathcal{P}(u)$ is uniform over all partitions of Σ_L into sets of size 2.*

We now state the new conjecture (and also throw in a stronger version with perfect completeness). Our main result is that it is equivalent to the Unique Games Conjecture.

Conjecture 1.6. Rich 2-to-1 Games Conjecture
For every constant $\varepsilon > 0$, there is a sufficiently large integer n such that Gap-Rich-2-to-$1_n[1 - \varepsilon, \varepsilon]$ *is NP-hard.*

Conjecture 1.7. Rich 2-to-1 Games Conjecture with Perfect Completeness
For every constant $\varepsilon > 0$, there is a sufficiently large integer n such that Gap-Rich-2-to-$1_n[1, \varepsilon]$ *is NP-hard.*

Theorem 1.8. Main Result
The Unique Games Conjecture 1.2 and the Rich 2-to-1 Games Conjecture 1.6 are equivalent.

The reduction from Unique Games to Rich 2-to-1 Games is straightforward, and is given in Appendix B for completeness. The reverse reduction requires new analytic results to analyze it. These results are stated in Section 3 and proved in Section 4. The reduction itself is presented in Section 5.

1.3 Motivation to Study the Rich 2-to-1 Games

We now explain how the notion of richness arises from natural (but admittedly technical) considerations. In short, the notion of richness is tailor-made so as to ensure the "sub-code covering" property; this property was identified and used in [19] and was crucial in the proof of the 2-to-2 Games Conjecture [16, 6] (however there are differences that are outlined below). We then comment on how the notion of richness might be useful towards proving the Unique Games Conjecture and towards proving hardness of approximation results with perfect completeness. These comments are speculative in nature.

Sub-code Covering Property

We describe, at a very high level, a typical PCP reduction starting with an instance of a Projection Game.[1] We admit that the description might not be friendly to a reader who is not already somewhat familiar with the area.

Let $\Psi = (L \cup R, E, \Sigma_L, \Sigma_R, \Phi)$ be an instance of Projection Game. In the reduction (or equivalently the PCP proof), each vertex $u \in L$ is replaced by a string $\mathsf{Enc}^*(u) \in [m]^{k_L}$ which is intended to be the encoding of the supposed label of u via an encoding scheme $\mathsf{Enc} : \Sigma_L \to [m]^{k_L}$. The encoding scheme is chosen a priori. Here $[m]$ is the proof alphabet (e.g. $\{0, 1\}$) and k_L is the encoding length. Similarly, each vertex $v \in R$ is replaced by a string $\mathsf{Enc}^*(v) \in [m]^{k_R}$ which is intended to be the encoding of the supposed label of v via the encoding scheme $\mathsf{Enc} : \Sigma_R \to [m]^{k_R}$. For convenience, we use the same notation, namely $\mathsf{Enc}(\cdot)$, to denote both encodings. Also, similar notation, namely $\mathsf{Enc}^*(\cdot)$ and $\mathsf{Enc}(\cdot)$, is used to emphasize their relationship: the latter is a true encoding whereas the former is a purported encoding.

The task of the PCP verifier is to check, given a purported proof and an edge $e = (u, v) \in E$,

- that the strings $\mathsf{Enc}^*(u)$ and $\mathsf{Enc}^*(v)$ in the purported proof are indeed codewords, i.e. that they are same as $\mathsf{Enc}(\sigma)$ for some label $\sigma \in \Sigma_L$ and $\mathsf{Enc}(\rho)$ for some label $\rho \in \Sigma_R$ respectively.
- that $\pi_e(\sigma) = \rho$ where $\pi_e : \Sigma_L \to \Sigma_R$ is the projection map defining the constraint.

These two tasks are referred to as the codeword test and the consistency test respectively and are often somehow incorporated into a single combined test (as seen below). Further, a combination of necessity and convenience dictates that:

- One needs to work with a relaxed conclusion that $\mathsf{Enc}^*(u)$ and $\mathsf{Enc}^*(v)$ are close to some codewords $\mathsf{Enc}(\sigma)$ and $\mathsf{Enc}(\rho)$ respectively so that $\pi_e(\sigma) = \rho$. This amounts to decoding or (more often) list-decoding the given strings $\mathsf{Enc}^*(u)$ and $\mathsf{Enc}^*(v)$.
- One needs that the codeword $\mathsf{Enc}(\rho)$ is a "sub-code" of the codeword $\mathsf{Enc}(\sigma)$ whenever $\pi_e(\sigma) = \rho$. Specifically, for every location $x \in [k_R]$ on the v-side, there is a location $\pi_e^{-1}(x) \in [k_L]$ on the u-side such that $\mathsf{Enc}(\sigma)[\pi_e^{-1}(x)] = \mathsf{Enc}(\rho)[x]$ whenever $\pi_e(\sigma) = \rho$.

Now we are ready to describe a typical PCP test. It picks $v \in R$ randomly and generates query locations $x_1, \ldots, x_k$ for the codeword tester of the purported codeword $\mathsf{Enc}^*(v)$ along with a predicate

$$P(\mathsf{Enc}^*(v)[x_1], \ldots, \mathsf{Enc}^*(v)[x_k])$$

that would determine whether the test accepts or rejects. However this test and the query locations are only virtual. To define the actual test and the query locations, one uses the property that the encoding $\mathsf{Enc}(\rho)$ of the supposed label ρ of $v \in R$ is a sub-code of the encoding $\mathsf{Enc}(\sigma_i)$ of the supposed label σ_i of a neighbor $u_i \in L$ of v, $e_i = (u_i, v)$. Thus, one may read-off the symbol $\mathsf{Enc}^*(v)[x_i]$ from the corresponding symbol $\mathsf{Enc}^*(u_i)[y_i]$ for appropriate location $y_i = \pi_{e_i}^{-1}(x_i)$ therein. More specifically, the test picks random, independent neighbors $u_1, \ldots, u_k \in L$ of v, and tests the predicate

$$P(\mathsf{Enc}^*(u_1)[y_1], \ldots, \mathsf{Enc}^*(u_k)[y_k]).$$

This completes the description of a typical PCP test. To make this approach "work" however, more is needed. To see the difficulty involved, let's assume that the (virtual) codeword test succeeds perfectly for

[1] This paradigm is referred to as the "Inner/Outer PCP" in literature, but we avoid the usage of this terminology.

every $v \in R$, i.e. that $\mathsf{Enc}^*(v) = \mathsf{Enc}(\rho(v))$ for some label $\rho(v)$ (that depends on v). Looking at things from the perspective of some fixed $u \in L$, this amounts to saying that the purported encoding $\mathsf{Enc}^*(u)$ has, as its sub-strings, correct sub-codewords $\mathsf{Enc}(\rho(v_j))$ for all neighbors $v_j \in R$ of u. Can we conclude now that $\mathsf{Enc}^*(u)$ is also a correct codeword or at least resembles a correct codeword? Not necessarily and that's the trouble.

It is possible that the sub-codewords $\mathsf{Enc}(\rho(v_j))$ (or rather the set of their locations) constitute only a negligible portion of the purported codeword $\mathsf{Enc}^*(u)$ "on the larger side" (or rather the set of its locations). If so, the consistency of $\mathsf{Enc}^*(u)$ with all its correct sub-codewords would not say anything about correctness of $\mathsf{Enc}^*(u)$ itself. Clearly, the disparity in the encoding lengths k_L and k_R on the two sides and the number of neighbors v for a fixed $u \in L$, both have bearing on this issue. In [19], the authors defined the "sub-code covering property" that is informally stated as follows.

Definition 1.9. *(Informal) The encoding scheme $\mathsf{Enc}(\cdot)$ along with the Projection Game structure is said to achieve sub-code covering property if for every fixed $u \in L$, the "pull-back distribution" on the (query) location $y \in [k_L]$ as described next is statistically close to the uniform distribution over $[k_L]$. The pull-back distribution is defined by picking a random neighbor $v \in R$ of u, $e = (u, v)$, then picking a uniformly random location $x \in [k_R]$ and letting $y = \pi_e^{-1}(x)$.*

In [19], the authors managed to achieve the sub-code covering property using Hadamard encoding (which sufficed for the application therein). This techniques was subsequently used in the proof of the 2-to-1 Games Conjecture [16, 6] using Grassmann encoding (which again sufficed for the application therein). The Hadamard and Grassmann codes have length polynomial in the alphabet size $|\Sigma_L|$ and $|\Sigma_R|$ and while there is still a big disparity between the encoding lengths on the two sides, it is possible to arrange for a vertex $u \in L$ to have sufficiently many neighbors $v \in R$ and achieve the sub-code covering property (we omit the details). A serious restriction however is that using Hadamard and Grassmann encodings requires the projections π_e as well as the PCP test to be linear (limiting the efficacy of this approach).

Long Code and Richness

In this paper, we attempt to work with the so-called Long encoding (defined below; this is extremely important in Unique Games based reductions). As is well-said, the Long code is too long. Its length is exponential in the alphabet size, making the disparity in encoding lengths on the two sides insurmountable (as far as we foresee). Still, we attempt to identify a scenario where the sub-code property is achievable using Long codes, possibly in a more relaxed sense. Indeed, we are able to do so when $|\Sigma_L| = 2|\Sigma_R|$, the projections π_e are 2-to-1, and the game is "rich" (meaning, for a fixed $u \in L$, for its random neighbor $v \in R, e = (u, v)$, the partition of Σ_L into sets of size 2 induced by the projection π_e is uniform among all possible such partitions). We informally state this observation below.

Lemma 1.10. *(Informal) The Long code along with the Rich 2-to-1 Game structure achieves a relaxed sub-code covering property in the following sense. For every fixed $u \in L$, the "pull-back distribution" on the (query) location $y \in [k_L]$ as described in Definition 1.9 has the property that for most locations $y \in [k_L]$, their probability under the pull-back distribution is not much larger than their probability under uniform distribution on $[k_L]$.*

Formally, let $\Sigma_L = [2n] = \{1, \ldots, 2n\}$ and $\Sigma_R = [n] = \{1, \ldots, n\}$. Fix a vertex $u \in L$ in a Rich 2-to-1 Game and consider its randomly chosen neighbor $v \in R$. Then, by definition of richness, $\pi = \pi_{(u,v)} : [2n] \to [n]$ is a uniformly random 2-to-1 map.

The m-ary Long code for the label of u corresponds to a function $F : [m]^{2n} \to [m]$ and the codeword for the label $i_0 \in [2n]$ corresponds to the i_0^{th} dictatorship function

$$\mathsf{Dict}_{i_0}(z) = \mathsf{Dict}_{i_0}(z_1, \ldots, z_{2n}) = z_{i_0}.$$

Similarly, the Long code for the label of v corresponds to a function $G : [m]^n \to [m]$ and the codeword for the label $j_0 \in [n]$ corresponds to the j_0^{th} dictatorship function

$$\mathsf{Dict}_{j_0}(x) = \mathsf{Dict}_{j_0}(x_1, \ldots, x_n) = x_{j_0}.$$

We observe that if one defines for $x \in [m]^n$, $\pi^{-1}(x) \in [m]^{2n}$ by letting $\pi^{-1}(x)_i = x_{\pi(i)}$ for all $i \in [2n]$, it indeed holds that

$$\mathsf{Dict}_{i_0}(\pi^{-1}(x)) = \mathsf{Dict}_{j_0}(x) \quad \text{whenever } \pi(i_0) = j_0.$$

In this sense, the encoding corresponding to v is a sub-code of the encoding corresponding to u. A location z from the pull-back distribution on $[m]^{2n}$ is sampled by first picking a uniformly random 2-to-1 map $\pi : [2n] \to [n]$, picking $x \in [m]^n$ uniformly, and letting $z = \pi^{-1}(x)$. Clearly, this distribution is supported only on $z \in [m]^{2n}$ for which each $s \in [m]$ appears an even number of times as its co-ordinate, and hence is statistically far from the uniform distribution on $[m]^{2n}$. On the other hand, we show that for "typical" $z \in [m]^{2n}$ (those for which all $s \in [m]$ occur roughly equal number of times as its coordinate), its probability under the pull-back distribution is at most a constant times its probability under the uniform distribution. We refer the reader to Lemma 4.2 for a formal statement.

