Electronic Colloquium on Computational Complexity, Report No. 110 (2012)

# APPROXIMATION RESISTANCE FROM PAIRWISE INDEPENDENT SUBGROUPS

SIU ON CHAN

Abstract. We show optimal (up to constant factor) NP-hardness for Max-$k$-CSP over any domain, whenever $k$ is larger than the domain size. This follows from our main result concerning predicates over abelian groups. We show that a predicate is approximation resistant if it contains a subgroup that is balanced pairwise independent. This gives an unconditional analogue of Austrin–Mossel hardness result, taking away their Unique-Games Conjecture assumption in exchange for an abelian subgroup structure.

Our main ingredient is a new technique to reduce soundness, which is inspired by XOR-lemmas. Using this technique, we also improve the NP-hardness of approximating Independent-Set on bounded degree graphs, Almost-Coloring, and Two-Prover-One-Round-Game.

## 1. Introduction

There is a huge gap between NP-hardness and algorithmic results for approximating Max-$k$-CSP, which is the task of finding the best assignment in a constraint satisfaction problem (CSP) of arity $k$. For boolean alphabet, the best algorithms by Charikar, Makarychev, and Makarychev [CMM09, MM12] have approximation ratio $\Omega(k/2^k)$, but the best NP-hardness results by Engebretsen, Holmerin, Samorodnitsky, and Trevisan [ST00, EH08] have hardness ratio $2^{O(\sqrt{k})}/2^k$, which is significantly larger by a factor of $\exp(\Omega(\sqrt{k}))$.

A related question is to identify predicates[1] that are approximation resistant, that is, predicates whose CSPs are NP-hard to approximate better than a random assignment. Approximation resistant predicates of arity 3 and 4 have been extensively studied in [Has05b] (see also [Hås01, Zwi98]), but only a handful of such predicates were known for higher arity. We are only aware of the following scattered examples: linear equations over abelian groups [Hås01], Engebretsen–Holmerin predicates [EH00], bipartite-graph predicates of Samorodnitsky and Trevisan [ST00], and those containing the bipartite-graph predicates [Has05b].

To make progress, many works obtained conditional results assuming Khot's Unique-Games Conjecture [Kho02b]. Under this conjecture, Samorodnitsky and Trevisan [ST09] showed that Max-$k$-CSP is NP-hard to approximate beyond $2k/2^k$, matching the best algorithm up to constant factor, and later Raghavendra [Rag08] obtained optimal inapproximability (and algorithmic) results for every CSP. Under the same conjecture, Austrin and Mossel [AM09] showed that a predicate is approximation resistant if it supports a balanced pairwise independent distribution. However, the UG conjecture remains uncertain. It is also desirable to look for new reduction techniques not relying on any conjecture.

In this work, we obtain the first general criterion for approximation resistant predicate (bypassing the conjecture), and settle the NP-hardness of Max-$k$-CSP (up to constant factor). We consider predicates $C$ over a domain $G$ which is an *abelian group*, such that $C$ is a *subgroup* satisfying a condition similar to Austrin and Mossel's [AM09]. We call such CSPs *additive* (see Section 4 for definitions).

*Date*: September 4, 2012.

[1] In this paper, we define a $k$-ary predicate $C$ over a domain $\Sigma$ as a subset of $\Sigma^k$, rather than a function $\Sigma^k \to \{0,1\}$. This justifies statements such as "a predicate *contains* another."



ISSN 1433-8092

**Theorem 1.1** (Main). *Let $k \geqslant 3$ be an integer, $G$ a finite abelian group, and $C$ a balanced pairwise independent subgroup of $G^k$. For some $\varepsilon = o_{n;k,|G|}(1)$,[2] it is* NP*-hard to $(1-\varepsilon, |C|/|G|^k + \varepsilon)$-decide*[3] Additive-CSP$(C)$.

A random assignment satisfies $|C|/|G|^k$ fraction of constraints in expectation, so our hardness ratio is tight. Our proof actually shows that any predicate containing such a subgroup $C$ is also approximation resistant, so these subgroups $C$ are hereditarily approximation resistant. Compared with Austrin and Mossel's [AM09], our result requires an abelian subgroup structure on the predicate, but avoids the UG Conjecture assumption.

Curiously, our result is also related to integrality gaps in the Lasserre semi-definite programming (SDP) hierarchy [Las01]. The existing (direct) construction of Lasserre gaps [Sch08, Tul09] seems to require both pairwise independence and abelian subgroup structure;[4] conversely, these two conditions are also sufficient for the construction (Appendix G). This observation has motivated our Theorem 1.1, even though the same hypothesis is exploited differently.

Theorem 1.1 settles the approximability of Max-$k$-CSP (up to constant factor), by choosing $C$ to be a hypergraph predicate of Samorodnitsky and Trevisan [ST09].

**Corollary 1.2.** *For any $k \geqslant 3$, there is $\varepsilon = o_{n;k}(1)$ such that it is* NP*-hard to $(1-\varepsilon, 2k/2^k + \varepsilon)$-decide* Max-$k$-CSP *over boolean domain.*

Our result also throws away the UG Conjecture assumption in Håstad's result [Hås09], showing that a random predicate is hereditarily approximation resistant with high probability, answering an open problem in that paper.

| **Problem** | NP**-Hardness** | |
|---|---|---|
| Max-$k$-CSP | $2^{O(\sqrt{k})}/2^k$ | [ST00, EH08] |
| (over $\mathbb{Z}_2$) | $2k/2^k$ | This work |
| Max-$k$-CSP | $q^{O(\sqrt{q})}/q^k$ | [Eng05] |
| (alphabet size $q$) | $q(q-1)k/q^k$ | This work |
| | $O(qk/q^k) \quad (k \geqslant q)$ | This work |
| 2-Prover-1-Round-Game | $1/R^{\Omega(1)}$ | [Raz98, Hol09, Rao08] |
| (alphabet size $R$) | $4/R^{1/6}$ | [KS11] |
| | $O(\log R)/\sqrt{R}$ | This work |
| Independent-Set | $1/D^{\Omega(1)}$ | [AFWZ95] |
| (degree bound $D$) | $\exp(O(\sqrt{\log D}))/D$ | [Tre01] |
| | $O(\log D)^4/D$ | This work |
| Almost-Coloring | $1/K^2$ | [DKPS10] |
| (almost $K$-colorable) | $1/\exp(\Omega(\log K)^2)$ | [KS12] |
| | $1/2^{K/2}$ | This work |

FIGURE 1. Summary of NP-hardness results.

[2]The notation $\varepsilon = o_{n;k,|G|}(1)$ means that for any fixed $k$, any fixed $|G|$, the quantity $\varepsilon$ goes to zero as $n$ goes to infinity.

[3]We say it is NP-hard to $(c,s)$-decide a CSP if given an instance $M$ of the CSP, it is NP-hard to decide whether the best assignment to $M$ satisfies at least $c$ fraction of constraints, or satisfies at most $s$ fraction. The parameters $c$ and $s$ are known as completeness and soundness, respectively. The hardness ratio is $s/c$.

[4]Abelian subgroup structure seems indispensable, as demonstrated in the Fourier analytic proof of Theorem G.4. For an abelian subgroup predicate, the balanced pairwise independence property is *necessary* for a random instance of its CSP to have exponential resolution complexity [CM08], which is an essential ingredient in the existing construction.

1.1. **Additional results.** Another way to state Corollary 1.2 is a Probabilistically Checkable Proof (PCP) that is optimally query-efficient. Query efficiency is measured by amortized query complexity, defined as $k/\log_2(c/s)$ when a PCP verifier read $k$ bits from a proof, and has completeness $c$ and soundness $s$ [BGS98, Section 2.2.2].

**Corollary 1.3.** *For every $k \geqslant 3$, for some $\varepsilon = o_{n;k}(1)$, there is a PCP for* NP *that reads $k$ bits, uses randomness $(1+\varepsilon)k \log n$, has completeness $1-\varepsilon$, and has amortized query complexity $1+(1+o_k(1))(\log k)/k+\varepsilon$.*

As an improvement over [ST00, EH08], our amortized query complexity is tight up to the $o_k(1)$ term unless $\mathrm{P} = \mathrm{NP}$ [Has05a, CMM09]. We also reduce the amortized free bit complexity of a PCP. A PCP has free bit complexity $f$ if on every choice of randomness, there are at most $2^f$ accepting local views (out of the $2^k$ possibilities for the $k$ bits read). Amortized free bit complexity is then $f/\log_2(c/s)$ [BGS98, Section 2.2.2]. Our PCP has amortized free bit complexity $(1+o_k(1))(\log k)/k$, up to additive $o_{n;k}(1)$.

Amortized free bit complexity has applications to a number of problems, including Independent-Set and Almost-Coloring.

**Theorem 1.4.** *For all sufficiently large $D$, there is $\nu = o_{n;D}(1)$ such that it* NP*-hard to approximate* Independent-Set *on degree-$D$ graphs beyond $O(\log D)^4/D + \nu$.*

The previous best NP-hardness ratio is $\exp(O(\sqrt{\log D}))/D$ by Trevisan [Tre01]. Our Theorem 1.4 is not far from factor $\Omega(\log D)/(D \log\log D)$ approximation algorithms of [Hal02, Hal98]. The best hardness ratio under the UG Conjecture is $O(\log D)^2/D$ [AKS11].

**Theorem 1.5.** *For any $K \geqslant 3$, there is $\nu = o_{n;K}(1)$ such that given a graph with an induced $K$-colorable subgraph of fractional size $1-\nu$, it is* NP*-hard to find an independent set of fractional size $1/2^{K/2} + \nu$.*

The previous best NP-hardness result by Khot and Saket [KS12] has soundness $\exp(-\Omega(\log K)^2)$, for some explicit function in big-Omega.[5] Our result has better soundness than the $4K^2/2^{K/2}$ hardness result for Approximate Graph Coloring in the Lasserre hierarchy [Tul09], albeit lacking "perfect completeness". Almost-2-Coloring has arbitrarily small constant soundness under the UG Conjecture [BK09, BK10]. For $K$-colorable graphs, Khot [Kho01] showed NP-hardness of finding independent set of fractional size $\exp(-\Omega(\log K)^2)$ for sufficiently large $K$. See [DKPS10, KS12] for additional references.

Back to inapproximability of CSPs, consider the large alphabet case. We can choose $C$ of Theorem 1.1 to be an O'Brien predicate of Austrin and Mossel [AM09, Theorem 1.2] (or a Guruswami–Raghavendra predicate [GR08] for a slightly weaker bound).

**Corollary 1.6.** *For any prime power $q$, any integer $k \geqslant 3$, there is $\varepsilon = o_{n;k,q}(1)$ such that it is* NP*-hard to $(1-\varepsilon, q(q-1)k/q^k + \varepsilon)$-decide* Max-$k$-CSP *over alphabet size $q$.*

The previous best NP-hardness result by Engebretsen has soundness $q^{O(\sqrt{q})}/q^k$ [Eng05]. Like [AM09], the soundness in our Corollary 1.6 can be improved to $O(qk/q^k) + \varepsilon$ for infinitely many $k$. Alternatively, one can plug in Håstad predicates (Appendix F) to tighten the hardness ratio for every $k \geqslant q$.

**Corollary 1.7.** *For any integers $q \geqslant 2$ and $k \geqslant q$, it is* NP*-hard (under randomized reduction) to approximate* Max-$k$-CSP *(over alphabet size $q$) beyond $O(qk/q^k)$.*

The best algorithm has a matching approximation ratio $\Omega(qk/q^k)$ when $k \geqslant \Omega(\log q)$ [MM12].

Our (proof of) Theorem 1.1 also reduces soundness of Two-Prover-One-Round-Game, or 2P1R-Game for short.

**Theorem 1.8.** *For any prime power $q$, there is $\varepsilon = o_{n;q}(1)$ such that it is* NP*-hard to $(1-\varepsilon, O(\log q/q)+\varepsilon)$-decide* 2P1R-Game *of alphabet size $q^2$.*

[5] Our Theorem 1.5 is actually stronger, improving [KS12, Theorem 1.1] for all $K \geqslant K_0$, where $K_0 \leqslant 128$.

In terms of alphabet size $R = q^2$, the hardness ratio is $O(\log R/\sqrt{R})$. The previous best inapproximability result by Khot and Safra [KS11] has soundness $4/R^{1/6}$ with alphabet size $R = q^6$. 2P1R-Games with perfect completeness have soundness $1/R^{\Omega(1)}$ [Raz98, Hol09, Rao08]. Hardness of 2P1R-Game is related to hardness of Quadratic-Programming [ABH+05], which was the original goal of Khot and Safra. Even though Theorem 1.8 improves soundness of the former problem, it does not imply any quasi-NP-hardness result for Quadratic-Programming.

## 2. Techniques

Despite progress on conditional results [ST09, GR08, AM09, Rag08], unconditional NP-hardness of Max-$k$-CSP have lagged behind. This is due to limitations of existing composition techniques, which were known since long code was introduced [BGS98, Section 3.4] more than 15 years ago. Here is an illustrative example: In the canonical composition of a *$k$-player* dictator test with a *two-party* Label-Cover instance, each player belongs to one of the two parties (as in Håstad's Max-3-Lin reduction). Replies from the same party may conspire and appear correct, even if the Label-Cover instance has no good assignment. With many more players than parties, soundness will not be $O(k/2^k)$ (assuming replies are boolean). To get around the barrier, previous works [Hås99, ST98, EH08] focused on strengthening the outer verifier and adjusting the composition step (say by creating more parties), as well as improving the inner-verifier analysis. A sequence of works [Tre98, ST98, ST00, EH08] brought soundness down to $2^{O(\sqrt{k})}/2^k$, which is still far from optimal.

In this work, we leapfrog the barrier with a new approach. We view an Additive CSP instance as a $k$-player game, and reduce soundness by a technique we call *direct sum*, which is inspired by XOR-lemmas. Direct sum is like parallel repetition, aiming to reduce soundness by asking each player multiple questions at once, except for receiving only a single answer from each player, namely the sum of answers to individual questions. Direct sum (or XOR-lemma) is invaluable to average-case complexity [GNW11] and central to communication complexity [BBCR10, She12], but has never been useful for hardness of approximation. As it turns out, a natural formulation of a multiplayer XOR-lemma is false (see Remark 5.2), which may explain its absence in the inapproximability literature.

Unable to decrease soundness directly, we instead demonstrate *randomness* of replies. Randomness means lack of correlation. The crucial observation is that correlation never increases with direct sum (Lemma 5.3). It remains to show that, in the Soundness case of a single game, we can isolate any player of our choice, so that his/her reply becomes uncorrelated with the other $k-1$ replies after secret shifting (Theorem 5.4). Then the direct sum of $k$ different games will isolate all players one by one, eliminating any correlation in their shifted replies.

We prove Theorem 5.4 using the canonical composition technique. To analyze the dictator test, we invoke an invariance principle of O'Donnell and Wright [OW12], with a small twist on the reason for matching second moments (Theorem A.1). Unlike previous works, we show invariance for the *correlation* (Definition 4.2) rather than the objective value. For the special case of Samorodnitsky–Trevisan hypergraph predicates, Theorem A.1 has an alternative proof using Gowers uniformity [ST09, Hås09], without the invariance principle (Appendix E).

Our approach also strengthens the NP-hardness ratios for other problems, with simple proofs. We improve the hardness of Two-Prover-One-Round-Game almost as a corollary (Section 8). Our low free-bit PCP also facilitates further reductions, improving hardness of Almost-Coloring (Section 7) and Independent-Set on bounded degree graphs (Appendix D). Despite its simplicity, the reduction to Almost-Coloring requires new ideas.

Previous reductions that bypassed the UG Conjecture for other problems [Kho02a, GRSW12, FGRW09, KM11] started from Khot's Smooth-Label-Cover [Kho02a]. By contrast, our reduction starts from the usual Label-Cover. In fact, the reduction in Theorem 1.1 maps a 3-SAT instance on

$n$ variables to an Additive CSP instance of size $n^{k(1+o_{n;k,|G|}(1))}$. Assuming the Exponential Time Hypothesis [IPZ01] (that deciding 3-SAT on $n$ variables requires $\exp(\Omega(n))$ time), our Theorem 1.1 implies certain Additive CSPs of arity $k$ remain "approximation resistant" against $\exp(n^{(1-o(1))/k})$ time algorithms — a conclusion unlikely to follow from the UG Conjecture because Unique-Games have subexponential time algorithms [ABS10].

## 3. Preliminaries

As usual, let $[q] = \{1, \ldots, q\}$. Denote $\ell^p$-norm of a vector $x \in \mathbb{R}^m$ by $\|x\|_{\ell^p} = (\sum_{i\in[m]} |x_i|^p)^{1/p}$.

Let $\triangle_q = \{x \in \mathbb{R}^q_{\geqslant 0} \mid \|x\|_{\ell^1} = 1\}$ denote the set of probability distributions over $[q]$.

Random variables are denoted by italic boldface letters, such as $\boldsymbol{x}$.

By the size of a constraint satisfaction problem (including Label-Cover), we mean the number of constraints/hyperedges (disregarding weights).

We recall basic facts about characters. A character $\chi$ of a finite abelian group $G$ is a homomorphism from $G$ to the circle group $\mathbb{T}$ of complex numbers of modulus one (under multiplication). The constant 1 function, denoted $\mathbf{1}$, is always a character, known as the trivial character. Any character $\chi$ of a power group $G^k$ has a unique decomposition as a product of characters $\chi_i : G \to \mathbb{T}$ in each coordinate. More precisely,

$$\chi(a_1, \ldots, a_k) = \chi_1(a_1) \ldots \chi_k(a_k) \tag{1}$$

for any $(a_1, \ldots, a_k) \in G^k$.

**Definition 3.1.** Given $j \in [k]$, a character $\chi$ of $G^k$ is *$j$-relevant* if its $j$-th component $\chi_j$ is non-trivial (i.e. not the constant 1 function).

Given two random variables $\boldsymbol{x}$ and $\boldsymbol{y}$ on a set $\Sigma$, their statistical distance $d(\boldsymbol{x}, \boldsymbol{y})$ is the statistical distance of their underlying distributions,

$$d(\boldsymbol{x}, \boldsymbol{y}) = \max_{A\subseteq\Sigma} |\mathbb{P}[\boldsymbol{x} \in A] - \mathbb{P}[\boldsymbol{y} \in A]|.$$

The following bound relating statistical distance and character distance is well known (e.g. [BV10, Claim 33]).[6]

**Proposition 3.2.** *If $|\mathbb{E}[\chi(\boldsymbol{x})] - \mathbb{E}[\chi(\boldsymbol{y})]| \leqslant \varepsilon$ for all characters $\chi$, then $2d(\boldsymbol{x}, \boldsymbol{y}) \leqslant \sqrt{|G| - 1} \cdot \varepsilon$.*

## 4. Additive CSPs

Let $G$ be an abelian group and $C$ a subset of $G^k$. An instance $M = ((V_1, \ldots, V_k), \boldsymbol{Q})$ of Additive-CSP$(C)$ is a distribution over constraints of the form $Q = (v, b)$, where $v = (v_1, \ldots, v_k) \in V_1 \times \cdots \times V_k$ is a $k$-tuple of variables and $b = (b_1, \ldots, b_k) \in G^k$ is a $k$-tuple of shifts. We think of an instance as a $k$-player game: a constraint is a question to the $k$ players, and an assignment $f_i : V_i \to G$ is a strategy of player $i$. Naturally, upon receiving a variable $v_i$, player $i$ responds with $f_i(v_i)$. A constraint $Q = (v, b)$ is satisfied if

$$f(v) - b \triangleq (f_1(v_1) - b_1, \ldots, f_k(v_k) - b_k) \in C.$$

The $k$ players try to satisfy the maximum fraction of constraints. The *value* of the game, denoted by $\mathrm{val}(M)$, is the maximum possible $\mathbb{P}[f(\boldsymbol{v}) - \boldsymbol{b} \in C]$ over $k$ assignments $f_i : V_i \to G$. Note that a game without shifts (equivalently, all shifts are the identity element $0_G$) is trivial, since players have a perfect strategy by always answering $0_G$. The shifts, unknown to the players, make the game challenging.

[6] [BV10] stated the result when $G$ is a finite field, but their proof can be easily adapted for general abelian groups.

**Definition 4.1.** A subset $C$ of $G^k$ is balanced pairwise independent if for every two distinct coordinates $i \neq j \in [k]$ and every two elements $a, b \in G$,

$$\mathbb{P}[\boldsymbol{c}_i = a, \boldsymbol{c}_j = b] = 1/|G|^2,$$

where $\boldsymbol{c} = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_k)$ is a uniformly random element from $C$.

We will often choose $C$ to be a subgroup of $G^k$. Examples of balanced pairwise independent subgroups include dual Hamming codes and Reed–Solomon codes of dimension at least two. Dual Hamming codes have been used to obtain inapproximability results based on the UG Conjecture [ST09] or in the Lasserre hierarchy [Tul09]. Reed–Solomon codes have appeared in low-degree tests.

Let $\mathcal{A}$ be the class of predicates over a balanced pairwise independent subgroup $C \subseteq G^k$ for some $k \geqslant 3$. In other words, these are the predicates satisfying the hypothesis of Theorem 1.1. These are also the predicates currently admitting a direct construction of Lasserre gaps (Appendix G). The class $\mathcal{A}$ is closely related to the bigger class $\mathcal{B}$ of predicates supporting a balanced pairwise independent distribution (possibly without any group structure on the domain), known to be approximation resistant under the UG Conjecture [AM09] and in weaker SDP hierarchies [BGMT12, TW12]. Even though $\mathcal{A}$ is a proper subclass of $\mathcal{B}$ [Tul12], many interesting predicates in $\mathcal{B}$ also belong to $\mathcal{A}$. In particular, all predicates constructed in [ST09, GR08] and O'Brien predicates in [AM09, Theorem 1.2] satisfy our abelian subgroup property.

When there is no prefect strategy, the shifted replies $f(\boldsymbol{v}) - \boldsymbol{b}$ may not have perfect correlation. We measure correlation of the best strategy by the following quantity.

**Definition 4.2.** Given Additive-CSP($C$) instance $M$ and character $\chi : G^k \to \mathbb{T}$, let

$$\|M\|_\chi \triangleq \max|\mathbb{E}\,\chi(f(\boldsymbol{v}) - \boldsymbol{b})| = \max|\mathbb{E}\,\chi(f_1(\boldsymbol{v}_1) - \boldsymbol{b}_1, \ldots, f_k(\boldsymbol{v}_k) - \boldsymbol{b}_k)|,$$

where the maximum is over $k$ assignments $f_i : V_i \to G$.

## 5. Direct sum

To make the game even more difficult for the players, we can take direct sum of instances.

**Definition 5.1.** Let $M = ((V_1, \ldots, V_k), \boldsymbol{Q})$ and $M' = ((V'_1, \ldots, V'_k), \boldsymbol{Q}')$ be Additive-CSP($C$) instances. Their direct sum $M \oplus M'$ is defined as $((V_1 \times V'_1, \ldots, V_k \times V'_k), \boldsymbol{Q} \oplus \boldsymbol{Q}')$. Player $i$ in $M \oplus M'$ receives a pair of variables $(v_i, v'_i) \in V_i \times V'_i$ from $M$ and $M'$.

The random question $\boldsymbol{Q} \oplus \boldsymbol{Q}'$ in $M \oplus M'$ is the direct sum of two independent random questions $\boldsymbol{Q}$ and $\boldsymbol{Q}'$, one from $M$ and the other from $M'$. By the direct sum $Q \oplus Q'$ of two questions $Q = (v, b)$ and $Q' = (v', b')$, we mean sending every player $i$ the variable pair $(v \oplus v')_i \triangleq (v_i, v'_i)$ and receiving a reply $g_i(v_i, v'_i)$. The shifts for $Q \oplus Q'$ is $b + b'$. To wit, $Q \oplus Q' = (v \oplus v', b + b')$.

We expect players' strategy to be independent across the two coordinates, that is $g_i(v_i, v'_i) = (f_i \oplus f'_i)(v_i, v'_i) \triangleq f_i(v_i) + f'_i(v'_i)$, where $f = (f_1, \ldots, f_k)$ is an assignment for $M$ and $f' = (f'_1, \ldots, f'_k)$ an assignment for $M'$. However, the players need not execute such a strategy. Bounding the value of $M \oplus M'$ in terms of the values of $M$ and $M'$ is thus a daunting task.

**Remark 5.2.** Common sense suggests that by repeatedly taking direct sum, the repeated game $M^{\oplus t}$ will have no strategy better than a random one, as long as the original game $M$ has no perfect strategy. More precisely, $\mathrm{val}(M^{\oplus t})$ should converge to the expected value of a random assignment as $t \to \infty$, provided $\|M\|_\chi < 1$ for all non-trivial characters $\chi$ (so that shifted replies are never contained in a proper subgroup of $G^k$). Such a result, if true, may be called a multiplayer XOR-lemma. This result turns out to be true for one- and two-player games, but is *false* for three-player games, as pointed out by Briët, Buhrman, Lee and Vidick [BBLV09]. A counterexample to the three-player XOR-lemma, known as Mermin's game, has a perfect quantum strategy but no perfect classical

strategy. Briët et. al. observed that certain perfect quantum strategies of the repeated game can be "rounded" to a non-trivial classical strategy, via a multilinear Grothendieck inequality. Amazingly, the counterexample was discovered via *quantum* considerations, even though the setting is entirely *classical*.