We have explained how the notion of richness is tailor-made to achieve the sub-code covering property for the Long code (albeit in a more relaxed sense). We now describe two motivations to study Rich 2-to-1 Games. Our comments are speculative, but we hope that these lead to fruitful research directions.

Hardness Results with Perfect Completeness?

From the discussion so far, it is evident that Rich 2-to-1 Games could be an excellent problem to reduce from. In particular, we show that it can be reduced to the Unique Games problem, and is equivalent to the latter. In light of this equivalence, why not just stick to the Unique Games Conjecture then? The additional advantage of using the Rich 2-to-1 Games Conjecture could be that this conjecture could hold even with perfect completeness. This could be useful towards proving hardness of approximation results where perfect completeness is essential. We cite couple of plausible candidates where hardness results could follow from the Rich 2-to-1 Games Conjecture with perfect completeness:

- Hardness of coloring 3-colorable graphs with a constant number of colors.
- Hardness of CSPs (constraint satisfaction problems) on satisfiable instances. A concrete example is the query-efficient dictatorship test with perfect completeness that is proposed and analyzed in [21, 4]. Therein, one does not know how to translate the dictatorship test to a hardness result, lacking a suitable, conjectured hard problem to reduce from.

We remark that such results could follow by developing the appropriate analytic machinery on specialized domains (minor adjustments to the reduction are needed). A concrete example (related to the problem of proving hardness of coloring 3-colorable graphs with a constant number of colors) is the multi-slice with an appropriate noise operator. Namely, $V = \left\{ x \in \{0,1,2\}^{6n} \;\middle|\; \text{the number of 0's, 1's and 2's in } x \text{ is } 2n \right\}$, with the noise operator T that acts on V in the following way: given x, randomly change half of the 0-valued

coordinates in x to 1's, and the rest into 2's, and similarly for the 1-valued and 2-valued coordinates. This operator can naturally be viewed as an averaging operator over functions, and one would need a "Majority is Stablest" type bound: if all of the low-degree influences of $f\colon V \to \{0,1\}$ are small, then $\langle f, Tf\rangle$ is bounded away from 0.

More ambitiously, one could hope that by developing the necessary analytical tools on such non-classical domains, any dictatorship test with perfect completeness could used to prove an NP-hardness result for the corresponding predicate, assuming Conjecture 1.7. We leave further investigation along this direction to future works.

Making Games Richer?

One might argue that since 2-to-1 Games are now known to be hard, we should now work towards showing that "rich" 2-to-1 Games are hard as well, showing in turn that the Unique Games are hard. It might be possible to consider "degree of richness" and design a sequence of reductions that successively achieve higher degree of richness, finally achieving full richness as in the definition of Rich 2-to-1 Games.

Formally, let $\mathcal{F}$ be a family of partitions of $[2n]$ into sets of size 2 each. A 2-to-1 Game is called $\mathcal{F}$-rich if for every fixed vertex $u \in L$, for its random neighbor $v \in R$, the partition of $[2n]$ induced by the projection $\pi = \pi_{(u,v)}$ is uniform over the family $\mathcal{F}$. We defined the game to be rich if it is $\mathcal{F}_{\mathsf{all}}$-rich, where $\mathcal{F}_{\mathsf{all}}$ is the family of all such partitions possible.

As is the case in the proof of the 2-to-1 Games Conjecture [16, 6], the 2-to-1 Games shown to be hard therein are $\mathcal{F}_{\mathsf{lin}}$-rich. Here $[2n]$ is identified with the additive group $GF(2)^k$ and $\mathcal{F}_{\mathsf{lin}}$ consists of one partition for every $b \in GF(2)^k, b \neq 0$ that induces the "linear pairing" $(x, x+b)$ for all $x \in GF(2)^k$. We float the idea to define a sequence of families

$$\mathcal{F}_0 = \mathcal{F}_{\mathsf{lin}} \subseteq \mathcal{F}_1 \ldots \subseteq \mathcal{F}_T = \mathcal{F}_{\mathsf{all}},$$

and design a sequence of reductions achieving $\mathcal{F}_j$-richness successively from $j = 0$ (which we now know) to $j = T$ (proving the Rich 2-to-1 Games Conjecture and hence the Unique Games Conjecture).

2 Preliminaries

Notation: We denote by $[n]$ the set $\{1, \ldots, n\}$ and by $[n]_d$ the set of ordered d-tuples of elements of $[n]$ consisting of distinct elements. The set of all permutations on $[n]$ is denoted by S_n and the set of all 2-to-1 mappings $\pi : [2n] \to [n]$ is denoted by $S_{2n,n}$.

We consider functions $f\colon [m]^n \to \mathbb{R}$. The distribution on $[m]^n$ is, by default, uniform (but we will have occasions to consider non-uniform distributions and if so, it will be clear from the context). A sample $x \in [m]^n$ will, by default, denote a uniform sample. For $p \geqslant 1$, the p-norm is defined in the standard manner, $\|f\|_p = \mathbb{E}_x\left[|f(x)|^p\right]^{1/p}$. The inner product of two functions is $\langle f, g\rangle = \mathbb{E}_x\left[f(x)g(x)\right]$.

Throughout the paper, $C(m), C(K,m), C(d,K,m)$ etc will denote a constant that depends on the respective parameters and this constant could change from time to time.

2.1 Basic Analytic Notions

We recall the standard way to express $f\colon [m]^n \to \mathbb{R}$ in the Fourier basis. Here, it is more convenient to define $[m] = \mathbb{Z}_m = \{0, 1, \ldots, m-1\}$ with the additive group structure. Let $\{Y_s : [m] \to \mathbb{R} \mid s \in [m]\}$

be an orthonormal set of random variables with $Y_0 \equiv 1$. One can then express a function $f\colon [m]^n \to \mathbb{R}$ uniquely as a "multi-linear" polynomial in random variables $\{X_{i,s} | 1 \leqslant i \leqslant n, s \in [m]\}$ where for each $1 \leqslant i \leqslant n$, the $\{X_{i,s}\}$ are copies of $\{Y_s\}$, and are independent for different i. A degree-d "monomial" looks like $\prod_{j=1}^{d} X_{i_j,s_j}$ with $s_j \neq 0$ and i_j distinct for $1 \leqslant j \leqslant d$. The degree of a polynomial is the maximum degree of its non-zero monomials.

Theorem 2.1. *(Hypercontractivity) Let $f\colon [m]^n \to \mathbb{R}$ be a function of degree at most d. Then for all $p \geqslant 2$,*

$$\|f\|_p \leqslant \sqrt{m(p-1)}^d \|f\|_2.$$

Definition 2.2. *The noise operator $T_{1-\varepsilon}$ acts on functions $f : [m]^n \to \mathbb{R}$ by defining*

$$T_{1-\varepsilon} f\,(x) = \mathop{\mathbb{E}}_{z \sim_{1-\varepsilon} x} [f(z)].$$

Here $z \sim_{1-\varepsilon} x$ denotes a random input z that is $(1-\varepsilon)$-correlated with x, i.e. independently for each coordinate $1 \leqslant i \leqslant n$, the i^{th} coordinate of z equals the i^{th} coordinate of x with probability $1-\varepsilon$ and is sampled uniformly from $[m]$ with probability ε.

Definition 2.3. *The influence of a coordinate $i \in [n]$ on a function $f\colon [m]^n \to \mathbb{R}$ is defined by (e_i denotes an input that is 1 in the i^{th} coordinate and zero otherwise)*

$$I_i[f] = \mathop{\mathbb{E}}_{x} \left[\left(f(x) - \mathop{\mathbb{E}}_{s \in [m]} [f(x + s e_i)] \right)^2 \right].$$

Lemma 2.4. *Let $f\colon [m]^n \to \mathbb{R}$ be a function and $i \in [n]$.*

$$\frac{1}{4} \mathop{\mathbb{E}}_{x, s \in [m]} \left[(f(x) - f(x + s e_i))^2 \right] \leqslant I_i[f] \leqslant \mathop{\mathbb{E}}_{x, s \in [m]} \left[(f(x) - f(x + s e_i))^2 \right].$$

Proof. Clearly, we have that $f(x) - \mathbb{E}_{s \in [m]} [f(x + s e_i)] = \mathbb{E}_{s \in [m]} [f(x) - f(x + s e_i)]$ and the right hand side follows. For the left hand side, write $x = (y, x_i)$ where y consists of the part of x except the i^{th} coordinate. Note that $\mathbb{E}_{s \in [m]} [f(x + s e_i)]$ only depends on y and denote it by $A(y)$. Then

$$\begin{aligned} I_i[f] = \mathop{\mathbb{E}}_{(y, x_i)} \left[(f(y, x_i) - A(y))^2 \right] &= \frac{1}{2} \mathop{\mathbb{E}}_{y, x_i, x_i'} \left[(f(y, x_i) - A(y))^2 + (A(y) - f(y, x_i'))^2 \right] \\ &\geqslant \frac{1}{4} \mathop{\mathbb{E}}_{y, x_i, x_i'} \left[(f(y, x_i) - f(y, x_i'))^2 \right], \end{aligned}$$

where we used $u^2 + v^2 \geqslant \frac{1}{2}(u+v)^2$. The last expectation is same as $\mathbb{E}_{x, s \in [m]} \left[(f(x) - f(x + s e_i))^2 \right]$ by letting $s = x_i - x_i'$ and we are done. □

Definition 2.5. *Let $f^{\leqslant d}$ and $f^{>d}$ denote the parts of f with degree at most d and larger than d respectively. The degree-d influence of a variable $i \in [n]$ on f is defined as $I_i^{\leqslant d}[f] = I_i[f^{\leqslant d}]$.*

We need the following noise-stability result of [7]. It upper-bounds the noise-stability of functions all of whose influences are low.

Theorem 2.6. *For every integer $m \geqslant 2$ and constants $\varepsilon, \theta > 0$, there is a sufficiently small constant $\delta > 0$, such that the following holds. Let $f : [m]^n \to [0,1]$ with $\mathbb{E}[f] \leqslant \theta$ and assume that for all $i \in [n]$, $I_i[f] \leqslant \delta$. Then*

$$\langle f, T_{1-\varepsilon} f \rangle \leqslant 2\Gamma_{1-\varepsilon}(\theta).$$

The function $\Gamma_{1-\varepsilon}(\theta)$ is defined in [14] and the only property we need is that for a fixed $\varepsilon > 0$, $\frac{\Gamma_{1-\varepsilon}(\theta)}{\theta} \to 0$ as $\theta \to 0$. A known upper bound is $\Gamma_{1-\varepsilon}(\theta) \leqslant C(\varepsilon)\, \theta^{2/(2-\varepsilon)}$.

Functions with range $[m]$: We also consider functions $F\colon [m]^n \to [m]$, which are more convenient to view as $F\colon [m]^n \to \Delta_m$ where Δ_m is the standard m-dimensional simplex, $\Delta_m = \{(t_0, \ldots, t_{m-1}) | \forall i\ t_i \geqslant 0,\ \sum_{i=0}^{m-1} t_i = 1\}$. The value $s \in [m]$ is then identified with the vertex $e_s \in \Delta_m$ of the simplex. Usually, we consider the function $F\colon [m]^n \to \Delta_m$ as a vector of $[0,1]$-valued functions $(F_0, F_1, \ldots, F_{m-1})$.

2.2 Hypercontractivity on the Symmetric Group and the 2-to-1 Mappings Domain

In this section, we give the basic background towards analyzing functions on the symmetric group and state the hypercontractive result we need. We consider functions $F\colon S_n \to \mathbb{R}$. For $S, T \in [n]_k$, $S = (i_1, \ldots, i_k)$, $T = (j_1, \ldots, j_k)$, we write $\pi(S) = T$ if $\pi(i_1) = j_1, \ldots, \pi(i_k) = j_k$. Let $1_{\pi(S)=T}$ be the indicator function on S_n indicating that $\pi(S) = T$.