Fortunately, we can bound the value of $M \oplus M'$ indirectly. As hinted earlier, we instead bound *correlation* of shifted replies. The following lemma shows that correlation can only decrease upon taking direct sum.

**Lemma 5.3.** *For any* Additive-CSP$(C)$ *instances $M$ and $M'$, any character $\chi : G^k \to \mathbb{T}$,*

$$\|M \oplus M'\|_\chi \leqslant \min\{\|M\|_\chi, \|M'\|_\chi\}.$$

*Proof.* Fix arbitrary assignments $f_i : V_i \times V_i' \to G$. The bias is

$$\left| \mathop{\mathbb{E}}_{\boldsymbol{Q}\boldsymbol{Q}'} \chi(f(\boldsymbol{v}, \boldsymbol{v}') - \boldsymbol{b} - \boldsymbol{b}') \right| \leqslant \mathop{\mathbb{E}}_{\boldsymbol{Q}} \left| \mathop{\mathbb{E}}_{\boldsymbol{Q}'} \chi(f(\boldsymbol{v}, \boldsymbol{v}') - \boldsymbol{b} - \boldsymbol{b}') \right|.$$

The RHS is at most $\|M'\|_\chi$, because for every question $\boldsymbol{Q}$ to $M$, we get assignments $g_i^{\boldsymbol{Q}}(v_i') = f_i(\boldsymbol{v}_i, v_i') - \boldsymbol{b}_i$ to $M'$. Since $f_i$'s are arbitrary, we have $\|M \oplus M'\|_\chi \leqslant \|M'\|_\chi$. The same argument also yields $\|M \oplus M'\|_\chi \leqslant \|M\|_\chi$. □

Of course, a simple induction shows that $\|M_1 \oplus \ldots \oplus M_\ell\|_\chi \leqslant \min_{i \in [\ell]} \|M_i\|_\chi$.

The following theorem will be proved in Appendix C, based on a dictator test described in the next section. See Definition 3.1 for $j$-relevant characters.

**Theorem 5.4.** *Let $C$ be a balanced pairwise independent subset of $G^k$. There are $\eta, \delta = o_{n;k,|G|}(1)$ such that for any $j \in [k]$, it is NP-hard to decide the following cases given an* Additive-CSP$(C)$ *instance $M_j$:*

(1) *Completeness:* $\mathrm{val}(M_j) \geqslant 1 - \eta$.
(2) *Soundness: $\|M_j\|_\chi \leqslant \delta$ for all $j$-relevant characters $\chi : G^k \to \mathbb{T}$.*

We can now prove Theorem 1.1. The reduction constructs $k$ instances $M_1, \ldots, M_k$, one for each $j \in [k]$, as guaranteed by Theorem 5.4. The reduction then outputs the direct sum instance $M = M_1 \oplus \ldots \oplus M_k$. If each $M_j$ has size at most $m$, then $M$ has size at most $m^k$, which is polynomial in $m$ for fixed $k$.

*Proof of Theorem 1.1.* **Completeness**. For every $j \in [k]$, let $f^{(j)} = (f_1^{(j)}, \ldots, f_k^{(j)})$ be an optimal assignment tuple for $M_j$. Consider the assignment tuple $g = (g_1, \ldots, g_k)$ for $M$ that is independent across the $k$ component games, that is

$$g_i(v_i^{(1)}, \ldots, v_i^{(k)}) = f_i^{(1)}(v_i^{(1)}) + \cdots + f_i^{(k)}(v_i^{(k)}),$$

Consider a question $\boldsymbol{R} = (\boldsymbol{u}, \boldsymbol{a}) = ((\boldsymbol{v}^{(1)}, \ldots, \boldsymbol{v}^{(k)}), \boldsymbol{b}^{(1)} + \cdots + \boldsymbol{b}^{(k)})$ in $M$. If each of its component question $(\boldsymbol{v}^{(j)}, \boldsymbol{b}^{(j)})$ is satisfied by $f^{(j)}$, then

$$g(\boldsymbol{u}) - \boldsymbol{a} = \textstyle\sum_j f^{(j)}(\boldsymbol{v}^{(j)}) - \boldsymbol{b}^{(j)} \in C,$$

because $C$ is closed under group operations. Hence $g$ also satisfies $\boldsymbol{R}$. Therefore $M$ has value at least $(1 - \eta)^k \geqslant 1 - k\eta$.

**Soundness**. Fix assignments $f_i : V_i \to G$. Let $\chi$ be a non-trivial character of $G^k$. Then $\chi$ is $j$-relevant for some $j \in [k]$, so

$$|\mathbb{E}\, \chi(f(\boldsymbol{v}) - \boldsymbol{b})| \leqslant \|M\|_\chi \leqslant \|M_j\|_\chi \leqslant \delta,$$

using Definition 4.2, Lemma 5.3 and Theorem 5.4. Let $\boldsymbol{a}$ be a uniformly random element from $G^k$, so $\mathbb{E}[\chi(\boldsymbol{a})] = 0$ for any non-trivial character $\chi$. By Proposition 3.2, $f(\boldsymbol{v}) - \boldsymbol{b}$ and $\boldsymbol{a}$ have statistical

distance

$$d(f(\boldsymbol{v}) - \boldsymbol{b}, \boldsymbol{a}) \leqslant \delta \cdot \sqrt{q^k}/2 =: \varepsilon.$$

Therefore

$$\mathbb{P}[f(\boldsymbol{v}) - \boldsymbol{b} \in C] \leqslant \mathbb{P}[\boldsymbol{a} \in C] + \varepsilon = |C|/|G|^k + \varepsilon. \qquad \square$$

## 6. Dictator test

Theorem 5.4 is based on a natural dictator test $T$, which we now describe. Throughout this section, $C$ is a balanced pairwise independent subset of $G$.

Suppose $G$ has $q$ elements. A function $f$ returning a random element from $G$ is considered as having codomain $\triangle_q$.

**Definition 6.1.** Given a string $x \in G^m$, an *$\eta$-noisy copy* is a random string $\dot{\boldsymbol{x}} \in G^m$, so that independently for each $s \in [m]$, $\dot{\boldsymbol{x}}_s = x_s$ with probability $1 - \eta$, and $\dot{\boldsymbol{x}}_s$ is set uniformly at random with probability $\eta$. For a function $f : G^m \to \triangle_q$, we also define the operator $\mathrm{T}_{1-\eta} f(x) = \mathbb{E}[f(\dot{\boldsymbol{x}})]$.

We will compose a $k$-player dictator test with a Label-Cover instance, which is a game between the clause party and the variable party. When assigning players to the parties, we single out player $j$ as the *lonely* player, who is in the variable party, while all other players are in the clause party. As usual, coordinates of a question are segmented into $R$ blocks, each of which has size 1 for the variable party and size $d$ for the clause party.

**Definition 6.2.** A $k$-player, $j$-lonely, $d$-blocked, $\eta$-noise correlated $C$-test is a specified by a block distribution $\mu$ on $G^{d_1} \times \cdots \times G^{d_k}$. Here dimension $d_i$ is $d$ for all $i \neq j$, and $d_j = 1$ for the lonely player $j \in [k]$.

The distribution $\mu$ is the uniform distribution of choosing length-$k$ tuples $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_d$ independently from $C$, conditioned on the tuples agreeing at position $j$. The tuples together represent an element in $G^{d_1} \times \cdots \times G^{d_k}$ because any position other than $j$ gets a sequence of $d$ elements from $G$, while position $j$ gets the common element of the tuples. We call an element from $G^{d_1} \times \cdots \times G^{d_k}$ a block.

Given $R \in \mathbb{N}$, the correlated test $T$ is a random variable $\boldsymbol{z}$ over $G^{d_1 R} \times \cdots \times G^{d_k R}$, which we think of as $R$ blocks. The random variable $\boldsymbol{z}$ is chosen in two steps.

(1) A random "matrix" $\boldsymbol{w} \in G^{d_1 R} \times \cdots \times G^{d_k R}$ is chosen from the product distribution $\mu^{\otimes R}$, so the $R$ blocks are independent of each other, and each block is distributed as $\mu$.
(2) Every column of $\boldsymbol{z}$ is an $\eta$-noisy copy of the corresponding column of $\boldsymbol{w}$.

We call $\boldsymbol{w}$ a matrix because we think of blocks as rows, and a string in $G^{d_i R}$ as the $i$-th column. Entries in this matrix have different lengths: an entry in column $j$ is an element from the base group $G$, while entries elsewhere are from the product group $G^d$.

Since $C$ is balanced pairwise independent, the $(t, i)$-entry of $\boldsymbol{z}$ is uniformly random over $G^{d_i}$. In fact, more is true: Looking only at column $j$ and any other column $i \in [k]$, the marginal distribution is uniform over $G \times G^{dR}$. We call this property "pairwise independence at column $j$". This property is weaker than pairwise independence, because columns $i$ and $i'$ need not be independent. To verify this property, it suffices to consider a block $\boldsymbol{y} \in G^{d_1} \times \cdots \times G^{d_k}$. For any $a \in G$ and $b \in G^d$, the event "$j$-th column of $\boldsymbol{y}$ equals $a$ and $i$-th column of $\boldsymbol{y}$ equals $b$" holds with probability

$$\mathbb{P}[\boldsymbol{y}_{(j)} = a, \boldsymbol{y}_{(i)} = b] = \mathbb{P}[\boldsymbol{y}_{(j)} = a] \cdot \mathbb{P}[\boldsymbol{y}_{(i),1} = b_1] \cdot \cdots \cdot \mathbb{P}[\boldsymbol{y}_{(i),d} = b_d],$$

where we have used pairwise independence of $C$ and conditional independence in the definition of $\mu$.

We measure correlation of players' replies $f_i$ by the Fourier coefficients of $f(\boldsymbol{z})$.

**Definition 6.3.** For a character $\chi : G^k \to \mathbb{T}$, define

$$\mathrm{Bias}_{T,\chi}(f) \triangleq |\mathbb{E}\,\chi(f(\boldsymbol{z}))| = \left|\mathbb{E}\,\chi(f_1(\boldsymbol{z}^{(1)}), \ldots, f_k(\boldsymbol{z}^{(k)}))\right|.$$

Inspired by [OW12], we also consider an uncorrelated version of the test in our analysis.

**Definition 6.4.** The uncorrelated test $T'$ is similar to $T$, only that its block distribution $\mu'$ is modified as follows. A block is chosen exactly the same as in $\mu$, and then the $j$-th entry is re-randomized to be a uniformly random element from $G$, independent of the other entries. Therefore $T'$ is a random variable $\boldsymbol{z}'$ chosen as in Definition 6.2, using $(\mu')^{\otimes R}$ for step (1) in place of $\mu^{\otimes R}$.

Let $\Sigma$ be any set (such as $G$). We will consider Hoeffding decomposition (or Efron–Stein decomposition) for functions $f : \Sigma^m \to \mathbb{R}^q$.

Given $f : \Sigma^m \to \mathbb{R}^q$, define $\|f\|_2^2 = \mathbb{E}_{\boldsymbol{x}\in\Sigma^m}[\|f(\boldsymbol{x})\|_{\ell^2}^2]$. Note that for $f : \Sigma^m \to \triangle_q$, $\|f\|_2^2 \leqslant 1$.

We need the following fact from [Mos10, Definition 2.10].

**Fact 6.5.** *Every function $f : \Sigma^m \to \mathbb{R}^q$ has a unique decomposition $f = \sum_{S\subseteq[m]} f^S$, where the functions $f^S : \Sigma^m \to \mathbb{R}^q$ satisfy*

(1) *$f^S$ depends only on $x_S \triangleq \{x_i\}_{i\in S}$.*
(2) *For any $S \not\subseteq T$ and any $x_T \in \Sigma^T$, $\mathbb{E}[f^S(\boldsymbol{x}) \mid \boldsymbol{x}_T = x_T] = 0$.*

As a result, we get an orthogonal decomposition, so that $\mathbb{E}_{\boldsymbol{x}\in\Sigma^m}\langle f^S(\boldsymbol{x}), f^T(\boldsymbol{x})\rangle = 0$ for any $S \neq T$, where $\langle\cdot,\cdot\rangle$ is the usual inner product in $\mathbb{R}^q$. Therefore $\|f\|_2^2 = \sum_{S\subseteq[m]}\|f^S\|_2^2$.

**Definition 6.6.** The influence and noisy influence of a subset $B \subseteq [m]$ are

$$\mathrm{Inf}_B[f] = \sum_{S:S\cap B\neq\emptyset} \|f^S\|_2^2 \qquad \text{and} \qquad \mathrm{Inf}_B^{(1-\eta)}[f] = \mathrm{Inf}_B[\mathrm{T}_{\sqrt{1-\eta}} f].$$

We also write $\mathrm{Inf}_t[f]$ for $\mathrm{Inf}_{\{t\}}[f]$.