Definition 2.7. *For $d = 0, \ldots, n$, let $V_d(S_n)$ be the linear subspace spanned by all functions*

$$\left\{ 1_{\pi(S)=T} \,\middle|\, S, T \in [n]_k,\ 0 \leqslant k \leqslant d \right\}.$$

We say that the "degree" of F is (at most) d if $F \in V_d(S_n)$.

Definition 2.8. *A degree-d function $F\colon S_n \to \mathbb{R}$ is called ε-pseudo-random if for any $S, T \in [n]_d$,*

$$\mathop{\mathbb{E}}_{\pi:\pi(S)=T} \left[F[\pi]^2 \right] \leqslant \varepsilon.$$

We need the following hypercontractive inequality from [10]. We will use it to show certain concentration properties of functions on the symmetric group (or more precisely on the 2-to-1 mappings domain defined next).

Theorem 2.9. *Let $F\colon S_n \to \mathbb{R}$ be a degree-d, ε-pseudo-random function. Then ($C(d) = 2^{40d^2}$ suffices)*

$$\mathop{\mathbb{E}}_{\pi} \left[F[\pi]^4 \right] \leqslant C(d)\varepsilon^2.$$

What we really need to analyze are functions on the 2-to-1 mappings domain $S_{2n,n}$, i.e. the set of 2-to-1 mappings $\pi\colon [2n] \to [n]$. We define the notion of degree of a function $F\colon S_{2n,n} \to \mathbb{R}$ in a similar manner. For $S \in [2n]_{2k}, T \in [n]_k$, $S = (i_1, i_1', \ldots, i_k, i_k')$, $T = (j_1, \ldots, j_k)$, we write $\pi(S) = T$ if $\pi(i_1) = \pi(i_1') = j_1, \ldots, \pi(i_k) = \pi(i_k') = j_k$. Let $1_{\pi(S)=T}$ be the indicator function on $S_{2n,n}$ indicating that $\pi(S) = T$.

Definition 2.10. *For $d = 0, \ldots, n$, let $V_d(S_n)$ be the linear subspace spanned by all functions*

$$\left\{ 1_{\pi(S)=T} \,\middle|\, S \in [2n]_{2k}, T \in [n]_k,\ 0 \leqslant k \leqslant d \right\}.$$

We say that the "degree" of F is (at most) d if $F \in V_d(S_n)$.

Definition 2.11. *A degree-d function $F\colon S_{2n,n} \to \mathbb{R}$ is called ε-pseudo-random if for any $S \in [2n]_{4d}$ and $T \in [n]_{2d}$ we have*

$$\mathop{\mathbb{E}}_{\pi:\pi(S)=T}\left[F[\pi]^2\right] \leqslant \varepsilon.$$

Remark 2.12. *Some comments regarding the above Definition 2.11. (1) For a degree-d function, we require the pseudo-randomness condition to hold for $|S| = 4d, |T| = 2d$ (unlike Definition 2.8). This is just to make sure that the proof of the next theorem works by reducing it to the case of the symmetric group. (2) The pseudo-randomness condition automatically implies the same conclusion whenever $|S| = 2k, |T| = k$, $k \leqslant 2d$. This is by randomly extending these tuples to $|S'| = 4d, |T'| = 2d$ and then averaging. (3) The pseudo-randomness condition automatically implies the same conclusion whenever the expectation is taken over $\pi \in S_{2n,n}$ such that $\pi(i_1) = j_1, \ldots, \pi(i_k) = j_k$ for $k \leqslant 2d$, $(i_1, \ldots, i_k) \in [2n]_k$, but $(j_1, \ldots, j_k)$ is allowed to be a multi-set of elements from $[n]$, each element occurring at most twice. This is by randomly filling in "mates" for $i_1, \ldots, i_k$ if missing and then averaging.*

Theorem 2.13. *Let $F\colon S_{2n,n} \to \mathbb{R}$ be a degree-d, ε-pseudo-random function. Then ($C(d) = 2^{160d^2}$ suffices)*

$$\mathop{\mathbb{E}}_{\pi}\left[F[\pi]^4\right] \leqslant C(d)\varepsilon^2.$$

Proof. We reduce to the case of the symmetric group by embedding S_{2n} into $S_{2n,n}$, mapping $\tilde{\pi} \to \pi$ in the following way. Given $\tilde{\pi} \in S_{2n}$, we define a 2-to-1 mapping π by $\pi(\tilde{\pi}(2j-1)) = \pi(\tilde{\pi}(2j)) = j$ for $j \in [n]$. This mapping is onto and is 2^n-to-1. In particular, sampling $\tilde{\pi} \in S_{2n}$ uniformly, the distribution of π is uniform over $S_{2n,n}$.

Let F be a degree-d function as in the statement of the theorem. Define $G\colon S_{2n} \to \mathbb{R}$ by $G(\tilde{\pi}) = F(\pi)$. We say that G is a "lifting" of F. We claim that G is a degree-$2d$ function. To see this, we simply observe that for a degree-k "monomial" function $1_{\pi(S)=T}$ on $S_{2n,n}$ with $S = (i_1, i_1', \ldots, i_k, i_k') \in [2n]_{2k}$, $T = (j_1, \ldots, j_k) \in [n]_k$, its lifting is exactly the sum of 2^k "monomial" functions on S_{2n}

$$\sum_{\tilde{S}} 1_{\tilde{\pi}(\tilde{S})=\tilde{T}}.$$

Here $\tilde{T} = (2j_1 - 1, 2j_1, \ldots, 2j_k - 1, 2j_k)$ and $\tilde{S} = (\{i_1, i_1'\}, \ldots, \{i_k, i_k'\})$, each ordering of $\{i_\ell, i_\ell'\}$ within the pair giving one possible $\tilde{S}$, hence 2^k possible $\tilde{S}$.

Finally, we claim that since F is ε-pseudo-random, so is G. Note that G is a degree-$2d$ function. For any $\tilde{S} = (i_1, \ldots, i_k) \in [2n]_k$ and $\tilde{T} = (j_1, \ldots, j_k) \in [2n]_k$, $k \leqslant 2d$, uniformly sampling $\tilde{\pi}$ such that $\tilde{\pi}(\tilde{S}) = \tilde{T}$ leads to uniformly sampling π such that $\pi(i_1) = \lceil \frac{j_1}{2} \rceil, \ldots, \pi(i_k) = \lceil \frac{j_k}{2} \rceil$. The "mates" for $i_1, \ldots, i_k$, if missing, can be randomly filled in and then the pseudo-randomness condition for G follows from that of F. Now one appeals to Theorem 2.9 and concludes by noting that $\mathbb{E}_\pi\left[F[\pi]^4\right] = \mathbb{E}_{\tilde{\pi}}\left[G[\tilde{\pi}]^4\right]$. □

3 Main Analytic Lemma

We now state our main analytic lemma. Let $\pi \in S_{2n,n}$ be a 2-to-1 map. We recall that for $x \in [m]^n$, its "pull-back" $\pi^{-1}(x) \in [m]^{2n}$ is defined as $\pi^{-1}(x)_i = x_{\pi(i)}$ for $i \in [2n]$. For a function $f : [m]^{2n} \to \mathbb{R}$, the "restriction" $f|_\pi : [m]^n \to \mathbb{R}$ is defined as (this is indeed restriction of f to the pull-back domain $\pi^{-1}([m]^n)$)

$$f|_\pi(x) = f(\pi^{-1}(x)).$$

Our main lemma states, loosely speaking, that if f is a low-degree, bounded function and if $\pi \in S_{2n,n}$ is a random 2-to-1 map, then the influential co-ordinates of f and those of the restricted function $f|_\pi$ are related. More specifically, it is unlikely to happen that $f|_\pi$ has some influential co-ordinate j without either of $i, i' \in \pi^{-1}(j)$ being influential for f.

Lemma 3.1. *Fix the alphabet size $m \geqslant 2$. For every constants $\delta, \zeta > 0$ and integer $d \geqslant 1$, there are sufficiently small constants $\gamma = \gamma(m, \delta, \zeta), \tau = \tau(d, m, \delta, \zeta) > 0$ such that the following holds. Suppose $f\colon [m]^{2n} \to [0,1]$ is a function such that $\|f^{>d}\|_2^2 \leqslant \gamma$. Then*

$$\Pr_{\pi}\left[\exists\, j \in [n] :\ I_j[f|_\pi] \geqslant \delta \quad \wedge \quad \max_{i \in \pi^{-1}(j)} I_i^{\leqslant d}[f] \leqslant \tau\right] \leqslant \zeta.$$

We first prove a very similar lemma stated below. It considers the special case when f itself has no degree-d influential variables at all. Its proof contains all the main ingredients and the above Lemma 3.1 is then proved with some minor modifications.

Lemma 3.2. *Fix the alphabet size $m \geqslant 2$. For every constants $\delta, \zeta > 0$ and integer $d \geqslant 1$, there are sufficiently small constants $\gamma = \gamma(m, \delta, \zeta), \tau = \tau(d, m, \delta, \zeta) > 0$ such that the following holds. Suppose $f\colon [m]^{2n} \to [0,1]$ is a function such that $\|f^{>d}\|_2^2 \leqslant \gamma$ and moreover that for all $i \in [2n]$, $I_i^{\leqslant d}[f] \leqslant \tau$. Then,*

$$\Pr_{\pi}\left[\exists\, j \in [n] :\ I_j[f|_\pi] \geqslant \delta\right] \leqslant \zeta.$$

Remark 3.3. *It is important that in the statements of the lemmas above, γ does not depend on d. When we apply these lemmas, f itself will be a smoothed version $T_{1-\varepsilon}h$ for some $[0,1]$-valued function h. Thus $\|f^{>d}\|_2^2 \leqslant \gamma = 2^{-\Omega(d/\varepsilon)}$ and in fact d will be chosen sufficiently large so as to make γ sufficiently small (so the dependence "in practice" is really the other way round).*

4 Proof of the Main Analytic Lemma

In this section, we prove Lemma 3.2 (and the proof of Lemma 3.1 follows by minor modifications). We will work, for the large part, with function g that is, roughly speaking, $f^{\leqslant d}$. However, for technical reasons, we will zero-out its values on a small set of "atypical" inputs that are outside a certain set $E \subseteq [m]^{2n}$. Formally, $g = f^{\leqslant d} 1_E$ where 1_E is the indicator of set E. Towards the end of the proof, we will relate influences of f and g. Motivation and overview of successive steps in the proof is presented as we go along.

4.1 The Pull-back Distribution

While trying to relate influences of a function $g : [m]^{2n} \to \mathbb{R}$ to those of its restrictions $g|_\pi$, a technical hurdle is that the "pull-back distribution" on $[m]^{2n}$ that we define next differs from the uniform distribution on $[m]^{2n}$. The pull-back distribution arises while considering the average of influences of $g|_\pi$ over the choice of π whereas the influences of g itself are defined with respect to the uniform distribution. We are able to show that the pull-back distribution resembles the uniform distribution on $[m]^{2n}$ in a loose, but controlled manner.

Definition 4.1. *The pull-back distribution $\nu_{2n,m}$ over $[m]^{2n}$ is defined by the following process: sample $\pi \in S_{2n,n}$, $x \in [m]^n$ and output $z = \pi^{-1}(x)$.*

Clearly, this distribution is supported only on $z \in [m]^{2n}$ for which each $s \in [m]$ appears an even number of times as its coordinate, and hence is statistically far from the uniform distribution on $[m]^{2n}$. On the other hand, we show that for "typical" $z \in [m]^{2n}$, its probability under the distribution $\nu_{2n,m}$ is at most a constant times its probability under the uniform distribution (this and an additional related fact is all we need).