Let $B(t) = \{s \in [dR] \mid (t-1)d < s \leqslant td\}$ denote the set of coordinates associated with block $t \in [R]$.

The following theorem will be proved in Appendix A, by adapting the invariance principle of [OW12]. The theorem says that functions $f_i$ without common influential blocks cannot distinguish between the correlated test $T$ from the uncorrelated version $T'$.

**Theorem 6.7.** *Let $T$ be the test from Definition 6.2. Let $f_i : G^{d_i R} \to \triangle_q$ satisfy*

$$\min\left\{\mathrm{Inf}_t^{(1-\eta)}[f_j], \max_{i\neq j}\{\mathrm{Inf}_{B(t)}^{(1-\eta)}[f_i]\}\right\} \leqslant \tau \quad \forall t \in [R].$$

*Then for any character $\chi : G^k \to \mathbb{T}$,*

$$\mathrm{Bias}_{T,\chi}(f) \leqslant \mathrm{Bias}_{T',\chi}(f) + \delta(q,k,\eta,\tau).$$

*Here $\delta(q,k,\eta,\tau) \leqslant \mathrm{poly}(kq/\eta)\cdot\tau^{\Omega(\eta/\log q)}$.*

We wish to show the term $\mathrm{Bias}_{T',\chi}(f)$ in Theorem 6.7 is negligible. This term is not small in general, if $f_i$ are constant functions. To combat this, we apply the standard trick of folding.

**Definition 6.8.** Given a function $f : G^m \to G$, its folded version $\tilde{f} : G^m \to \triangle_q$ is the function which, upon receiving $x \in G^m$, picks a random $\boldsymbol{y} \in G$ and returns $f(x + (\boldsymbol{y}, \ldots, \boldsymbol{y})) - \boldsymbol{y}$.

The folding shift $\boldsymbol{y}$ is the same shift appearing in a constraint of Additive CSP.

Consider applying the uncorrelated test $T'$ to functions $f_i$'s, where $f_j$ is folded. For any $j$-relevant character $\chi$,

$$\mathrm{Bias}_{T',\chi}(f) = |\mathbb{E}[\chi_j(f_j(\boldsymbol{z}^{(j)}))]\,\mathbb{E}[\chi_J(f_J(\boldsymbol{z}^{(J)}))]|,$$

where $J = [k] \setminus \{j\}$ denotes all players or columns other than $j$. The term $\mathbb{E}[\chi_j(f_j(\boldsymbol{z}^{(j)}))]$ is zero, because folding forces $f_j(\boldsymbol{z}^{(j)})$ to be uniformly random over $G$. Thus

$$\mathrm{Bias}_{T',\chi}(f) = 0.$$

Our preceding discussion implies the following bound on the bias of $T$ for folded functions.

**Theorem 6.9.** *Let $\chi : G^k \to \mathbb{T}$ be a $j$-relevant character. Suppose functions $f_i : G^{d_i R} \to \triangle_q$ satisfy*

$$\min\left\{\mathrm{Inf}_t^{(1-\eta)}[f_j], \max_{i\neq j}\{\mathrm{Inf}_{B(t)}^{(1-\eta)}[f_i]\}\right\} \leqslant \tau \quad \forall t \in [R].$$

*Assume further $f_j$ is folded. Then* $\mathrm{Bias}_{T,\chi}(f) \leqslant \delta(q,k,\eta,\tau) \leqslant \mathrm{poly}(kq/\eta)\cdot\tau^{\Omega(\eta/\log q)}$.

The test $T$ can be turned into an NP-hardness reduction by standard techniques (Appendix C).

## 7. Almost-Coloring

In this section, we prove Theorem 1.5. In our opinion, our proof is simpler than [DKPS10, KS12].

We construct a PCP with small covering parameter apart from small fraction of randomness. Our notion of covering parameter is a variant of Feige and Kilian's [FK98]. We then turn the PCP into an FGLSS graph [FGL+96].

Let $M$ be an Additive-CSP($C$) instance. We say that $M$ has *covering parameter* $K$ if there are $K$ assignments $f^{(1)},\ldots,f^{(K)}$ covering every question $(\boldsymbol{v},\boldsymbol{b})$ of $M$, that is for every $c \in C$, some $f^{(t)}$ satisfies $f^{(t)}(\boldsymbol{v}) - \boldsymbol{b} = c$.

**Proposition 7.1.** *Let $C$ be a balanced pairwise independent subset of $G^k$. There is an* Additive-CSP($C$) *instance $M_C$ with covering parameter $|C|$.*

*Proof.* Let $K = |C|$. Enumerate tuples $c^{(1)},\ldots,c^{(K)}$ in $C$. There is only one question $Q = (v, 0_{G^k})$ in $M_C$. The variable tuple $v = (v_1,\ldots,v_k)$ has components

$$v_i = \left(c_i^{(1)},\ldots,c_i^{(K)}\right) \in G^K.$$

Consider the matching dictator assignment $f^{(t)} = (f_1^{(t)},\ldots,f_k^{(t)})$, where $f_i^{(t)} : G^K \to G$ is given by $f^{(t)}(w) = w_t$. Then $f^{(t)}(v) = c^{(t)}$, and the $K$ matching dictator assignments cover $Q$. □

We recall the definition of an FGLSS graph, specialized for Additive CSPs.

**Definition 7.2.** Given an Additive-CSP($C$) instance $M$, its FGLSS graph $H$ has a vertex $(Q,c)$ for every question $Q = (v,b)$ of $M$ and every $c \in C$. A vertex $(Q,c)$ represents an accepting configuration for $M$. The vertex has weight $w(Q,c) = \mathbb{P}[\boldsymbol{Q} = Q]/|C|$. Two vertices $((v,b),c)$ and $((v',b'),c')$ are connected if their corresponding configurations are conflicting, that is $v_i = v'_i$ and $b_i + c_i \neq b'_i + c'_i$ for some $i \in [k]$.

Denote by $\mathrm{val}(H)$ the maximum fractional size $w(S) \triangleq \sum_{u\in S} w(u)$ of an independent set $S$ in $H$ (a vertex subset $S$ is an independent set if no edge in $H$ has both endpoints in $S$).

The value of $M$ determines the fractional size of a maximum independent set in $H$.

**Proposition 7.3** ([FGL+96, Lemma 3.5])**.** $\mathrm{val}(M) = \mathrm{val}(H)/|C|$.

From now on, $C$ will be a subgroup (not just a subset). Let $M$ be the instance from Theorem 1.1, which either has value at least $1-\eta$ or at most $|C|/|G|^k + \varepsilon$. We construct a PCP $M'$ which is the direct sum $M_C \oplus M$. The output instance is the FGLSS graph $H$ for $M'$.

*Proof of Theorem 1.5.* **Completeness**. There are $K$ assignments $g^{(1)},\ldots,g^{(K)}$ covering $1-\eta$ fraction of questions $(\boldsymbol{v},\boldsymbol{b})$ of $M'$. Indeed, we can take $g^{(t)} = f^{(t)} \oplus f$, where $f^{(t)}$ is a dictator assignment assignment from Proposition 7.1 and $f$ is an assignment satisfying $1-\eta$ questions of $M$. Then for

any question $\boldsymbol{Q} = (\boldsymbol{v}, \boldsymbol{b})$ of $M$ satisfied by $f$ and any question $\boldsymbol{Q}_C$ of $M_C$, the question $\boldsymbol{Q}_C \oplus \boldsymbol{Q}$ is covered by the $g^{(t)}$'s, since the map $c \mapsto c + z$ is a permutation of $C$ whenever $z = f(\boldsymbol{v}) - \boldsymbol{b} \in C$.

In the FGLSS graph $H$, the $K$ assignments $g^{(t)}$'s correspond to $K$ independent sets containing $1 - \eta$ fraction of vertices in total.

**Soundness**. By the proof of Theorem 1.1, $M'$ inherits the soundness property from $M$. By Proposition 7.3, no independent set in $H$ has fractional size more than

$$\frac{1}{|C|}\left(\frac{|C|}{|G|^k} + \varepsilon\right) = \frac{1}{|G|^k} + \frac{\varepsilon}{|C|}.$$

To get the result, fix $C$ to be a hypergraph predicate of Samorodnitsky and Trevisan [ST09]. Then $K \leqslant 2k$, so soundness is $1/2^k \leqslant 1/2^{K/2}$, up to additive $\varepsilon/|C|$. □

## 8. Two-Prover-One-Round-Game

We prove Theorem 1.8 in this section.

Let $M = ((V_1, \ldots, V_k), \boldsymbol{Q})$ be an instance of Additive-CSP($C$). We convert $M$ into a two-prover-one-round game $L_M = ((U, W), \boldsymbol{P})$ between the clause player and the variable player. The variable player receives a variable $u \in U \triangleq V_1 \cup \cdots \cup V_k$, and the clause player receives a clause $Q \in W \triangleq \mathrm{supp}(\boldsymbol{Q}) \subseteq (V_1 \times \cdots \times V_k) \times G^k$. In the new game $L_M$, a clause $\boldsymbol{Q} = (\boldsymbol{v}, \boldsymbol{b})$ is chosen from $M$, and a variable $\boldsymbol{u}$ is chosen uniformly at random from $\boldsymbol{v} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$, so that $\boldsymbol{u} = \boldsymbol{v}_{\boldsymbol{j}}$ for a random index $\boldsymbol{j} \in [k]$. The clause player responds with a satisfying assignment $g(\boldsymbol{Q}) \in C$ to $\boldsymbol{Q}$; the variable player responds with an assignment $f(\boldsymbol{u}) \in G$ to $\boldsymbol{u}$. The players win if their replies agree,

$$g(\boldsymbol{Q})_{\boldsymbol{j}} = f(\boldsymbol{u}) - \boldsymbol{b}_{\boldsymbol{j}}.$$

Then $L_M$ is a two-prover-one-round projection game[7], with alphabet size $|C|$.

Consider the instance $L_M$ when $M$ is the output instance of Theorem 1.1. It is straightforward to show that $\mathrm{val}(L_M) \geqslant 1 - \varepsilon$ if $\mathrm{val}(M) \geqslant 1 - \varepsilon$. For the Soundness case, we again consider randomness in variable player's reply. Define $h(v) = (f(v_1), \ldots, f(v_k)) \in G^k$ for $v \in V_1 \times \cdots \times V_k$.

Recall the multiplicative Chernoff bound (e.g. [SSS95, Theorem 2(I)]).

**Proposition 8.1.** *Suppose $\boldsymbol{Y}$ is a sum of independent $\{0,1\}$-valued random variables. Let $\mu = \mathbb{E}[\boldsymbol{Y}]$. Then for any $\lambda \geqslant 1$,*

$$\mathbb{P}[\boldsymbol{Y} \geqslant (1+\lambda)\mu] \leqslant \exp(-\lambda\mu/3).$$

*Proof of Theorem 1.8.* **Soundness**. For a fixed question $Q = (v, b)$, the winning probability (over the random index $\boldsymbol{j}$) is precisely

$$\mathrm{agr}(g(Q), h(v) - b) \triangleq \mathbb{P}[g(Q)_{\boldsymbol{j}} = (h(v) - b)_{\boldsymbol{j}}].$$

We can approximate the random variable $h(\boldsymbol{v}) - \boldsymbol{b}$ with a random variable $\boldsymbol{a}$ that is uniform over $G^k$. Then for any potential answer $c \in C \subseteq G^k$ of the clause player, the fractional agreement $\mathrm{agr}(c, \boldsymbol{a})$ is a random variable $\boldsymbol{Y}/k$, where $\boldsymbol{Y}$ is Binomial with parameters $k$ and $1/q$. Write $t = O(\log(q|C|)) \cdot k/q$, and assume $k \geqslant q$. By multiplicative Chernoff bound (Proposition 8.1),

$$\mathbb{P}[\mathrm{agr}(c, \boldsymbol{a}) \geqslant t/k] = \mathbb{P}[\boldsymbol{Y} \geqslant t] \leqslant 1/q|C|.$$

It follows by union bound that

$$\mathbb{P}[\exists c \in C,\ \mathrm{agr}(c, \boldsymbol{a}) \geqslant t/k] \leqslant 1/q.$$

[7] That is, the reply of the clause player determines the *only* correct reply for the variable player.

Therefore $\mathrm{val}(L_M)$ is bounded by

$$\begin{aligned}\mathbb{E}[\mathrm{agr}(g(\boldsymbol{Q}), h(\boldsymbol{v}) - \boldsymbol{b})] &\leqslant t/k + \mathbb{P}[\exists c \in C,\ \mathrm{agr}(c, h(\boldsymbol{v}) - \boldsymbol{b}) \geqslant t/k] \\ &\leqslant O(\log(q|C|)/q) + 1/q + d(h(\boldsymbol{v}) - \boldsymbol{b}, \boldsymbol{a}).\end{aligned}$$

As in the proof of Theorem 1.1, the statistical distance $d(h(\boldsymbol{v}) - \boldsymbol{b}, \boldsymbol{a}) = o_{n;k,|G|}(1)$ and is negligible.

To bound the first term, we can choose $k = q$ and $C$ to be Reed–Solomon code over $\mathbb{F}_q$ of dimension two, so that $|C| = q^2$. □

## 9. Open problems

Our PCP in Corollary 1.3 has optimal query complexity, but lacks perfect completeness. Getting perfect completeness is an interesting open problem. Our PCP has large blow-up in size due to the use of long code, while previous query-efficient PCP has a smaller variant using Hadamard code [Kho01]. Getting a small PCP with optimal query-efficiency is another natural problem (it requires something different from Hadamard code [ST09, Lov08]).

**Acknowledgements.** I am indebted to the following people for numerous helpful discussions and continual encouragement: Siu Man Chan, Ilias Diakonikolas, Elchanan Mossel, Prasad Raghavendra, Grant Schoenebeck, Piyush Srivastava, Luca Trevisan, Madhur Tulsiani, Satish Rao, and Yi Wu. I am grateful to Luca Trevisan for pointing out the composition limitations, and bringing [Hås09] to my attention. I also thank Urmila Mahadev and Thomas Vidick for sharing their expertise on quantum multiplayer games. I sincerely thank Johan Håstad for allowing me to include his predicates in this paper, and Subhash Khot for providing the full version of [KS12]. I thank Thomas Watson for many suggestions that greatly enhance the presentation of the paper. Special thanks to Piyush Srivastava for collaboration on related aspects of this research.

The author was supported by Elchanan Mossel's NSF award DMS-1106999 and DOD ONR grant N000141110140 and by Satish Rao's NSF award CCF-1118083.

## Appendix A. Invariance principle

Building on [MOO10, Mos10, OW12], we first prove an invariance theorem when functions have no influential coordinates. Our proof (and presentation) closely follows O'Donnell and Wright's [OW12, Section A]. To justify the matching second moments condition, we need "pairwise independence at column $j$" (see Section 6).

As before, denote by $q$ the size of $G$. Let $B_i(t) \subseteq [d_i R]$ denote the set of coordinates in block $t \in [R]$, so $B_j(t) = \{t\}$ and $B_i(t) = \{s \in [dR] \mid (t-1)d < s \leqslant td\}$ for $i \neq j$.

Write $F_i \triangleq \mathrm{T}_{1-\eta} f_i$. We will use vector notations repeatedly, for example

$$F(z) \triangleq (F_1(z^{(1)}), \ldots, F_k(z^{(k)})).$$

For a function $f_i : G^{d_i R} \to \triangle_q$, introduce the operators

$$\mathrm{L}_t f_i = \sum_{S : S \cap B_i(t) \neq \emptyset} f_i^S \qquad \mathrm{E}_t f_i = \sum_{S : S \cap B_i(t) = \emptyset} f_i^S.$$

We also apply these operators to $F$ component-wise, so $\mathrm{L}_t F \triangleq (\mathrm{L}_t F_1, \ldots, \mathrm{L}_t F_k)$.

Our invariance principle shows that low-influence functions cannot distinguish between the correlated random variable $\boldsymbol{z}$ (Definition 6.2) and its uncorrelated version $\boldsymbol{z}'$ (Definition 6.4).

**Theorem A.1** (Invariance principle). *Let $\Psi : \triangle_q^k \to \mathbb{R}$ be a $C^3$ function satisfying*

$$|\partial^{(\beta)} \Psi| \leqslant A \quad \forall |\beta| = 3.$$

*Suppose $f_i : G^{d_i R} \to \triangle_q$ are functions satisfying*

$$\max_{i\in[k]}\{\mathrm{Inf}^{(1-\eta)}_{B_i(t)}[f_i]\} \leqslant \tau \quad \forall t \in [R].$$

*Then*

$$|\mathbb{E}[\Psi(F(\boldsymbol{z}))] - \mathbb{E}[\Psi(F(\boldsymbol{z}'))]| \leqslant 18(kq)^3 A\tau^{c(q,\eta)}c_\eta,$$

*where $c_\eta = \frac{2}{\eta}\ln\left(\frac{1}{\eta}\right)$ and $c(q,\eta) = \Theta(\eta/\log q)$.*

As in all Lindeberg-style proofs of the invariance principle, we consider random variables that are hybrids of $\boldsymbol{z}$ and $\boldsymbol{z}'$. For $t = 0, \ldots, R$, the $t$-th hybrid is $\boldsymbol{z}^{(t)} = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_t, \boldsymbol{z}'_{t+1}, \ldots, \boldsymbol{z}'_R)$, where every $\boldsymbol{z}_s$ is distributed according to $\mu$ and every $\boldsymbol{z}'_s$ according to $\mu'$, independently. Recall that we think of each $\boldsymbol{z}_t$ as a row of the "matrix" $\boldsymbol{z}^{(t)}$.

*Proof.* We bound

$$|\mathbb{E}[\Psi(F(\boldsymbol{z}))] - \mathbb{E}[\Psi(F(\boldsymbol{z}'))]| \leqslant \sum_{t\in[R]} \mathrm{err}_t,$$

where the error for switching from $\boldsymbol{z}^{(t-1)}$ to $\boldsymbol{z}^{(t)}$ is

$$\mathrm{err}_t \triangleq \left|\mathbb{E}[\Psi(F(\boldsymbol{z}^{(t-1)}))] - \mathbb{E}[\Psi(F(\boldsymbol{z}^{(t)}))]\right|.$$

Let $\boldsymbol{F} = \mathrm{E}_t F(\boldsymbol{z}^{(t)})$, $\boldsymbol{H} = \mathrm{L}_t F(\boldsymbol{z}^{(t-1)})$, and $\boldsymbol{K} = \mathrm{L}_t F(\boldsymbol{z}^{(t)})$. Note that $\boldsymbol{F}$ is independent of block $t$, as guaranteed by the Hoeffding decomposition (Fact 6.5).

Apply Taylor's theorem to $\Psi$, centered at $\boldsymbol{F}$, out to the third partial derivatives:

$$\Psi(x+y) = \sum_{|\beta|<3} \frac{\Psi^{(\beta)}(x)}{\beta!}y^\beta + \sum_{|\beta|=3} r_\beta(x,y)y^\beta,$$

where the remainder term satisfies

$$|r_\beta(x,y)| \leqslant \frac{|\beta|A}{\beta!}.$$

We now show that the linear part cancels, and so does the quadratic part. That is, for any multi-index $\beta \in (\mathbb{N}^q)^k$, $|\beta| \leqslant 2$,

$$\mathbb{E}\left[\frac{\Psi^{(\beta)}(\boldsymbol{F})}{\beta!}\boldsymbol{H}^\beta\right] = \mathbb{E}\left[\frac{\Psi^{(\beta)}(\boldsymbol{F})}{\beta!}\boldsymbol{K}^\beta\right].$$

Since $\Phi^{(\beta)}(\boldsymbol{F})/\beta!$ is independent of the value of $\boldsymbol{z}_t$ or $\boldsymbol{z}'_t$, we only need to verify that when everything except $\boldsymbol{z}_t$ and $\boldsymbol{z}'_t$ is fixed, $\boldsymbol{H}^\beta$ and $\boldsymbol{K}^\beta$ are distributed identically.

When $|\beta| = 0$, the statement is trivial. Split $\beta$ into $(\beta_j, \beta_J)$, where $\beta_j$ contains all $q$ indices $(j,l) \in [k]\times[q]$ related to the $j$-th column, and $\beta_J$ contains the rest. If $|\beta_j| = 0$, then $\boldsymbol{H}^\beta$ is independent of the $j$-th column of $\boldsymbol{z}'_t$, and likewise $\boldsymbol{K}^\beta$ is independent of the $j$-th column of $\boldsymbol{z}_t$. So the expectations equal because $\boldsymbol{z}_t$ and $\boldsymbol{z}'_t$ have identical marginal distributions on other columns. If $|\beta_J| = 0$, the argument is similar, and now $\boldsymbol{H}^\beta$ and $\boldsymbol{K}^\beta$ are independent of the $i$-th column of $\boldsymbol{z}'_t$ or $\boldsymbol{z}_t$, for all $i \neq j$. Ths covers the cases $|\beta| \leqslant 1$, and a portion of the case of $|\beta| = 2$.

What remains is $|\beta_j| = |\beta_J| = 1$. Suppose $(i,l)$ is the index such that $\beta(i,l) = 1$ and $i \neq j$. Then $\boldsymbol{H}^\beta$ can only depend on the $i$- and $j$-th columns, and likewise for $\boldsymbol{K}^\beta$. Since $\boldsymbol{z}_t$ and $\boldsymbol{z}'_t$ have identical marginals on columns $j$ and $i$ (by pairwise independence at column $j$), the expectations agree.

We thus have the bound

$$\mathrm{err}_t \leqslant \sum_{|\beta|=3} \frac{3A}{\beta!}(\mathbb{E}[|\boldsymbol{H}|^\beta] + \mathbb{E}[|\boldsymbol{K}|^\beta]).$$

There are $\leqslant (kq)^3$ cubic error terms, one for each of $\boldsymbol{H}$ and $\boldsymbol{K}$. Suppose $|\beta_j| = 1$ and $|\beta_J| = 2$ (other cases are analogous). In this case $\beta_j$ selects one component $l_1$ from $(\boldsymbol{H})_j$ and two components, $l_2$ and $l_3$, from $(\boldsymbol{H})_{i_2}$ and $(\boldsymbol{H})_{i_3}$.

$$\begin{aligned}\mathbb{E}[|\boldsymbol{H}|^\beta] &= \mathbb{E}[|(\boldsymbol{H})_{j,l_1}(\boldsymbol{H})_{i_2,l_2}(\boldsymbol{H})_{i_3,l_3}|] \\ &\leqslant \mathbb{E}[|(\boldsymbol{H})_{j,l_1}|^3]^{1/3}\,\mathbb{E}[|(\boldsymbol{H})_{i_2,l_2}|^3]^{1/3}\,\mathbb{E}[|(\boldsymbol{H})_{i_3,l_3}|^3]^{1/3},\end{aligned}$$

by Hölder's inequality. Effectively, this breaks the dependence between the $(\boldsymbol{H})_{i,l}$'s. Summarizing,

$$\mathbb{E}[|\boldsymbol{H}|^\beta] \leqslant \mathbb{E}[|\boldsymbol{H}|^3]^{\beta/3}.$$

We now consider the contribution of a single factor $\mathbb{E}[|(\boldsymbol{H})_{i,l}|^3]$. Since $(\boldsymbol{H})_{i,l}$ takes values in $[0,1]$, we have $\mathbb{E}[|(\boldsymbol{H})_{i,l}|^3] \leqslant \mathbb{E}[|(\boldsymbol{H})_{i,l}|^r]$ for $r \leqslant 3$.

$$\mathbb{E}[|(\boldsymbol{H})_{i,l}|^r] = \mathbb{E}[|\mathrm{T}_{\sqrt{1-\eta}}\mathrm{T}_{\sqrt{1-\eta}}\mathrm{L}_t(f_i)_l|^r] \leqslant \mathbb{E}[|\mathrm{T}_{\sqrt{1-\eta}}\mathrm{L}_t(f_i)_l|^2]^{r/2}$$

where $r = 2 + 2c$ and $c = c(q,\eta)$ is a small number so that $\mathrm{T}_{\sqrt{1-\eta}}$ is $(2,r)$-hypercontractive. By [Ole03, Wol07], we can take $c = \Theta(\eta/\log q)$. Now

$$\mathbb{E}[|\mathrm{T}_{\sqrt{1-\eta}}\mathrm{L}_t(f_i)_l|^2]^{r/2} = \mathrm{Inf}^{(1-\eta)}_{B_i(t)}[(f_i)_l]^{1+c}.$$

To simplify notation, define $\mathrm{Inf}^{(1-\eta)}_{B(t)}[f]^\beta = \prod_i \mathrm{Inf}^{(1-\eta)}_{B_i(t)}[(f_i)_l]^{\beta_{il}}$. Then we have just shown

$$\mathbb{E}[|\boldsymbol{H}|^\beta] \leqslant \left(\mathrm{Inf}^{(1-\eta)}_{B(t)}[f]^\beta\right)^{(1+c)/3}.$$

We consider the contribution of the power $1/3$ and $c/3$ separately. The contribution from the power $1/3$ is

$$\mathrm{Inf}^{(1-\eta)}_{B(t)}[f]^{\beta/|\beta|} \leqslant (\beta/|\beta|)\cdot \mathrm{Inf}^{(1-\eta)}_{B(t)}[f].$$

by AM-GM inequality.