Lemma 4.2. *A point $z \in [m]^{2n}$ is called K-roughly balanced if every value $s \in [m]$ appears in $\frac{2n}{m} \pm \sqrt{K \log m \, \frac{n}{m}}$ of the coordinates of z. For a K-roughly balanced point z,*

$$\nu_{2n,m}(z) \leqslant C(K,m)\, m^{-2n}.$$

Proof. Let A_s be the set of coordinates of z that are equal to $s \in [m]$, let $a_s = |A_s|$, and let $a_s = \frac{2n+v_s}{m}$. We may assume that all sets A_s are even-sized, since otherwise $\nu_{n,m}(z) = 0$. In this case, $\nu_{n,m}(z)$ is equal to m^{-n} times the probability that for a random $\pi \in S_{2n,n}$, z happens to be in the range of π^{-1}, or equivalently that π matches off each set A_s within itself. By Lemma A.5, this probability is

$$\frac{\binom{n}{\frac{a_0}{2},\ldots,\frac{a_{m-1}}{2}}}{\binom{2n}{a_0,\ldots,a_{m-1}}}.$$

Since z is K-roughly balanced, we have that $|v_s| \leqslant \sqrt{(K \log m)\, m\, n}$. Using Lemma A.4, the ratio between the two multinomial coefficients is at most $C(m)\, 2^{K \log m \cdot m} m^{-n} = C(K,m) m^{-n}$. □

Remark 4.3. *We will often use the lemma above with additional conditioning on the choice of π, say for example that π is sampled uniformly with the condition $\pi(2n-1) = \pi(2n) = n$. The lemma continues to hold. The distribution $\tilde{v}_{2n,m}$ on inputs $z \in [m]^{2n}$ is now supported on z where every $s \in [m]$ occurs an even number of times as its coordinate and moreover that $z_{2n-1} = z_{2n}$. Writing $z = (\tilde{z}, z_{2n-1}, z_{2n})$, if z is K-roughly balanced, then $\tilde{z} \in [m]^{2n-2}$ is $(K+1)$-roughly balanced and the probability that $\tilde{z}$ is output is $C(K,m)$ times its probability under uniform distribution.*

The lemma above immediately implies the following. It is then used to relate influences of $g : [m]^{2n} \to \mathbb{R}$ to those of $g|_\pi$ (the latter in expectation).

Lemma 4.4. *Let $h \colon [m]^{2n} \to [0,\infty)$ be a function supported only on K-roughly balanced inputs. Then*

$$\mathop{\mathbb{E}}_{z \sim \nu_{2n,m}}[h(z)] \leqslant C(K,m) \mathop{\mathbb{E}}_{z \in_R [m]^{2n}}[h(z)].$$

We now show how this is useful. Let $g : [m]^{2n} \to \mathbb{R}$ and consider a random choice of $\pi \in S_{2n,n}$ such that $\pi(2n-1) = \pi(2n) = n$. Such π can be chosen at random by first choosing $\pi' \in S_{2(n-1),(n-1)}$ at random, letting $\pi = \pi'$ on $[2(n-1)]$, and then extending by letting $\pi(2n-1) = \pi(2n) = n$. We wish to consider the expected influence of the n^{th} coordinate on the restriction $g|_\pi$.

Remark 4.5. *Here we specifically consider the n^{th} coordinate of $g|_\pi$ under the requirement $\pi(2n-1) = \pi(2n) = n$. This is for notational convenience only and is without loss of generality. The same results hold for any given j^{th} coordinate of $g|_\pi$ under the requirement that $\pi(i) = \pi(i') = j$ for any given $i \neq i' \in [2n]$.*

Lemma 4.6. *Let $g : [m]^{2n} \to \mathbb{R}$ be a function supported only on K-roughly balanced inputs. Then*

$$\mathop{\mathbb{E}}_{\pi}[I_n[g|_\pi]] \leqslant C(K,m)\left(I_{2n-1}[g] + I_{2n}[g]\right).$$

Proof. Let e_{2n} be the input with the $(2n)^{th}$ coordinate 1 and all other coordinates zero. Let e_{2n-1} be similarly defined and let $e = e_{2n-1} + e_{2n}$. By Lemma 2.4,

$$\mathop{\mathbb{E}}_{\pi}\left[I_n[g|_\pi]\right] \leqslant \mathop{\mathbb{E}}_{\substack{\pi, x\in[m]^n \\ s\in[m]}}\left[\Big(g(\pi^{-1}(x)) - g(\pi^{-1}(x) + s\ e)\Big)^2\right].$$

Let $z = \pi^{-1}(x)$ so that z is distributed according to the distribution $\tilde{\nu}_{2n,m}$ (see Remark 4.3). Since g is supported only on K-roughly balanced inputs, the term above is non-zero only if z is $(K+1)$-roughly balanced. Hence by Lemma 4.4, the above expectation is at most

$$C(K,m) \cdot \mathop{\mathbb{E}}_{z\in_R[m]^{2n},\ s}\left[\Big(g(z) - g(z + s\ e)\Big)^2\right].$$

Note that we think of $z \in_R [m]^{2n}$ as uniformly distributed now onwards. Using $(a-b)^2 \leqslant 2(a-c)^2 + 2(c-b)^2$, the last expectation is at most twice

$$\mathop{\mathbb{E}}_{z,s}\left[\Big(g(z) - g(z + s\ e_{2n-1})\Big)^2\right] + \mathop{\mathbb{E}}_{z,s}\left[\Big(g(z + s\ e_{2n-1}) - g(z + s\ e_{2n-1} + s\ e_{2n}))\Big)^2\right].$$

Since the distribution of $z \in [m]^{2n}$ is uniform, so is the distribution of $z + s\ e_{2n-1}$ and hence these expectations are equal (up to a factor 4) to $I_{2n-1}[g]$ and $I_{2n}[g]$ respectively. □

4.2 The Function G on $S_{2n,n}$ and its Pseudo-randomness

We seek to show that under appropriate conditions, if a function $g : [m]^{2n} \to \mathbb{R}$ has all influences low, then with high probability over the choice of π, the same is true for the restriction $g|_\pi$. We begin by a (somewhat imprecise) proof-sketch.

Suppose that g has all influences low, say at most τ. By above Lemma 4.6, the expected value of the influence $I_n[g|_\pi]$, over the choice of π, is at most $O(\tau)$. We would like to show that in fact $I_n[g|_\pi]$ is at most $O(\tau)$ with high probability over the choice of π. We would then argue that the same holds for the influence $I_j[g|_\pi]$ for every $1 \leqslant j \leqslant n$ (since consideration of the n^{th} coordinate was just for notational convenience), then take a union bound over all $1 \leqslant j \leqslant n$, and conclude that the restriction $g|_\pi$ has all influences low.

However, such an argument requires strong probabilistic guarantees. It is natural to seek an upper bound on the higher moments of the random variable $G[\pi] = I_n[g|_\pi]$. We are able to do this, but only in a rather convoluted manner. We show that $G[\pi]$ is pseudo-random as a function on $S_{2n,n}$ (or strictly speaking, on $S_{2(n-1),n-1}$ since $\pi(2n-1) = \pi(2n) = n$ is pre-defined) in the sense of Definition 2.11. Concretely, we show that for small d and any sets $|A| = 2(d-1), |B| = d-1$, the conditional second moment $\mathbb{E}_{\pi(A)=B}\left[G[\pi]^2\right]$ remains bounded by $O(1)$ times (the unconditional second moment) $\mathbb{E}\left[G[\pi]^2\right]$. For notational convenience (only), one can think of

$$A = \{2(n-(d-1))-1, 2(n-(d-1)), \ldots, 2(n-1)-1, 2(n-1)\}, \quad B = \{n-(d-1), \ldots, n-1\},$$

and the event $\pi(A) = B$ denotes the event that $\pi(2(n-j)-1) = \pi(2(n-j)) = n-j$ for $1 \leqslant j \leqslant d-1$ (and in addition, $\pi(2n-1) = \pi(2n) = n$ is pre-defined, corresponding to $j = 0$).

This pseudo-randomness property then implies that the fourth moment $\mathbb{E}\left[G[\pi]^4\right]$ is upper bounded by $O(1)$ times (the square of the second moment) $\mathbb{E}\left[G[\pi]^2\right]^2$. This gives sufficiently strong guarantees to make the "with high probability" and union bound arguments to go through. Towards implementing the details of this proof, we need the following ad hoc sounding lemma. We then show how to use it and prove the desired pseudo-randomness property.

Lemma 4.7. *Let a pair of inputs $z_1, z_2 \in [m]^{2n}$ be chosen by two different methods:*

- *Choose a random $\pi \in S_{2n,n}$, then choose $x_1, x_2 \in [m]^n$ at random, and then define $z_i = \pi^{-1}(x_i)$. Let $\mu(z_1, z_2)$ denote the probability that the pair (z_1, z_2) is output.*
- *Let*

$$A = \{2(n-(d-1))-1, 2(n-(d-1)), \ldots, 2n-1, 2n\}, \quad B = \{n-(d-1), \ldots, n\},$$

and the event $\pi(A) = B$ denotes the event that $\pi(2(n-j)-1) = \pi(2(n-j)) = n-j$ for $0 \leqslant j \leqslant d-1$. Let $\mu_{\mathsf{cond}}(z_1, z_2)$ denote the probability that the pair (z_1, z_2) is output by the method above, but conditional on the event $\pi(A) = B$.

Then if the pair (z_1, z_2) is "typical", we have

$$\mu_{\mathsf{cond}}(z_1, z_2) \leqslant C(d, m)\, \mu(z_1, z_2),$$

where the pair (z_1, z_2) is "typical" if among the multi-set $\{(z_1(i), z_2(i)) | 1 \leqslant i \leqslant 2n\}$ of their coordinates, each of the m^2 patterns in $[m] \times [m]$ appears at least $\frac{2n}{20m^2}$ times.

Proof. Among the multi-set $\{(z_1(i), z_2(i)) | 1 \leqslant i \leqslant 2n\}$, let the number of occurrences of the m^2 possible patterns be $v_1, \ldots, v_{m^2}$. We may assume that these numbers are all even since otherwise the pair (z_1, z_2) will never be output. The probability $\mu(z_1, z_2)$ is equal to (using Lemma A.5)

$$\frac{\binom{n}{\frac{v_1}{2}, \ldots, \frac{v_{m^2}}{2}}}{\binom{2n}{v_1, \ldots, v_{m^2}}}. \tag{1}$$

Denote by $u_1, \ldots, u_{m^2}$ the number of occurrences of these patterns that appear in the $2d$ coordinates of A so that $2d = u_1 + \ldots + u_{m^2}$. The probability $\mu_{\mathsf{cond}}(z_1, z_2)$ is equal to (using Lemma A.5 again)

$$\frac{\binom{n-d}{\frac{v_1-u_1}{2}, \ldots, \frac{v_{m^2}-u_{m^2}}{2}}}{\binom{2n-2d}{v_1-u_1, \ldots, v_{m^2}-u_{m^2}}}. \tag{2}$$

Applying Lemma A.3, we see that the numerator of (2) is at most $C(d, m)$ times the numerator of (1), and its denominator is at least $c(d, m)$ times the denominator of (1) for some $c(d, m) > 0$, and hence we conclude that $\mu_{\mathsf{cond}}(z_1, z_2) \leqslant C(d, m)\mu(z_1, z_2)$ for a typical pair (z_1, z_2). □

Lemma 4.8. *Let $g : [m]^{2n} \to \mathbb{R}$ be a function supported only on K-roughly balanced inputs. Let*

$$A = \{2(n-(d-1))-1, 2(n-(d-1)), \ldots, 2(n-1)-1, 2(n-1)\}, \quad B = \{n-(d-1), \ldots, n-1\}.$$

Then the random variable $G[\pi] = I_n[g|_\pi]$ satisfies ($\pi(2n-1) = \pi(2n) = n$ is pre-defined)

$$\mathop{\mathbb{E}}_{\pi(A)=B}\left[G[\pi]^2\right] \;\leqslant\; C(d, m)\, \mathbb{E}\left[G[\pi]^2\right] \;+\; 2^{-\Omega(\frac{n}{m^2})} \cdot C(d, K, m) \cdot \|g\|_4^4.$$

Proof. Denote $h(z) = (g(z) - \mathbb{E}_{s \in [m]}[g(z + s\, e)])^2$ where the last two coordinates of e equal 1 and the rest are zero. By definition, $\mathbb{E}_{\pi(A)=B}\left[G[\pi]^2\right]$ equals

$$\mathop{\mathbb{E}}_{\pi(A)=B}\left[I_n[g|_\pi]^2\right] = \mathop{\mathbb{E}}_{\pi(A)=B}\left[\mathop{\mathbb{E}}_{x_1, x_2}\left[h(\pi^{-1}(x_1))h(\pi^{-1}(x_2))\right]\right] = \mathop{\mathbb{E}}_{(z_1, z_2) \sim \mu_{\mathsf{cond}}}[h(z_1)h(z_2)]. \tag{3}$$

We wish to upper bound this expression in terms of $\mathbb{E}\left[G[\pi]^2\right]$ which may be written similarly as

$$\mathbb{E}\left[G[\pi]^2\right] = \mathop{\mathbb{E}}_{(z_1,z_2)\sim\mu}[h(z_1)h(z_2)],$$

the two expectations being similar, but under different distributions μ_{cond} and μ respectively. The proof proceeds by splitting the expectation in (3) into two parts, over the pairs (z_1, z_2) that are typical versus that are atypical. For the first part, we upper bound using the above Lemma 4.7 and hence are able to "switch" to the distribution μ. We now show how to upper bound the second part; this is by using Cauchy-Schwartz carefully and noting that only a negligible number of pairs are atypical. Let 1_{Bad} denote the indicator of the event that the pair (z_1, z_2) is atypical. We note that the probability of this event is at most $2^{-\Omega(\frac{n}{m^2})}$. We wish to upper bound

$$\mathop{\mathbb{E}}_{(z_1,z_2)\sim\mu_{\mathsf{cond}}}[h(z_1)h(z_2)1_{\mathsf{Bad}}(z_1,z_2)] = \mathop{\mathbb{E}}_{(z_1,z_2)\sim\mu_{\mathsf{cond}}}[h(z_1)1_{\mathsf{Bad}}(z_1,z_2)\cdot h(z_2)1_{\mathsf{Bad}}(z_1,z_2)].$$

By Cauchy-Schwartz, this is upper bounded by

$$\mathop{\mathbb{E}}_{(z_1,z_2)\sim\mu_{\mathsf{cond}}}\left[h(z_1)^2 1_{\mathsf{Bad}}(z_1,z_2)\right]. \tag{4}$$

Since $g(z_1)$ is non-zero only on K-roughly balanced inputs z_1, the same holds for $h(z_1)$ (possibly replacing K by $K+1$; we ignore this minor point). We may thus assume that z_1 is K-roughly balanced. Provided that z_1 is K-roughly balanced, the probability that (z_1, z_2) is atypical remains $2^{-\Omega(\frac{n}{m^2})}$. We note in addition that

$$h(z_1)^2 \leqslant C(m)\cdot \mathop{\mathbb{E}}_{s\in[m]}\left[g(z_1 + s\ e)^4\right],$$

and that since z_1 is K-roughly balanced, its probability under μ_{cond} is at most $C(d, K, m)$ times that under the uniform distribution on $[m]^{2n}$ (by Lemma 4.2; the conditioning $\pi(A) = B$ may give additional factor of m^d). Putting these observations together, we upper bound (4), as desired, by

$$2^{-\Omega(\frac{n}{m^2})}\cdot C(d,K,m)\cdot \mathop{\mathbb{E}}_{z\in_R[m]^{2n}}\left[g(z)^4\right].$$

□

4.3 Using Hypercontractivity on $S_{2n,n}$

We now present the key hypercontractive argument, almost completing the proof as far as the function $g = f^{\leqslant d}1_E$ is concerned. In subsequent sections, we carry out the final steps relating influences of f and g.

Lemma 4.9. *Let $f : [m]^{2n} \to [0,1]$ be a bounded function and $E \subseteq [m]^{2n}$ be the set of K-roughly balanced inputs. Define*

$$g = f^{\leqslant d}\cdot 1_E,$$

i.e. g is the low-degree part of f, but in addition zeroed out on the imbalanced inputs. Then

$$\Pr_{\pi}\left[I_n[g|_\pi] \geqslant \delta\right] \leqslant \frac{C(d,K,m)}{\delta^4}\left(I_{2n-1}[g]^4 + I_{2n}[g]^4\right).$$

Proof. We use Lemma 4.8 to conclude that

$$\mathop{\mathbb{E}}_{\pi(A)=B}\left[G[\pi]^2\right] \leqslant C(d,m)\,\mathbb{E}\left[G[\pi]^2\right] + 2^{-\Omega(\frac{n}{m^2})}\cdot C(d,K,m)\cdot\|g\|_4^4.$$

Towards bounding the second term, we observe

$$\|g\|_4^4 = \|f^{\leqslant d}1_E\|_4^4 \leqslant \|f^{\leqslant d}\|_4^4 \leqslant (3m)^{2d}\|f^{\leqslant d}\|_2^4 \leqslant (3m)^{2d}\|f\|_2^4 \leqslant (3m)^{2d}.$$

Here we used the fact that f is a bounded function and Theorem 2.1. The $2^{-\Omega(\frac{n}{m^2})}$ factor in the second term makes the term negligible. This term does not really affect subsequent arguments, so for the clarity of presentation, we take the liberty to ignore it henceforth. Thus we have

$$\mathop{\mathbb{E}}_{\pi(A)=B}\left[G[\pi]^2\right] \leqslant C(d,m)\,\mathbb{E}\left[G[\pi]^2\right].$$

Since this holds for any $|A| = 2d-2, |B| = d-1$, the function $G[\pi]$, as a function on $S_{2(n-1),n-1}$, is pseudo-random in the sense of Definition 2.11 (we stress again that $\pi(2n-1) = \pi(2n) = n$ is pre-defined). Moreover, the degree of $G[\pi]$ is (at most) $2d$. The subtle explanation is as follows.[2] By definition, $G[\pi]$ is the average of

$$(g(\pi^{-1}(x)) - g(\pi^{-1}(x+s\,e)))^2 = g(\pi^{-1}(x))^2 + g(\pi^{-1}(x+s\,e))^2 - 2g(\pi^{-1}(x))g(\pi^{-1}(x+s\,e)) \quad (5)$$

over some distribution over x, s, so it is enough to argue about the degree for each fixed x, s. If either of the inputs $\pi^{-1}(x)$ or $\pi^{-1}(x+s\,e)$ falls outside of the set E, their g-value is zero and can be dropped from consideration. Otherwise their g-values are given by the degree-d function $f^{\leqslant d} : [m]^{2n} \to \mathbb{R}$. Thus (5) can be written as a linear combination of monomials of degree at most $2d$ and any monomial, say on coordinates $i_1, \ldots, i_{2d}$, is determined by $\pi(i_1), \ldots, \pi(i_{2d})$ when regarded as a function on $S_{2(n-1),n-1}$.

Thus $G[\pi]$ is a degree-$2d$ pseudo-random function and we can apply Lemma 2.13 to upper bound its fourth moment as

$$\mathop{\mathbb{E}}_{\pi}\left[G[\pi]^4\right] \leqslant C(d,m)\mathop{\mathbb{E}}_{\pi}\left[G(\pi)^2\right]^2.$$

Finally, by $(\frac{3}{2}, 3)$-Holder's inequality,

$$\mathop{\mathbb{E}}_{\pi}\left[G[\pi]^2\right] = \mathop{\mathbb{E}}_{\pi}\left[G[\pi]^{\frac{2}{3}}\cdot G[\pi]^{\frac{4}{3}}\right] \leqslant (\mathop{\mathbb{E}}_{\pi}[G[\pi]])^{\frac{2}{3}}(\mathop{\mathbb{E}}_{\pi}\left[G[\pi]^4\right])^{\frac{1}{3}},$$

which yields, using the bound on the fourth moment,

$$\mathop{\mathbb{E}}_{\pi}\left[G[\pi]^2\right] \leqslant C(d,m)\mathop{\mathbb{E}}_{\pi}[G[\pi]]^2,$$

and then

$$\mathop{\mathbb{E}}_{\pi}\left[G[\pi]^4\right] \leqslant C(d,m)\mathop{\mathbb{E}}_{\pi}[G[\pi]]^4.$$

Using Markov and Lemma 4.6, we conclude as desired, that

$$\Pr_{\pi}[G[\pi] \geqslant \delta] \leqslant \frac{\mathbb{E}_\pi\left[G[\pi]^4\right]}{\delta^4} \leqslant \frac{C(d,m)}{\delta^4}\mathop{\mathbb{E}}_{\pi}[G[\pi]]^4 \leqslant \frac{C(d,K,m)}{\delta^4}\,(I_{2n-1}[g]^4 + I_{2n}[g]^4).$$

□

[2] To be in strict accordance with Definition 2.11, one actually argues here that $G[\pi]$ has degree $d^* = 2d$ and the pseudo-randomness condition holds for all $|A| = 4d^*, |B| = d^*$. We have avoided this minor point for ease of presentation.

Lemma 4.10. *Let $f : [m]^{2n} \to [0,1]$ be a bounded function and $E \subseteq [m]^{2n}$ be the set of K-roughly balanced inputs. Define $g = f^{\leqslant d} \cdot 1_E$ as in the statement of Lemma 4.9. Then*

$$\Pr_{\pi}\left[\exists j : I_j[g|_\pi] \geqslant \delta\right] \leqslant \frac{C(d,K,m)}{\delta^4} \cdot \sum_{i=1}^{2n} I_i[g]^4.$$

Proof. We use Lemma 4.9, but note that the consideration of the n^{th} coordinate and the requirement that $\pi(2n-1) = \pi(2n) = n$ is only for notational convenience. What we have actually proved is that for any $1 \leqslant j \leqslant n$ and any $1 \leqslant i \neq i' \leqslant 2n$,

$$\Pr_{\pi:\pi(i)=\pi(i')=j}\left[I_j[g|_\pi] \geqslant \delta\right] \;\leqslant\; \frac{C(d,K,m)}{\delta^4}\,(I_i[g]^4 + I_{i'}[g]^4).$$

Fixing j and taking average over all $1 \leqslant i \neq i' \leqslant 2n$ gives

$$\Pr_{\pi}\left[I_j[g|_\pi] \geqslant \delta\right] \;\leqslant\; \frac{C(d,K,m)}{\delta^4}\,\mathop{\mathbb{E}}_{\pi}\left[\sum_{i\in\pi^{-1}(j)} I_i[g]^4\right].$$

Now taking a union bound over all $1 \leqslant j \leqslant n$ gives the result. □

The above Lemma 4.10 shows, morally speaking, that with high probability over the choice of π, all influences of $g|_\pi$ are low provided that all influences of g are low. We could upper bound $\sum_{i=1}^{2n} I_i[g]^4$ by $\tau(g)^3 I[g]$ where $\tau(g)$ is the maximum influence $I_i[g]$ and $I[g] = \sum_{i=1}^{2n} I_i[g]$ is the total influence. The total influence, since g is morally speaking same as $f^{\leqslant d}$, should be $O(d)$. However, the fact that $g = f^{\leqslant d}1_E$ is a truncation of $f^{\leqslant d}$ complicates matters and we have to go through a somewhat tedious argument.