Therefore

$$\mathbb{E}[|\boldsymbol{H}|^\beta] \leqslant \tau^c \frac{\beta}{|\beta|}\cdot \mathrm{Inf}^{(1-\eta)}_{B(t)}[f].$$

There is a similar inequality for $\boldsymbol{K}$.

Consequently, the total error is at most

$$\sum_{t\in[R]} \mathrm{err}_t \leqslant 6(kq)^3 A\tau^c \max_{|\beta|=3}\frac{\beta}{\beta!}\cdot\sum_{t=1}^R \mathrm{Inf}^{(1-\eta)}_{B(t)}[f] \leqslant 18(kq)^3 A\tau^c c_\eta,$$

where the last inequality comes from the following bound on total influence.

**Fact A.2** ([OW12, Fact A.2]). *Let $c_\eta = \frac{2}{\eta}\ln(\frac{1}{\eta})$. Then for any $d, R \in \mathbb{N}$ and any $f : \Sigma^{dR} \to \mathbb{R}$,*

$$\sum_{t=1}^R \mathrm{Inf}^{(1-\eta)}_{B(t)}[f] \leqslant c_\eta \|f\|_2^2. \qquad \square$$

We now relax the influence assumption, allowing one party to have larger influence. We need the following proposition to take care of such coordinates, which is inspired by [Mos10, Lemma 6.7].

**Definition A.3.** A function $\Psi : \triangle_q^k \to \mathbb{R}^m$ is $L$-Lipschitz (in each coordinate) if for every $x, y \in \triangle_q^k$ and $l \in [k]$ satisfying $x_i = y_i \in \triangle_q$ for all $i \neq l$, we have

$$\|\Psi(x) - \Psi(y)\|_{\ell^2} \leqslant L\|x_l - y_l\|_{\ell^2}.$$

**Proposition A.4.** *Let $\Psi : \triangle_q^k \to \mathbb{R}^m$ be L-Lipschitz. Let $F_i : G^{d_i R} \to \triangle_q$ be functions, and $j \in [k]$. Suppose for the t-th block, we have $\mathrm{Inf}_t[F_j] \leqslant \tau$ or $\max_{i \neq j}\{\mathrm{Inf}_{B_i(t)}[F_i]\} \leqslant \tau$. Then*

$$\left\| \mathbb{E}[\Psi(F(\boldsymbol{z}^{(t-1)}))] - \mathbb{E}[\Psi(F(\boldsymbol{z}^{(t)}))] \right\|_{\ell^2} \leqslant 2(k-1)L\sqrt{\tau}.$$

*Proof.* Consider the case $\max_{i \neq j}\{\mathrm{Inf}_{B_i(t)}[F_i]\} \leqslant \tau$. Define $H_j \triangleq F_j$, and $H_i \triangleq \mathrm{E}_t F_i$ for $i \neq j$. We bound

$$\left\| \mathbb{E}[\Psi(F(\boldsymbol{z}^{(t)}))] - \mathbb{E}[\Psi(H(\boldsymbol{z}^{(t)}))] \right\|_{\ell^2} \leqslant \mathbb{E}[\|\Psi(F(\boldsymbol{z}^{(t)})) - \Psi(H(\boldsymbol{z}^{(t)}))\|_{\ell^2}].$$

To bound the RHS, we change $F_i$ into $H_i$ one-by-one for every $i \in [k]$. Let $\boldsymbol{w}_{(h)}$ denote the $h$-th column of $\boldsymbol{z}^{(t)}$, and define the hybrid functions

$$\boldsymbol{H}^{(i)} = (F_1(\boldsymbol{w}_{(1)}), \ldots, F_i(\boldsymbol{w}_{(i)}), H_{i+1}(\boldsymbol{w}_{(i+1)}), \ldots, H_k(\boldsymbol{w}_{(k)})).$$

The error incurred for $i \neq j$ is

$$\mathbb{E}[\|\Psi(\boldsymbol{H}^{(i)}) - \Psi(\boldsymbol{H}^{(i-1)})\|_{\ell^2}] \leqslant L\,\mathbb{E}[\|F_i(\boldsymbol{w}_{(i)}) - H_i(\boldsymbol{w}_{(i)})\|_{\ell^2}].$$

We further bound

$$\mathbb{E}[\|F_i(\boldsymbol{w}_{(i)}) - H_i(\boldsymbol{w}_{(i)})\|_{\ell^2}] \leqslant \mathbb{E}[\|F_i(\boldsymbol{w}_{(i)}) - H_i(\boldsymbol{w}_{(i)})\|_{\ell^2}^2]^{1/2} \leqslant \sqrt{\tau},$$

by Jensen's inequality and the fact that $\mathbb{E}[\|F_i(\boldsymbol{w}_{(i)}) - H_i(\boldsymbol{w}_{(i)})\|_{\ell^2}^2] = \mathrm{Inf}_{B_i(t)}[F_i]$. For $i = j$, there is no error incurred. Summarizing,

$$\left\| \mathbb{E}[\Psi(F(\boldsymbol{z}^{(t)}))] - \mathbb{E}[\Psi(H(\boldsymbol{z}^{(t)}))] \right\|_{\ell^2} \leqslant L(k-1)\sqrt{\tau}.$$

A similar inequality holds with $\boldsymbol{z}^{(t)}$ replaced with $\boldsymbol{z}^{(t-1)}$.

We will be done (for the case we started off) if we can show

$$\mathbb{E}[\Psi(H(\boldsymbol{z}^{(t-1)})] = \mathbb{E}[\Psi(H(\boldsymbol{z}^{(t)}))].$$

This equality holds because $H$ does not depend on the $(t,i)$-entries for any $i \neq j$, and the variables $\boldsymbol{z}^{(t-1)}$ and $\boldsymbol{z}^{(t)}$ have the same joint marginal on all the other entries.

The other case $\mathrm{Inf}_t[F_j] \leqslant \tau$ is analogous. Instead we define $H_i = F_i$ for $i \neq j$, and define $H_j = \mathrm{E}_t F_j$. We get a better bound

$$\left\| \mathbb{E}[\Psi(F(\boldsymbol{z}^{(t)}))] - \mathbb{E}[\Psi(H(\boldsymbol{z}^{(t)}))] \right\|_{\ell^2} \leqslant L\sqrt{\tau}$$

for this case, since the error comes only from changing $F_j$ into $H_j$. The equality $\mathbb{E}[\Psi(H(\boldsymbol{z}^{(t-1)})] = \mathbb{E}[\Psi(H(\boldsymbol{z}^{(t)}))]$ also holds, since $H_j$ does not depend on the $(t,j)$-entry, and the variables $\boldsymbol{z}^{(t-1)}$ and $\boldsymbol{z}^{(t)}$ have the same joint marginal on all other entries. □

**Theorem A.5.** *Consider the same setting as in Theorem A.1, except the influence condition becomes*

$$\min\left\{ \mathrm{Inf}_t^{(1-\eta)}[f_j], \max_{i \neq j}\{\mathrm{Inf}_{B_i(t)}^{(1-\eta)}[f_i]\} \right\} \leqslant \tau \quad \forall t \in [R].$$

*Also assume $\Psi$ is L-Lipschitz. Then*

$$|\mathbb{E}[\Psi(F(\boldsymbol{z}))] - \mathbb{E}[\Psi(F(\boldsymbol{z}'))]| \leqslant O(kq)^3 A \tau^{c(q,\eta)/4} c_\eta + O(k^2) L \tau^{1/4} c_\eta.$$

*Proof.* Call $t$ non-influential if $\max_{i \in [k]}\{\mathrm{Inf}_{B_i(t)}^{(1-\eta)}[f_i]\} \leqslant \tau_0 \triangleq \tau^{1/4}$, otherwise $t$ is half-influential.

By Fact A.2, there are at most $kc_\eta/\tau_0$ half-influential blocks. By Proposition A.4, these blocks together contribute an error bounded by

$$2L(k-1)\sqrt{\tau} \cdot kc_\eta/\tau_0 \leqslant 2L(k-1)kc_\eta\tau^{1/4}.$$

For non-influential blocks, we bound their error by the analysis in Theorem A.1. □

We will apply Theorem A.5 with the function $\Psi$ defined in Appendix B. We get the following theorem.

**Theorem 6.7.** *Let $T$ be the test from Definition 6.2. Let $f_i : G^{d_i R} \to \triangle_q$ satisfy*

$$\min\left\{\mathrm{Inf}_t^{(1-\eta)}[f_j], \max_{i\neq j}\{\mathrm{Inf}_{B(t)}^{(1-\eta)}[f_i]\}\right\} \leqslant \tau \quad \forall t \in [R].$$

*Then for any character $\chi : G^k \to \mathbb{T}$,*

$$\mathrm{Bias}_{T,\chi}(f) \leqslant \mathrm{Bias}_{T',\chi}(f) + \delta(q,k,\eta,\tau).$$

*Here $\delta(q,k,\eta,\tau) \leqslant \mathrm{poly}(kq/\eta)\cdot\tau^{\Omega(\eta/\log q)}$.*

## Appendix B. Derivative bounds

We will apply the invariance principle to the function $\Psi = \Psi_\chi : \triangle_q^k \to \mathbb{R}^2$ encoding a character $\chi : G^k \to \mathbb{T}$ (or rather, to each of the two output coordinates of $\Psi$). Here $\Psi$ is defined to be the multilinearized version of $\chi$, with the two output coordinates of $\Psi$ being the real and imaginary parts of $\chi$. Therefore

$$\Psi(x) = \sum_{a_1,\ldots,a_k\in G} x_{1,a_1}\ldots x_{k,a_k}\mathcal{C}(\chi(a_1,\ldots,a_k)), \tag{2}$$

where $\mathcal{C} : \mathbb{C} \to \mathbb{R}^2$ is the bijection between the complex plane and $\mathbb{R}^2$ given by $s+it \mapsto (s,t)$ for $s,t\in\mathbb{R}$.

We now bound the third derivatives of $\Psi$. For a multi-index $\beta \in (\mathbb{N}^q)^k$ with $|\beta| > 0$, the derivative $\partial^{(\beta)}\Psi(x)$ vanishes unless $\beta$ selects components from different parts. In other words, if $|\beta| = 3$ then the derivative is zero unless there are components $(i_1,a_1),(i_2,a_2),(i_3,a_3) \in [k]\times[q]$ with distinct $i_j$'s such that $\beta(i_j,a_j) = 1$. Without loss of generality, assume $i_j = j$ for $j = 1,2,3$. Then

$$\partial^{(\beta)}\Psi(x) = \sum_{a_4,\ldots,a_k\in G} x_{4,a_4}\ldots x_{k,a_k}\mathcal{C}(\chi(a_1,\ldots,a_k)).$$

This means $\partial^{(\beta)}\Psi(x)$ is precisely the $\chi$-Fourier coefficient of the product distribution $e_{a_1}\otimes e_{a_2}\otimes e_{a_3}\otimes x_4\otimes\ldots\otimes x_k$. Let $\boldsymbol{y}$ be a random variable with this product distribution, then $\|\partial^{(\beta)}\Psi(x)\|_{\ell^2} = |\mathbb{E}[\chi(\boldsymbol{y})]| \leqslant 1$.

We also need to bound the Lipschitz constant of $\Psi$. We will show that

$$\|\Psi(x_1,\ldots,x_k) - \Psi(y_1,\ldots,y_k)\|_{\ell^2} \leqslant \|x_l - y_l\|_{\ell^1}.$$

for every $l\in[k]$, every $x_1,\ldots,x_k,y_1,\ldots,y_k \in \triangle_q$ such that $x_i = y_i$ for all $i\neq l$. It then follows that $\Psi$ is $\sqrt{q}$-Lipschitz, thanks to the Cauchy–Schwarz inequality $\|x\|_{\ell^1} \leqslant \sqrt{q}\|x\|_{\ell^2}$ for $x\in\mathbb{R}^q$. Write $z_i = x_i$ for $i\neq l$ and $z_l = x_l - y_l$. Expand

$$\|\Psi(x_1,\ldots,x_k) - \Psi(y_1,\ldots,y_k)\|_2 = \left|\sum_{a_1,\ldots,a_k} z_{1,a_1}\ldots z_{k,a_k}\chi(a_1,\ldots,a_k)\right| = \prod_{i\in[k]}\left|\sum_a z_{i,a}\chi_i(a)\right|,$$

where the characters $\chi_i : G\to\mathbb{T}$ are the components of $\chi$. Then

$$\left|\sum_a z_{i,a}\chi_i(a)\right| \leqslant \sum_a |z_{i,a}| = \|z_i\|_{\ell^1}.$$

This yields the desired bound since $\|z_i\|_{\ell^1} = 1$ for $i\neq l$.