4.4 Relating Influences of f and g

Lemma 4.11. *Let $f : [m]^{2n} \to [0,1]$ be a bounded function and $E \subseteq [m]^{2n}$ be the set of K-roughly balanced inputs. Define $g = f^{\leqslant d} \cdot 1_E$. Then for any coordinate $1 \leqslant i \leqslant 2n$,*

$$I_i[g] \;\leqslant\; C(m)\, I_i^{\leqslant d}[f] \;+\; C(d,m)\, n^{-\frac{3}{8}}.$$

Proof. By definition, $g = f^{\leqslant d} \cdot 1_E$ and $I_i[g]$ equals (possibly up to a factor m)

$$\mathop{\mathbb{E}}_{z\in_R[m]^{2n}}\left[|f^{\leqslant d}(z)\cdot 1_E(z) - f^{\leqslant d}(z+e_i)\cdot 1_E(z+e_i)|^2\right].$$

Now if both z and $z+e_i$ are in E, the term inside is same as $|f^{\leqslant d}(z) - f^{\leqslant d}(z+e_i)|^2$ and it contributes to the influence $I_i^{\leqslant d}[f]$. So only additional contribution to $I_i[g]$ on top of $I_i^{\leqslant d}[f]$ is due to inputs z such that $z \in E$, but $z + e_i \notin E$ (or vice versa). Let ∂E denote the set of such z so that it constitutes at most $\frac{C(m)}{\sqrt{n}}$ fraction of inputs in $[m]^{2n}$. The additional contribution to $I_i[g]$ is now upper bounded as (using $(4, \frac{4}{3})$-Holder)

$$\mathbb{E}\left[f^{\leqslant d}(z)^2 1_{\partial E}\right] \leqslant \mathbb{E}\left[f^{\leqslant d}(z)^8\right]^{\frac{1}{4}} \mathbb{E}\left[1_{\partial E}^{\frac{4}{3}}\right]^{\frac{3}{4}} \leqslant C(d,m)\left(\frac{C(m)}{\sqrt{n}}\right)^{\frac{3}{4}}.$$

We used $\mathbb{E}\left[f^{\leqslant d}(z)^8\right] \leqslant C(d,m)\mathbb{E}\left[f^{\leqslant d}(z)^2\right]^4 \leqslant C(d,m)$ that follows from Theorem 2.1. This completes the proof. □

Lemma 4.12. *Let $f : [m]^{2n} \to [0,1]$ be a bounded function and $E \subseteq [m]^{2n}$ be the set of K-roughly balanced inputs. Define $g = f^{\leqslant d} \cdot 1_E$. Then except with probability ζ over the choice of π, we have*

$$\max_{1 \leqslant j \leqslant n} I_j[f|_\pi] \;\leqslant\; 3 \cdot \max_{1 \leqslant j \leqslant n} I_j[g|_\pi] + \delta.$$

This holds as long as $K = O(\log \frac{1}{\delta\zeta})$ is sufficiently large and $\|f^{>d}\|_2^2 \leqslant \gamma = \gamma(m, \delta, \zeta)$ is sufficiently small. We emphasize that γ does not depend on d.

Proof. We write $f = g + h + q$ where $g = f^{\leqslant d} \cdot 1_E$, $h = f^{>d} \cdot 1_E$, and $q = f \cdot 1_{\overline{E}}$. Clearly, for any coordinate $1 \leqslant j \leqslant n$ (using $(a+b+c)^2 \leqslant 3(a^2+b^2+c^2)$),

$$I_j[f|_\pi] \leqslant 3 \cdot (I_j[g|_\pi] + I_j[h|_\pi] + I_j[q|_\pi]) \,.$$

We will show that except with "small" probability over the choice of π, both $\|h|_\pi\|_2^2$ and $\|q|_\pi\|_2^2$ are "small". Since these are upper bounds on $I_j[h|_\pi]$ and $I_j[q|_\pi]$ respectively, the lemma follows. We will just show that $\mathbb{E}_\pi\left[\|h|_\pi\|_2^2\right]$ and $\mathbb{E}_\pi\left[\|q|_\pi\|_2^2\right]$ are "small" (i.e. $\ll \delta\zeta$ and this determines the quantitative constraints on K and γ) and then use Markov. Indeed,

- Towards upper-bounding $\mathbb{E}_\pi\left[\|q|_\pi\|_2^2\right]$, we note that f is bounded in $[0,1]$ and

  $$\mathop{\mathbb{E}}_{\pi}\left[\|q|_\pi\|_2^2\right] = \mathop{\mathbb{E}}_{\pi, x \in [m]^n}\left[f(\pi^{-1}(x))\, 1_{\overline{E}}(\pi^{-1}(x))\right] \leqslant \mathop{\mathbb{E}}_{\pi, x \in [m]^n}\left[1_{\overline{E}}(\pi^{-1}(x))\right],$$

  and the probability that $\pi^{-1}(x)$ is imbalanced is at most $2^{-\Omega(K)}$.

- Towards upper-bounding $\mathbb{E}_\pi\left[\|h|_\pi\|_2^2\right]$, we argue that ($z = \pi^{-1}(x)$)

  $$\mathop{\mathbb{E}}_{\pi}\left[\|h|_\pi\|_2^2\right] = \mathop{\mathbb{E}}_{\pi, x \in [m]^n}\left[f^{>d}(z)^2 1_E(z)\right] \leqslant C(K,m)\, \|f^{>d}\|_2^2 \leqslant C(K,m)\, \gamma.$$

  In the second step, since one is concerned only with K-roughly balanced inputs, one can "switch" to uniform distribution over input thanks to Lemma 4.2.

□

4.5 Finishing the Proof

Proof of Lemma 3.2

We now complete the proof of Lemma 3.2. Let $f : [m]^{2n} \to [0,1]$ be a function as therein with $\|f^{>d}\|_2^2 \leqslant \gamma$ and for all $i \in [2n]$, $I_i^{\leqslant d}[f] \leqslant \tau$. Let $g = f^{\leqslant d} 1_E$ where E is the set of K-roughly balanced inputs. The parameters K, γ, τ are chosen as needed by the proof.

By Lemma 4.11, we get an upper bound as below. We note that the total influence of $f^{\leqslant d}$ is $O(d)$ and its maximum influence is at most τ by hypothesis.

$$\sum_{i=1}^{2n} I_i[g]^4 \leqslant C(d,m) \left(\sum_{i=1}^{2n} I_i^{\leqslant d}[f]^4 \;+\; \frac{1}{\sqrt{n}} \right) \leqslant C(d,m)\, \tau^3.$$

This gives, by Lemma 4.10, that

$$\Pr_\pi\left[\exists j : I_j[g|_\pi] \geqslant \delta\right] \;\leqslant\; \frac{C(d,K,m)}{\delta^4}\, \tau^3.$$

Finally, by Lemma 4.12, except with probability ζ over the choice of π, it holds that

$$\max_{1 \leqslant j \leqslant n} I_j[f|_\pi] \;\leqslant\; 3 \cdot \max_{1 \leqslant j \leqslant n} I_j[g|_\pi] + \delta.$$

Putting the two conclusions together, we conclude that

$$\Pr_\pi\left[\exists j : I_j[f|_\pi] \geqslant 4\delta\right] \;\leqslant\; \zeta + \frac{C(d,K,m)}{\delta^4}\,\tau^3 \;\leqslant\; 2\zeta,$$

completing the proof of Lemma 3.2. In terms of quantitative constraints on the parameters, $K = K(\delta,\zeta), \gamma = \gamma(m,\delta,\zeta)$ are determined by Lemma 4.10 and $\tau = \tau(d,m,\delta,\zeta)$ needs to obey the very last inequality above.

Proof of Lemma 3.1

Proof of Lemma 3.1 follows along the same lines with one minor change. In this case, $f^{\leqslant d}$ could have influential variables and we treat them separately. Let

$$L = \{i \in [2n] \mid I_i^{\leqslant d}[f] \geqslant \tau\}.$$

In the proof of Lemma 4.10, instead of considering the event $I_j[g|_\pi] \geqslant \delta$, we consider a more refined event $I_j[g|_\pi] \geqslant \delta \;\wedge\; \pi^{-1}(j) \cap L = \emptyset$. Basically the same argument gives

$$\Pr_\pi\left[\exists j : I_j[g|_\pi] \geqslant \delta \;\wedge\; \pi^{-1}(j) \cap L = \emptyset\right] \leqslant \frac{C(d,K,m)}{\delta^4} \cdot \sum_{i \in [2n]\setminus L} I_i[g]^4.$$

The latter sum is bounded as before by $C(d,m)\tau^3$ so that

$$\Pr_\pi\left[\exists j : I_j[g|_\pi] \geqslant \delta \;\wedge\; \pi^{-1}(j) \cap L = \emptyset\right] \;\leqslant\; \frac{C(d,K,m)}{\delta^4}\,\tau^3.$$

As before, except with probability ζ over the choice of π, it holds that

$$\max_{1 \leqslant j \leqslant n} I_j[f|_\pi] \;\leqslant\; 3 \cdot \max_{1 \leqslant j \leqslant n} I_j[g|_\pi] + \delta.$$

Putting the two conclusions together, we conclude that

$$\Pr_\pi\left[\exists j : I_j[f|_\pi] \geqslant 4\delta \;\wedge\; \pi^{-1}(j) \cap L = \emptyset\right] \;\leqslant\; \zeta + \frac{C(d,K,m)}{\delta^4}\,\tau^3 \;\leqslant\; 2\zeta,$$

completing the proof of Lemma 3.1.

5 The Reduction

We now prove Theorem 1.8 that the Rich 2-to-1 Games Conjecture is equivalent to the Unique Games Conjecture. The reduction from Unique Games to Rich 2-to-1 Games as well as its analysis are standard and are sketched in Appendix B. The reduction from Rich 2-to-1 Games to Unique Games is also standard and is presented in this section. Its analysis however needs new analytic tools, specifically Lemma 3.1.

We are given a Rich 2-to-1 Games instance $\Psi = (L \cup R, E, \Sigma_L, \Sigma_R, \Phi)$ with $\Sigma_L = [2n]$, $\Sigma_R = [n]$, completeness (at least) $1-\eta$, and soundness (at most) η. The reduction outputs an instance of Unique Games

with alphabet $[m]$, completeness (at least) $1-5\varepsilon$, and soundness (at most) ε. For given ε, first m needs to be taken sufficiently large, then η sufficiently small, and in turn n sufficiently large. The instance of Unique Games produced is linear, i.e. its alphabet $[m]$ is identified with $\mathbb{Z}_m$, the additive group of integers modulo m, and the constraints are linear equations.

As is standard, we replace a vertex $u \in L$ with the (supposed) m-ary long code of the (supposed) label of u. The positions in the long code correspond to the variables of the Unique Games instance. An assignment to these variables corresponds to a function $F_u : [m]^{2n} \to [m]$. The intention is that if $i \in [2n]$ is the label of u (in the 2-to-1 Games instance), then $F_u(x) = F_u(x_1, \ldots, x_{2n}) = x_i$ is the corresponding dictatorship function. In our reduction, the long codes for labels of vertices $v \in R$ do not appear explicitly, but it will be convenient to imagine them "virtually".

The PCP test is straightforward: one performs a two-query "noise-test" on the virtual long code of a vertex $v \in R$, but actually reads off the queries from the long codes of neighbors u, w of the vertex v respectively (the virtual long code for v is "contained" in that of u as well as w). Each test is viewed as a Unique Games constraint and this defines the Unique Games instance produced by the reduction. Formally, a test/equation is produced as follows:

- Sample $v \in R$ at random, $a \in [m]^n$ at random, and $b \in [m]^n$ that is $1-\varepsilon$ correlated with a.
- Sample two neighbours u, w of v independently at random.
- Set
$$A = \pi^{-1}_{(u,v)}(a), \quad B = \pi^{-1}_{(w,v)}(b) \in [m]^{2n}.$$
- Finally, sample $x \in [m]^{2n}$ that is $1-\varepsilon$ correlated with A, and $y \in [m]^{2n}$ that is $1-\varepsilon$ correlated with B. Output the equation
$$F_u(x) = F_w(y).$$

Folding. As is standard, we can assume that the functions $F_u : [m]^{2n} \to [m]$ that appear in the PCP proof are folded, meaning $F_u(x + se) = F_u(x) + s$ where $e \in [m]^{2n}$ is the all 1 vector. In particular, F_u is then balanced, i.e. takes all values in $[m]$ equally often. Technically, folding is enforced by keeping only one of the inputs in the set $\{x + se \,|s \in [m]\}$ as a representative and inferring values at other inputs from the representative. The effect of folding is that the equations produced are of the type $p = q + s$ instead of just $p = q$ where p, q are the Unique Games variables in the output instance and $s \in [m]$.

5.1 Completeness

If the 2-to-1 Games instance Ψ has a labeling $\sigma\colon L \to [2n]$, $\rho\colon R \to [n]$ that satisfies at least $1-\eta$ fraction of the constraints, we show that the Unique Games instance is (at least) $1-2\eta-3\varepsilon \geqslant 1-5\varepsilon$ satisfiable for η sufficiently small.