## Appendix C. Hardness reduction

In this section, we prove Theorem 5.4. Our reduction closely follows those in previous works [Hås01, OW12], with one notable difference to Håstad's reduction: we allow different strategies from different players, so our output instance is $k$-partite. We will need this feature for the direct sum operation.

As usual, we will reduce from Label-Cover $\mathsf{LC}_{R,dR}$. An instance of $\mathsf{LC}_{R,dR}$ is a weighted bipartite graph $((U,V),\boldsymbol{e})$. Vertices from $U$ are variables with domain $[R]$, and vertices from $V$ are variables with domain $[dR]$. Every edge $\boldsymbol{e}=(\boldsymbol{u},\boldsymbol{v})\in U\times V$ has an associated $d$-to-1 map $\pi_{\boldsymbol{e}}:[dR]\to[R]$. Given an assignment $A:U\to[R],V\to[dR]$, the constraint on $\boldsymbol{e}$ is satisfied if $\pi_{\boldsymbol{e}}(A(\boldsymbol{v}))=A(\boldsymbol{u})$.

The following theorem of Moshkovitz and Raz asserts hardness of Label-Cover [MR10, DH10].

**Theorem C.1.** *For some $0<c<1$ and some $g(n)=\Omega(\log n)^c$, for any $\sigma=\sigma(n)\geqslant\exp(-g(n))$, there are $d,R\leqslant\exp(\mathrm{poly}(1/\sigma))$ such that the problem of deciding a* 3-SAT *instance with $n$ variables can be Karp-reduced in $\mathrm{poly}(n)$ time to the problem of $(1,\sigma)$-deciding a $\mathsf{LC}_{R,dR}$ instance $L$ of size $n^{1+o(1)}$. Furthermore, $L$ is a bi-regular bipartite graph with left- and right-degrees $\mathrm{poly}(1/\sigma)$.*

Our reduction from Label-Cover to Additive-CSP$(C)$ produces an instance that is a $k$-uniform, $k$-partite hypergraph on the vertex set $V_1\cup\cdots\cup V_k$. The $j$-th vertex set $V_j$ is $U\times G^R$, obtained by replacing each vertex in $U$ with a $G$-ary hypercube. Any other vertex set $V_i$ is a copy of $V\times G^{dR}$. All vertices are variables with domain $G$ (that has $q$ elements). We think of an assignment to variables in $u\in V_j$ as a function $f_{j,u}:G^R\to G$, and likewise an assignment to variables in $v\in V_i$ as a function $g_{i,v}:G^{dR}\to G$.

For every constraint $\boldsymbol{e}=(\boldsymbol{u},\boldsymbol{v})$, the reduction introduces $C$-constraints on the (folded versions[8] of) assignments $f_{j,\boldsymbol{u}}$ and $g_{i,\boldsymbol{v}}$, as specified by a dictator test $T$ under blocking map $\pi_{\boldsymbol{e}}$.

The following theorem, together with Theorem C.1, implies Theorem 5.4.

**Theorem C.2.** *Let $T$ be the test from Definition 6.2. Suppose $\sigma\leqslant\delta\eta^2\tau^2/(k-1)$, where $\tau=\tau(q,k,\eta,\delta)=(\eta\delta/kq)^{O(\log q)/\eta}$ is chosen to satisfy $\delta\leqslant\mathrm{poly}(kq/\eta)\cdot\tau^{\Omega(\eta/\log q)}$ in Theorem 6.9.*

*The problem of $(1,\sigma)$-deciding a $\mathsf{LC}_{R,dR}$ instance $L$ can be Karp-reduced to the problem of deciding the following cases given an* Additive-CSP$(C)$ *instance $M_j$:*

1. *Completeness:* $\mathrm{val}(M_j)\geqslant 1-\eta$.
2. *Soundness: $\|M_j\|_\chi\leqslant 2\delta$ for all $j$-relevant characters $\chi$.*

*Further, if $L$ has size $m$, then $M_j$ has size $m\cdot q^{O(kdR)}$.*

*Proof.* **Completeness**. Let $A$ be an assignment to the Label-Cover instance with value 1. Consider the assignment $f_{j,u}(z)=z_{A(u)}$ and $g_{i,v}(z)=z_{A(v)}$. These are matching dictators since $A$ satisfies the constraint on $\boldsymbol{e}$. Therefore for every $\boldsymbol{e}$, at least $1-k\eta$ fraction of the associated $C$-constraints from $T$ are satisfied by $f_{j,\boldsymbol{u}}$ and $g_{i,\boldsymbol{v}}$'s.

**Soundness**. We prove the contrapositive. Let $\chi:G^k\to\mathbb{T}$ be a $j$-relevant character. Suppose there are folded assignments $f_{i,v}:G^{d_iR}\to\triangle_q$ for $M_j$ causing the bias to exceed $2\delta$. Then

$$\|M\|_\chi=\left|\mathop{\mathbb{E}}_{\boldsymbol{e}}\mathop{\mathbb{E}}_{\boldsymbol{z}}\chi(f_{\boldsymbol{e}}(\boldsymbol{z}))\right|\leqslant\mathop{\mathbb{E}}_{\boldsymbol{e}}\left|\mathop{\mathbb{E}}_{\boldsymbol{z}}\chi(f_{\boldsymbol{e}}(\boldsymbol{z}))\right|,$$

where $f_{\boldsymbol{e}}=(f_{1,\boldsymbol{w}_1},\ldots,f_{k,\boldsymbol{w}_k})$ with $\boldsymbol{w}_i=\boldsymbol{v}$ for $i\neq j$ and $\boldsymbol{w}_j=\boldsymbol{u}$. The RHS is at most

$$\mathop{\mathbb{E}}_{\boldsymbol{e}}\mathrm{Bias}_{T,\chi}(f_{\boldsymbol{e}}).$$

Therefore at least $\delta$ fraction of the edges $\boldsymbol{e}$ satisfy $\mathrm{Bias}_{T,\chi}(f_{\boldsymbol{e}})>\delta$. We call such edges good.

[8] For simplicity, we use active folding in the sense of [OWZ11].

For any good edge $e$, some $t_e \in [R]$ satisfies

$$\mathrm{Inf}_{t_e}^{(1-\eta)}[f_{j,u}], \max_{i \neq j} \left\{ \mathrm{Inf}_{\pi_e^{-1}(t_e)}^{(1-\eta)}[f_{i,v}] \right\} \geqslant \tau, \tag{3}$$

by Theorem 6.9.

We use the following randomized decoding procedure to generate an assignment $\boldsymbol{A}$ for the LC instance. For every $u \in U$, choose $S \subseteq [R]$ with probability $\|f_{j,u}^S\|_2^2$. (These numbers sum to at most 1 by the discussion following Fact 6.5. For the remaining probability, pick $S$ arbitrarily.) Then pick $\boldsymbol{A}(u)$ as a uniformly random element in $S$ (or assign arbitrarily if $S = \emptyset$). To get a label $\boldsymbol{A}(v)$, we first pick a random position $\boldsymbol{i} \in [k]$ different from $j$, then go on as before using $\|g_{\boldsymbol{i},v}^S\|_2^2$ as the probability distribution.

Then for any $B \subseteq [R]$ and any $u \in U$,

$$\begin{aligned}
\mathbb{P}[\boldsymbol{A}(u) \in B] &\geqslant \sum_{S : S \cap B \neq \emptyset} \|f_{j,u}^S\|_2^2 \cdot |S \cap B|/|S| \\
&\geqslant \sum_{S : S \cap B \neq \emptyset} \|f_{j,u}^S\|_2^2 \cdot \eta(1-\eta)^{|S|/|S \cap B|} \\
&\qquad \text{(since } \alpha \geqslant \eta(1-\eta)^{1/\alpha} \text{ for } \alpha > 0 \text{ and } 0 \leqslant \eta \leqslant 1) \\
&\geqslant \eta \cdot \mathrm{Inf}_B^{(1-\eta)}[f_{j,u}].
\end{aligned}$$

And similarly

$$\mathbb{P}[\boldsymbol{A}(v) \in B] \geqslant \eta \cdot \mathop{\mathbb{E}}_{\boldsymbol{i} \neq j} \mathrm{Inf}_B^{(1-\eta)}[f_{\boldsymbol{i},v}].$$

For a good edge, let $f_{i_e,v}$ be a function maximizing the influence on the LHS of (3).

$$\begin{aligned}
\mathbb{P}[\boldsymbol{A}(u) = \pi_e(\boldsymbol{A}(v))] &\geqslant \mathbb{P}[\boldsymbol{A}(u) = t_e \text{ and } \boldsymbol{A}(v) \in \pi_e^{-1}(t_e)] \\
&= \mathbb{P}[\boldsymbol{A}(u) = t_e] \cdot \mathbb{P}[\boldsymbol{A}(v) \in \pi_e^{-1}(t_e)] \\
&\geqslant \frac{\eta^2}{k-1} \cdot \mathrm{Inf}_{t_e}^{(1-\eta)}[f_{j,u}] \cdot \mathrm{Inf}_{\pi_e^{-1}(t_e)}^{(1-\eta)}[f_{i_e,v}] \geqslant \frac{\eta^2 \tau^2}{k-1}.
\end{aligned}$$

Therefore the expected fraction of constraints in $L$ satisfied by $\boldsymbol{A}$ exceeds $\delta\eta^2\tau^2/(k-1) \geqslant \sigma$. □

## Appendix D. Independent-Set

We prove Theorem 1.4 in this section. In the Independent-Set problem, a graph $H$ is given, and the goal is to find the largest independent set in $H$. The application of low free-bit PCP for Independent-Set is well known [ST09], but the actual hardness ratio is not explicitly computed before, so we include a proof for completeness.

Our proof closely follows [Tre01, Section 6]. We will construct an FGLSS graph $H$ (Definition 7.2) for our PCP, and reduce degree by replacing bipartite complete subgraphs in $H$ with "bipartite $\delta$-dispersers" (close relatives of bipartite expanders). The degree bound $O(\delta^{-1}\log(\delta^{-1}))$ for dispersers determines the hardness ratio. Unlike [Tre01], we do not use efficient deterministic constructions of dispersers, since none of the known constructions matches the degree bound offered by probabilistic ones. Luckily, bipartite complete subgraphs in $H$ have size bounded by a function of $1/\varepsilon$ and $1/\eta$, so we can find good dispersers by exhaustive search.

*Proof of Theorem 1.4.* By Corollary 1.2, there is a PCP $\Pi$ with completeness $c = 1 - \eta$, soundness $s = 2k/2^k + \varepsilon$, and free bit complexity at most $\log_2(2k)$. Construct the FGLSS graph $H$ for $\Pi$.

Following [DS05, Proposition 8.1], we now turn $H$ into an unweighted graph $H'$ (equivalently, vertices in $H'$ have equal weight), by duplicating vertices. Suppose $H$ is a weighted independent set instance of size $m$ with minimum weight $\lambda$ and maximum weight $\kappa$, and $0 < \sigma \leqslant \lambda$ be a granularity

parameter. Construct an unweighted instance $H'$ of size $O(m\kappa^2/\sigma^2)$ as follows: Replicate each vertex $u$ in $H$ of weight $w(u)$ by $\lfloor w(u)/\sigma \rfloor$ copies in $H'$; if $u$ and $v$ are connected in $H$, connect all copies of $u$ to all copies of $v$ in $H'$. Then weights are roughly preserved: any vertex $u$ of weight $w(u)$ in $H$ will have copies of total weight $w(u)(1 \pm O(\lambda/\sigma))$ in $H'$. Therefore, it is not hard to see that objective value is roughly preserved, $\mathrm{val}(H') = \mathrm{val}(H)(1 \pm O(\lambda/\sigma))$. Further, any vertex $u$ in $H$ has at most $\kappa/\sigma$ copies in $H'$.

As observed in [Tre01], the graph $H$ is a union of bipartite complete subgraphs. More precisely, for every index $i$ in the proof for $\Pi$, there is a bipartite complete subgraph between the sets $Z_i$ and $O_i$ of configurations, where configurations in $Z_i$ query index $i$ and expect an answer of zero, and configurations in $O_i$ query index $i$ and expect an answer of one. Further, the set of edges in $H$ is the union of all such bipartite complete subgraphs over index $i$. This bipartite complete subgraph structure is preserved by the vertex duplication process.

Also, the sets $Z_i$ and $O_i$ in $H$ have the same total weight, and in fact there is a weight-preserving bijection between $Z_i$ and $O_i$. This bijection is inherited from the corresponding bijection of the subgroup $C$, thanks to its balanced property. As a result, in the instance $H'$ after duplication of vertices, the vertex sets $Z_i$ and $O_i$ have the same size $\ell_i$.