Indeed, define for any $u \in L$, the long-code assignment $F_u(x) = x_{\sigma(u)}$. Since the edges $(u, v), (w, v)$ are distributed uniformly, with probability at least $1-2\eta$, both edges are satisfies by the labeling, i.e. $\pi_{(u,v)}(\sigma(u)) = \rho(v) = \pi_{(w,v)}(\sigma(w))$. Whenever this happens, the test accepts with probability at least $1-3\varepsilon$ since the failure to accept can be attributed to one of three events: strings a and b differing on the co-ordinate $\rho(v)$, strings x and A differing on the co-ordinate $\sigma(u)$, or strings y and B differing on the co-ordinate $\sigma(w)$.

5.2 Soundness

We will show that if the 2-to-1 Games instance has soundness at most η (to be chosen sufficiently small later), then the probability that the test accepts is upper bounded by ε.

Let $F_u : [m]^{2n} \to [m]$ be the folded functions given as assignment to the Unique Games instance. In a standard manner, we view the functions as $F_u : [m]^{2n} \to \Delta_m$ where $\Delta_m = \{(t_0, \ldots, t_{m-1}) | t_i \geqslant 0, \sum_{i=0}^{m-1} t_i = 1\}$ is the standard m-dimensional simplex. Each function F_u is then thought of as a vector $(F_{u,0}, \ldots, F_{u,m-1})$ where each $F_{u,r}$ is a $\{0,1\}$-valued function and $\mathbb{E}[F_{u,r}] = \frac{1}{m}$ since F_u is folded and balanced. Moreover, the acceptance criterion of the test, i.e. $F_u(x) = F_w(y)$, can be written arithmetically as $\sum_{r=0}^{m-1} F_{u,r}(x) \cdot F_{w,r}(y)$. Hence the probability that the test accepts can be written as (the expectation is over all the choices made)

$$\mathop{\mathbb{E}}_{v,u,w,a,b,A,B,x,y} \left[\sum_{r=0}^{m-1} F_{u,r}(x) \cdot F_{w,r}(y) \right] = \sum_{r=0}^{m-1} \mathop{\mathbb{E}}_{v,u,w,a,b,A,B,x,y} [F_{u,r}(x) \cdot F_{w,r}(y)]. \tag{6}$$

Henceforth we will fix the index $0 \leqslant r \leqslant m-1$ and then show an upper bound on the expectation on the right (with the overall upper bound being m times that). For notational convenience, we drop the subscript r and define $\{0,1\}$-valued functions $f_u = F_{u,r}$. Thus the goal is to upper bound (note that $a, b \in [m]^n$ and $A, B, x, y \in [m]^{2n}$)

$$\mathop{\mathbb{E}}_{\substack{v,u,w \\ a \sim_{1-\varepsilon} b}} \left[\mathop{\mathbb{E}}_{x \sim_{1-\varepsilon} A,\ y \sim_{1-\varepsilon} B} [f_u(x) \cdot f_w(y)] \right] = \mathop{\mathbb{E}}_{\substack{v,u,w, \\ a \sim_{1-\varepsilon} b}} [T_{1-\varepsilon} f_u(A) \cdot T_{1-\varepsilon} f_w(B)] = \mathop{\mathbb{E}}_{\substack{v,u,w, \\ a \sim_{1-\varepsilon} b}} [g_u(A) \cdot g_w(B)],$$

where $g_u = T_{1-\varepsilon} f_u$. We note that g_u is $[0,1]$-valued and $\mathbb{E}[g_u] = \mathbb{E}[f_u] = \frac{1}{m}$. We further define $g_{u,v} = g_u|_{\pi_{(u,v)}}$ and we still have $\mathbb{E}[g_{u,v}] = \frac{1}{m}$ (see Appendix C for this subtle point). The expectation can be rewritten as

$$\mathop{\mathbb{E}}_{\substack{v,u,w, \\ a \sim_{1-\varepsilon} b}} [g_{u,v}(a) \cdot g_{w,v}(b)] = \mathop{\mathbb{E}}_{v,u,w} [\langle g_{u,v}, T_{1-\varepsilon} g_{w,v} \rangle] = \mathop{\mathbb{E}}_{v} [\langle h_v, T_{1-\varepsilon} h_v \rangle], \tag{7}$$

where in the last step we used the fact that the choices of u, w are independent (for a fixed v) and defined $h_v = \mathbb{E}_u[g_{u,v}]$. We note that $\mathbb{E}[h_v] = \frac{1}{m}$ as well. We now show, by way of contradiction, that if the expectation in (7) is at least $\beta = \frac{\varepsilon}{m}$, then one can define a labeling to the 2-to-1 Games instance that satisfies more than η fraction of its constraints. It then follows that (7) is bounded by β and hence (6) (i.e. the acceptance probability of PCP test) by $m\beta = \varepsilon$ as desired.

Assume therefore that the expectation in (7) is at least β. By an averaging argument, for at least $\frac{\beta}{2}$ fraction of vertices $v \in R$, the inner product $\langle h_v, T_{1-\varepsilon} h_v \rangle$ is at least $\frac{\beta}{2}$. Let $R_{\mathsf{Good}} \subseteq R$ be the subset of such vertices. That is, for $v \in R_{\mathsf{Good}}$,

$$\langle h_v, T_{1-\varepsilon} h_v \rangle \geqslant \frac{\beta}{2}. \tag{8}$$

Using Theorem 2.6, the function h_v then must have an influential co-ordinate, and moreover since $h_v = \mathbb{E}_u[g_{u,v}]$, so does the function $g_{u,v}$ for a good fraction of the neighbors $u \in L$. In light of this observation, we hope to come up with a labeling to vertices $v \in R$ and $u \in L$ by choosing an influential co-ordinate of the function h_v and the function g_u respectively (we need to use the fact that g_u is smooth or low-degree). This strategy works thanks to our main technical Lemma 3.1.

Indeed, for $v \in R_{\mathsf{Good}}$, define its label $\rho(v)$ to be an arbitrary co-ordinate $j \in [n]$ such that $I_j[h_v] \geqslant \delta$. Such a coordinate exists since $\langle h_v, T_{1-\varepsilon} h_v \rangle \geqslant \frac{\beta}{2}$ and using Theorem 2.6. One needs to take m sufficiently

large so that $\mathbb{E}[h_v] = \frac{1}{m} = \theta$ is sufficiently small to bring the bound $2\Gamma_{1-\varepsilon}(\theta)$ in Theorem 2.6 below $\frac{\beta}{2} = \frac{\varepsilon}{2}\frac{1}{m} = \frac{\varepsilon}{2}\theta$. One then needs to take the influence parameter δ therein sufficiently small.

Since $h_v = \mathbb{E}_u[g_{u,v}]$ and $\rho(v)$ has influence at least δ on h_v, it follows that for at least $\frac{\delta}{2}$ fraction of neighbors $u \in L$ of v we have $I_{\rho(v)}[g_{u,v}] \geqslant \frac{\delta}{2}$. Let $N_{\mathsf{Good}}(v)$ denote the subset of such neighbors. We emphasize that for a random choice of edge (u, v), we have $v \in R_{\mathsf{Good}}$ and $u \in N_{\mathsf{Good}}(v)$ with probability at least $\frac{\beta}{2}\frac{\delta}{2}$.

Now, by the main Lemma 3.1, except with probability $\zeta \ll \frac{\beta\delta}{8}$ over the choice of the edge (u, v), it is the case that whenever $I_j[g_{u,v}] \geqslant \frac{\delta}{2}$ for some $j \in [n]$, one has $I_i[g_u] \geqslant \tau$ for some $i \in \pi^{-1}(j), \pi = \pi_{(u,v)}$. We note that since $g_u = T_{1-\varepsilon}f_u$, its Fourier mass beyond degree d is at most $\gamma = 2^{-\Omega(d/\varepsilon)}$, which can be made sufficiently small by taking d sufficiently large. Finally, τ is taken to be sufficiently small so that the lemma applies. It follows that with probability $\frac{\beta\delta}{8}$, all these events happen simultaneously:

$$v \in R_{\mathsf{Good}}, \quad u \in N_{\mathsf{Good}}(v), \quad I_{\rho(v)}[g_{u,v}] \geqslant \frac{\delta}{2}, \quad I_i^{\leqslant d}[g_u] \geqslant \tau \text{ for some } i \in \pi^{-1}(\rho(v)).$$

Thus if we defined a label for $u \in L$ by making a list of all co-ordinates with degree-d influence at least τ on g_u and then picked one label at random from this list, it would agree with $\rho(v)$ (via $\pi = \pi_{(u,v)}$) with probability at least $\Omega(\frac{\tau}{d})$ (the list size is $O(\frac{d}{\tau})$ since the total degree-d influence is at most d). This gives a labeling to the 2-to-1 Games instance that satisfies overall $\Omega(\frac{\beta\delta\tau}{d})$ fraction of its constraints. Choosing the soundness η of the 2-to-1 Games instance to be even lower a priori completes the proof.

References

[1] 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.

[2] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *J. ACM*, 45(1):70–122, 1998.

[3] Boaz Barak, Pravesh K. Kothari, and David Steurer. Small-set expansion in shortcode graph and the 2-to-2 conjecture. In *10th Innovations in Theoretical Computer Science Conference, ITCS 2019, January 10-12, 2019, San Diego, California, USA*, pages 9:1–9:12, 2019.

[4] Amey Bhangale, Subhash Khot, and Devanathan Thiruvenkatachari. An improved dictatorship test with perfect completeness. In *37th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2017, December 11-15, 2017, Kanpur, India*, pages 15:1–15:23, 2017.

[5] Irit Dinur, Subhash Khot, Guy Kindler, Dor Minzer, and Muli Safra. On non-optimally expanding sets in grassmann graphs. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 940–951, 2018.

[6] Irit Dinur, Subhash Khot, Guy Kindler, Dor Minzer, and Muli Safra. Towards a proof of the 2-to-1 games conjecture? In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 376–389, 2018.

[7] Irit Dinur, Elchanan Mossel, and Oded Regev. Conditional hardness for approximate coloring. *SIAM J. Comput.*, 39(3):843–873, 2009.

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

[9] Uriel Feige and László Lovász. Two-prover one-round proof systems: Their power and their problems (extended abstract). In *Proceedings of the 24th Annual ACM Symposium on Theory of Computing, May 4-6, 1992, Victoria, British Columbia, Canada*, pages 733–744, 1992.

[10] Yuval Filmus, Guy Kindler, Noam Lifshitz, and Dor Minzer. Hypercontractivity for global functions on the symmetric group. *in preparation*, 2019.

[11] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the 17th Annual IEEE Conference on Computational Complexity, Montréal, Québec, Canada, May 21-24, 2002*, page 25, 2002.

[12] Subhash Khot. Inapproximability of NP-complete problems, discrete fourier analysis, and geometry. In *Proceedings of the International Congress of Mathematicians 2010*, pages 2676–2697, 2010.

[13] Subhash Khot. On the unique games conjecture. In *Proceedings of the 25th Annual IEEE Conference on Computational Complexity, CCC 2010, Cambridge, Massachusetts, June 9-12, 2010*, pages 99–121, 2010.

[14] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for max-cut and other 2-variable csps? *SIAM J. Comput.*, 37(1):319–357, April 2007.

[15] Subhash Khot, Dor Minzer, Dana Moshkovitz, and Muli Safra. Small set expansion in the johnson graph. *Electronic Colloquium on Computational Complexity (ECCC)*, 25:78, 2018.

[16] Subhash Khot, Dor Minzer, and Muli Safra. On independent sets, 2-to-2 games, and grassmann graphs. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 576–589, 2017.

[17] Subhash Khot, Dor Minzer, and Muli Safra. Pseudorandom sets in grassmann graph have near-perfect expansion. In *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018*, pages 592–601, 2018.

[18] Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within 2-epsilon. *J. Comput. Syst. Sci.*, 74(3):335–349, 2008.

[19] Subhash Khot and Muli Safra. A two-prover one-round game with strong soundness. *Theory of Computing*, 9(28):863–887, 2013.

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

[21] Suguru Tamaki and Yuichi Yoshida. A query efficient non-adaptive long code test with perfect completeness. *Random Struct. Algorithms*, 47(2):386–406, 2015.

[22] Luca Trevisan. On Khot's unique games conjecture. *Bulletin of the AMS*, 49(1):91–111, 2012.

A Standard Facts and Calculations

Let the entropy function be $H(p_1, \ldots, p_r) = \sum_{i=1}^{r} p_i \log(1/p_i)$. By Stirling's formula one has:

Fact A.1. *For positive integers $v_1, \ldots, v_r$ that sum up to n we have*

$$\binom{n}{v_1, \ldots, v_r} = \Big(1/(2\pi)^{(r-1)/2} + o_n(1)\Big) \sqrt{\frac{n}{v_1 \cdots v_r}}\, 2^{H(\frac{v_1}{n}, \ldots, \frac{v_r}{n}) \cdot n}$$

Lemma A.2. *For any $\varepsilon_1, \ldots, \varepsilon_r \in [-1, 1]$ such that $\sum_{i=1}^{r} \varepsilon_i = 0$, we have*

$$H\left(\frac{1+\varepsilon_1}{r}, \ldots, \frac{1+\varepsilon_r}{r}\right) \geqslant \log r - \frac{1}{r}\sum_{i=1}^{r} \varepsilon_i^2.$$

Proof. By definition of entropy, using that the ε_i sum to zero, and that $\log(1+\varepsilon_i) \leqslant \varepsilon_i$, we have

$$\begin{aligned} H\left(\frac{1+\varepsilon_1}{r}, \ldots, \frac{1+\varepsilon_r}{r}\right) &= \sum_{i=1}^{r} \frac{1+\varepsilon_i}{r} \log\left(\frac{r}{1+\varepsilon_i}\right) \\ &= \log r - \sum_{i=1}^{r} \frac{1+\varepsilon_i}{r} \log(1+\varepsilon_i) \\ &\geqslant \log r - \sum_{i=1}^{r} \frac{1+\varepsilon_i}{r}\, \varepsilon_i \\ &= \log r - \frac{1}{r}\sum_{i=1}^{r} \varepsilon_i^2. \end{aligned}$$

□

Lemma A.3. *For any positive integers r, t, there are constants $0 < c(r,t) < C(r,t)$ such that the following holds for large enough n. Let $v_1, \ldots, v_r \geqslant \frac{n}{10r}$ be integers that sum up to n, let $u_1, \ldots, u_r$ be non-negative integers that are each at most t, and denote $u = u_1 + \ldots + u_r$. Then*

$$c(r,t)\binom{n}{v_1, \ldots, v_r} \leqslant \binom{n-u}{v_1 - u_1, \ldots, v_r - u_r} \leqslant C(r,t)\binom{n}{v_1, \ldots, v_r}.$$

Proof. By definition, the ratio between the two multinomial coefficients is equal to

$$\frac{\binom{n}{v_1, \ldots, v_r}}{\binom{n-u}{v_1-u_1, \ldots, v_r-u_r}} = \frac{n!}{(n-u)!} \prod_{i=1}^{r} \frac{(v_i - u_i)!}{v_i!} = \frac{n(n-1)\cdots(n-u+1)}{\prod_{i=1}^{r} v_i(v_i-1)\cdots(v_i-u_i+1)}.$$

We first argue that this is upper bounded by $C(r,t)$. Indeed, the numerator is at most n^u, whereas the denominator is at least

$$\prod_{i=1}^{r} (v_i - u_i)^{u_i} \geqslant \prod_{i=1}^{r} \Big(\frac{n}{10r} - t\Big)^{u_i} \geqslant \prod_{i=1}^{r} \Big(\frac{n}{20r}\Big)^{u_i} = \Big(\frac{n}{20r}\Big)^{u}.$$

Hence the ratio between the multinomial coefficients it at most $(20r)^u \leqslant (20r)^{rt} \stackrel{def}{=} C(r,t)$. Similarly, for the lower bound, the numerator is at least $(n-u)^u \geqslant (n-tr)^u \geqslant (\frac{n}{2})^u$, whereas the denominator is at most $\prod_{i=1}^{r} v_i^{u_i} \leqslant n^u$, so the ratio between the multinomial coefficients is at least $2^{-u} \geqslant 2^{-rt} = c(r,t)$. □

Lemma A.4. *For any positive integer r, there is a constant $C(r) > 0$ such that the following holds for large enough n. Let $v_1, \dots, v_r$ be integers of absolute value at most $\sqrt{K \cdot r \cdot n}$ that sum to zero, and such that for every $1 \leqslant i \leqslant r$, the integer $n + v_i$ is divisible by r. Then*

$$\frac{\binom{n}{\frac{n+v_1}{r}, \dots, \frac{n+v_r}{r}}}{\binom{2n}{\frac{2(n+v_1)}{r}, \dots, \frac{2(n+v_r)}{r}}} \leqslant C(r) 2^{Kr} r^{-n}.$$

Proof. Using Fact A.1, the left hand side is equal to

$$C(r) \cdot \frac{2^{H\left(\frac{n+v_1}{nr}, \dots, \frac{n+v_r}{nr}\right) \cdot n}}{2^{H\left(\frac{n+v_1}{nr}, \dots, \frac{n+v_r}{nr}\right) \cdot 2n}} = C(r) \cdot 2^{-H\left(\frac{n+v_1}{nr}, \dots, \frac{n+v_r}{nr}\right) \cdot n}.$$

Using Fact A.2,

$$H\left(\frac{n+v_1}{nr}, \dots, \frac{n+v_r}{nr}\right) \geqslant \log r - \frac{1}{r} \sum_{i=1}^{r} \left(\frac{v_i}{n}\right)^2 \geqslant \log r - \frac{1}{r} \sum_{i=1}^{r} \frac{Kr}{n} = \log r - \frac{Kr}{n}.$$

Hence, we ge the desired upper bound of $C(r) \cdot 2^{-n \log r + Kr} = C(r) 2^{Kr} r^{-n}$. □

Let $A_1, \dots, A_r$ be a partition of $[2n]$ into r even-sized sets. We say a mapping $\pi \in S_{2n,n}$ is consistent with $A_1, \dots, A_r$ if matching given by π matches off each set A_i within itself (or equivalently that $\pi^{-1}(\pi(A_i)) = A_i$).

Lemma A.5. *Let $A_1, \dots, A_r$ be a partition of $[2n]$ into even-sized sets, and denote their sizes by $a_1, \dots, a_r$. Then*

$$\Pr_{\pi \in S_{2n,n}} [\pi \textit{ is consistent with } A_1, \dots, A_r] = \frac{\binom{n}{\frac{a_1}{2}, \dots, \frac{a_r}{2}}}{\binom{2n}{a_1, \dots, a_r}}.$$

Proof. We count the number of π that are consistent with the partition. To begin with, the number of matchings of $[2n]$ that match off every A_i within itself is

$$\prod_{i=1}^{r} \frac{a_i!}{2^{\frac{1}{2}a_i} (\frac{a_i}{2})!} = 2^{-n} \prod_{i=1}^{r} \frac{a_i!}{(\frac{a_i}{2})!}.$$

Given such a matching, each matched pair should be mapped to a distinct element of $[n]$, so there are $n!$ choices of π that can be produced from the matching. In total, the number of π that are consistent with the partition is $n!\ 2^{-n} \prod_{i=1}^{r} \frac{a_i!}{(\frac{a_i}{2})!}$. To get the desired probability, we divide this number by the total number of $\pi \in S_{2n,n}$, which is $2^{-n}(2n)!$. □

B Reduction from Unique Games to Rich 2-to-1 Games

In this section, we give a reduction from Unique Games to Rich 2-to-1 Games.

Given a Unique Games instance $\phi = (L \cup R, E, \Sigma, \Phi)$, we construct a Rich 2-to-1 Games instance $\psi = (U \cup V, E', \Sigma_U, \Sigma_V, \Psi)$ as follows. First, note that we may assume that the size of the alphabet in ϕ, namely $|\Sigma|$, is even [3], and we assume $\Sigma = [2k]$ for $k \in \mathbb{N}$ henceforth.

To construct the instance ψ, let U be a copy of L and set $\Sigma_U = \Sigma$. Also, define $V = R \times S_{2k,k}$ (recall that $S_{2k,k}$ is the set of all 2-to-1 mappings from $[2k]$ to $[k]$) and $\Sigma_V = [k]$; a label j of a vertex (v, π) should be thought of as "one of the labels in $\pi^{-1}(j)$" for v. The vertices $u \in U$ and $(v, \pi) \in V$ are adjacent in ψ if (u, v) is an edge in ϕ, and the constraint on them is given by

$$\Psi(u, (v, \pi)) = \left\{ (i, j) \mid \exists \sigma_v \in \pi^{-1}(j) \text{ such that } (i, \sigma_v) \in \Phi(u, v) \right\}.$$

This completes the description of the reduction.

The completeness of the reduction, as well as the fact that ψ is a Rich 2-to-1 Games instance, are both easy to see. For the soundness, given assignments $A \colon U \to \Sigma_U$ and $B \colon V \to \Sigma_V$ satisfying δ-fraction of the constraints in ψ, one can construct assignments A', B' for ψ satisfying at least $\delta/2$ fraction of the constraints, as follows: take $A' \equiv A$, and for B' is the randomized assignment defined on each $v \in V$ by taking $\pi \in S_{2k,k}$ uniformly, picking an element $\sigma \in \pi^{-1}(B(v, \pi))$ uniformly and setting $B'(v) = \sigma$.

C Symmetries induced by Folding

In the soundness analysis in Section 5.2, we used certain symmetry properties that are ensured by folding. We point these out here for clarity. Let $F : [m]^{2n} \to [m]$ be a folded function. Viewing $F : [m]^{2n} \to \Delta_m$ as a function into the simplex, the folding condition amounts to saying (here $\mathbf{c}$ stands for the center of the simplex)

$$\forall z \in [m]^{2n}, \quad \frac{1}{m} \sum_{s \in [m]} F(z + se) = \mathbf{c} = \left(\frac{1}{m}, \ldots, \frac{1}{m} \right) \in \Delta_m.$$

Let $G = T_{1-\varepsilon} F$ (thought of as $T_{1-\varepsilon}$ applied coordinate-wise). We claim that G also satisfies

$$\forall z \in [m]^{2n}, \quad \frac{1}{m} \sum_{s \in [m]} G(z + se) = \mathbf{c}. \tag{9}$$

Indeed,

$$\begin{aligned} \sum_{s \in [m]} G(z + se) &= \sum_{s \in [m]} \mathop{\mathbb{E}}_{y \sim_{1-\varepsilon} z + se} [F(y)] = \sum_{s \in [m]} \mathop{\mathbb{E}}_{y \sim_{1-\varepsilon} z} [F(y + se)] \\ &= \mathop{\mathbb{E}}_{y \sim_{1-\varepsilon} z} \left[\sum_{s \in [m]} F(y + se) \right] = m\mathbf{c}. \end{aligned}$$

Finally, we observe that if $\pi : [2n] \to [n]$ is a 2-to-1 map and $H = G|_\pi : [m]^n \to \Delta_m$ (again, coordinate-wise), then applying (9) to $z = \pi^{-1}(x)$, we see that (here e' is the n-dimensional all 1 vector).

$$\forall x \in [m]^n, \quad \frac{1}{m} \sum_{s \in [m]} H(x + se') = \mathbf{c}.$$

In particular, if f_r, g_r, h_r are r^{th} coordinate functions of F, G, H respectively, $0 \leqslant r \leqslant m-1$, then we have $\mathbb{E}[f_r] = \mathbb{E}[g_r] = \mathbb{E}[h_r] = \frac{1}{m}$.

[3] Otherwise, we may construct a Unique Games instance ϕ' on the same graph whose alphabet is $\Sigma \times \{0, 1\}$, and for each edge (u, v), the constraint $\Phi'(u, v)$ demands that the label (σ_u, b_u) of u and (σ_v, b_v) of v satisfy $(\sigma_u, \sigma_v) \in \Phi(u, v)$ and $b_u = b_v$. It is easy to see that the value of both instances is equal.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il