We now replace the bipartite complete subgraph between $O_i$ and $Z_i$ with a bipartite disperser on $([\ell_i], [\ell_i])$, for all index $i$. The graph after replacement is $H''$.

**Proposition D.1.** *For every $\delta > 0$ and any $\ell \geqslant 1$, there is a bipartite graph on $(([\ell], [\ell]), E)$ of degree at most $d = O(\delta^{-1} \log(\delta^{-1}))$ such that for any $A, B \subseteq [\ell]$, $|A| \geqslant \lfloor \delta\ell \rfloor$ and $|B| \geqslant \lfloor \delta\ell \rfloor$, some edge in $E$ goes between $A$ and $B$, so $(A \times B) \cap E \neq \emptyset$.*

A random bipartite graph is well-known to be a $\delta$-disperser (for completeness, we include a proof below). We can therefore find (and verify) a disperser deterministically by exhaustive search in time $\exp(\mathrm{poly}(\ell_i))$.

To bound $\ell_i$, we first bound the maximum size $W$ of $Z_i$ in $H$ (measured by the number of vertices, disregarding weights). Then $W$ times the maximum number of copies of a vertex will upperbound $\ell_i$. It is not hard to see that $W = O_{\varepsilon,k}(1)$, where $O_{\varepsilon,k}(1)$ denotes a quantity bounded by a function of $\varepsilon$ and $k$. Indeed, $W$ is at most $2^f \Delta(M)$, where $\Delta(M)$ is the maximum number of constraints incident on a variable in the instance $M$ of Theorem 1.1 (disregarding weight on constraints). To bound $\Delta(M)$, observe that $\Delta(L) = O_{\varepsilon,k}(1)$ for the Label-Cover instance $L$ of Theorem C.1. Also, $\Delta(M_j) = O_{\varepsilon,k}(1)$, where $M_j$ is the instance from Theorem 5.4. Further, direct sum preserves boundedness of $\Delta$, since $\Delta(M \oplus M') = \Delta(M)\Delta(M')$. This shows that $W = O_{\varepsilon,k}(1)$.

We bound the number of copies of a vertex in the replication step by $\kappa/\sigma$. To bound $\kappa/\sigma$, we first bound the ratio $\rho(M) = \kappa(M)/\sigma(M)$ of the maximum weight constraint to minimum weight constraint in a CSP instance $M$. Then $\rho(L) = 1$ for the Label-Cover instance $L$ in Theorem C.1, because $L$ is a bi-regular bipartite graph. After composing with the dictator test, $\rho(M_j)$ is at most $O_{\varepsilon,\eta,k}(1)$. Finally, $\rho(M \oplus M') = \rho(M)\rho(M')$. Hence the ratio $\kappa/\lambda$ for the FGLSS graph $H$ is $O_{\varepsilon,\eta,k}(1)$. If we pick $\sigma = \varepsilon\lambda$, then $\ell_i = O_{\varepsilon,\eta,k}(1)$.

The disperser replacement step increases the objective value by at most $k\delta$ [Tre01]. We will therefore choose $\delta = s2^{-f}/k$, and the degree bound for $H''$ becomes $D = O(k/\delta \cdot \log(1/\delta)) = O(k^3 2^k)$. The hardness ratio is $O(c/s) = O(k/2^k) = O(\log D)^4/D$. □

*Proof of Proposition D.1.* We may assume $\ell \geqslant \delta^{-1} \log(\delta^{-1})$ (otherwise, just take the bipartite complete graph). Assume for now that $\delta\ell$ is an integer.

Denote by $U, V$ the two vertex subsets of size $\ell$. We pick a random degree-$d$ bipartite (multi)-graph on $(U, V)$, generated as the union of $d$ independent random perfect matchings.

Consider $A \subseteq U$ of size $\delta\ell$ and $B \subseteq V$ of size $\delta\ell$. The probability that in a perfect matching, all edges from $A$ miss $B$ is $\binom{(1-\delta)\ell}{\delta\ell} / \binom{\ell}{\delta\ell} \leqslant (1-\delta)^{\delta\ell}$. Hence $A$ shares no edges with $B$ with probability at

most $(1-\delta)^{d\delta\ell}$. Taking union bound over all choices of $A$ and $B$, the random graph is a $\delta$-disperser except with probability at most

$$\binom{\ell}{\delta\ell}\binom{\ell}{\delta\ell}(1-\delta)^{d\delta\ell} \leqslant \left(\frac{e^2}{\delta^2}(1-\delta)^d\right)^{\delta\ell},$$

where we have used $\binom{n}{r} \leqslant (en/r)^r$. The quantity in bracket on the RHS is less than 1 when $d = O(\delta^{-1}\log(\delta^{-1}))$.

When $\delta\ell$ is not an integer, it is easy to get the same conclusion using $\ell \geqslant \delta^{-1}\log(\delta^{-1})$ and appropriate approximations. □

## Appendix E. Samorodnitsky–Trevisan hypergraph predicates

Let $k = 2^r - 1$. The Samorodnitsky–Trevisan hypergraph predicate of arity $k$ is the dual Hamming code $C$ of block length $k$ and dimension $r$ over $\mathbb{F}_2$ [ST09]. The main result of this section is an alternative proof of Theorem A.1 for these predicates, with stronger bounds. This stronger version is not needed in this paper.

**Theorem E.1.** *Let $T$ be the $j$-lonely, $\eta$-noise correlated test for the hypergraph predicate of arity $k$. Let $\chi : \mathbb{Z}_2^k \to \mathbb{T}$ be a non-trivial character. Suppose functions $f_i : \mathbb{Z}_2^{d_i R} \to \triangle_2$ satisfy*

$$\max_{i\neq j}\{\mathrm{Inf}^{(1-\eta)}_{B(t)}[f_i]\} \leqslant \tau \quad \forall t \in [R].$$

*Assume further $f_i$'s are folded. Then* $\mathrm{Bias}_{T,\chi}(f) \leqslant \mathrm{poly}(k/\eta)\cdot\sqrt{\tau}$.

Note that the strategy $f_j$ of the lonely player needs not have small influence. On the other hand, all strategies $f_i$'s are assumed to be folded. Our analysis builds on [Hås09, Section 3], incorporating ideas from the Max-3-Lin analysis of [Hås01, Section 5] to handle the lonely player.

If $\chi = \chi_j$ depends only on the $j$-th coordinate, then $\mathrm{Bias}_{T,\chi}(f) = |\mathbb{E}[\chi_j(f_j(\boldsymbol{z}^{(j)}))]| = 0$, because folding forces $f_j(\boldsymbol{z}^{(j)})$ to be uniformly random. Therefore it suffices to consider the case $\chi$ is $j'$-relevant for some $j' \neq j$.

For convenience, we identify the $k = 2^r - 1$ players as non-empty subsets of $[r]$. Due to symmetry of players in $C$, we may assume player $\{r\} \subseteq [r]$ is lonely.

Using notations from Definition 6.2, reinterpret the correlated distribution $\mu$ over $\prod_{\emptyset\neq U\subseteq[r]}\mathbb{Z}_2^{d_U}$ as follows. Choose $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_r \in \mathbb{Z}_2^d$ independently, conditioned on $\boldsymbol{x}_r$ having the same element on all $d$ coordinates. Then define $\boldsymbol{z}^U = \sum_{i\in U}\boldsymbol{x}_i$ for $U \neq \{r\}$ and $\boldsymbol{z}^{\{r\}} = \pi(\boldsymbol{x}_r)$, where $\pi : \mathbb{Z}_2^d \to \mathbb{Z}_2$ is projection to (say) the first coordinate. This completes the description of a block $\boldsymbol{z} = (\boldsymbol{z}^U)_{\emptyset\neq U\subseteq[r]} \in \prod_{\emptyset\neq U\subseteq[r]} G^{d_U}$, where $d_{\{r\}} = 1$ and and $d_U = d$ for $U \neq \{r\}$.

We need the following propositions.

**Proposition E.2** ([Hås09, Theorem 3.4])**.** *Let $(f_U)_{U\subseteq[r]}$ be a tuple of functions from $\mathbb{Z}_2^m$ to $[-1,1]$ satisfying $\mathrm{Inf}_t[f_U] \leqslant \tau$ for all $t \in [m]$ and all non-empty $U \subseteq [r]$. Suppose $\min_{U\neq\emptyset}|\mathbb{E}[f_U]| \leqslant \delta$. Then*

$$\left|\mathop{\mathbb{E}}_{\boldsymbol{y}_1,\ldots,\boldsymbol{y}_r\in\mathbb{Z}_2^m}\left[\prod_{U\subseteq[r]} f_U\left(\sum_{i\in U}\boldsymbol{y}_i\right)\right]\right| \leqslant \delta + (2^r-2)\sqrt{\tau}.$$

**Proposition E.3** ([ST09, Lemma 4])**.** *Let $f, g : \mathbb{Z}_2^m \to [-1,1]$ be functions, and define $h(x) = f(x)g(x)$. Then for every $t \in [m]$,*

$$\mathrm{Inf}_t[h] \leqslant 2(\mathrm{Inf}_t[f] + \mathrm{Inf}_t[g]).$$

We also need the following simple fact that follows easily by Fourier analysis (proof omitted).

**Fact E.4.** *For any $f : \mathbb{Z}_2^m \to \mathbb{R}$, any $y \in \mathbb{Z}_2^m$, let $g^y(x) = f(x+y)$. Then for any $B \subseteq [m]$,*

$$\mathrm{Inf}_B[g^y] = \mathrm{Inf}_B[f].$$

Given a function $f : \Sigma^m \to \mathbb{R}^q$ and a linear map $\mathrm{X} : \mathbb{R}^q \to \mathbb{R}$, consider the composed function $\mathrm{X}(f) : \Sigma^m \to \mathbb{R}$. It is easy to see that this linear map "commutes" with Hoeffding decomposition.

**Proposition E.5.** *The Hoeffding decomposition of $\mathrm{X}(f) = \sum_{S\subseteq[m]} (\mathrm{X}(f))^S$ has components*

$$(\mathrm{X}(f))^S = \mathrm{X}(f^S).$$

*Proof.* Expanding $f$ in its Hoeffding decomposition, and using linearity of X, we can write

$$\mathrm{X}(f) = \mathrm{X}\left(\sum_{S\subseteq[m]} f^S\right) = \sum_{S\subseteq[m]} \mathrm{X}(f^S). \tag{4}$$

Clearly, $\mathrm{X}(f^S)$ depends only on $x_S$. Also, given any $S \not\subseteq T$ and $x_T \in \Sigma^T$,

$$\mathop{\mathbb{E}}_{\boldsymbol{x}\in G^m}[\mathrm{X}(f^S)(\boldsymbol{x}) \mid \boldsymbol{x}_T = x_T] = 0,$$

by linearity of X. Therefore (4) is also a Hoeffding decomposition of $\mathrm{X}(f)$, and the result follows because the decomposition is unique. □

Definition 6.6 and Proposition E.5 imply that

$$\mathrm{Inf}_S[\mathrm{X}(f)] \leqslant \|\mathrm{X}\|_{\mathrm{op}}^2 \cdot \mathrm{Inf}_S[f] \tag{5}$$

where

$$\|\mathrm{X}\|_{\mathrm{op}} \triangleq \sup_{y\neq 0} \frac{|\mathrm{X}(y)|}{\|y\|_{\ell^2}}.$$

*Proof of Theorem E.1.* Recall that $\chi$ can be decomposed as a product of characters (see (1) in Section 3). For non-empty $U \subseteq [r]$, let $\chi_U : \triangle_2 \to [-1, 1]$ be the multilinearized version of the $U$-component of $\chi$.[9] Then $\mathrm{Bias}_{T,\chi}(f)$ equals the magnitude of the quantity

$$\mathbb{E} \prod_{\emptyset\neq U\subseteq[r]} \chi_U(F_U(\boldsymbol{z}^U)), \tag{6}$$

where $F_U = \mathrm{T}_{1-\eta} f_U$. For $x \in \mathbb{Z}_2^m$ and non-empty $U \subseteq [r-1]$, let

$$g_U^x(y) = \chi_U(F_U(y))\chi_{U'}(F_{U'}(x+y))$$

where $U' = U \cup \{r\}$. Also let

$$g_\emptyset^x(y) = \chi_{\{r\}}(F(\pi(x)))$$

where $\pi : \mathbb{Z}_2^{dR} \to \mathbb{Z}_2^R$ projects every block to its first coordinate. Then (6) can be rewritten as

$$\mathop{\mathbb{E}}_{\boldsymbol{x}_r} \mathop{\mathbb{E}}_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{r-1}} \prod_{U\subseteq[r-1]} g_U^{\boldsymbol{x}_r}\left(\sum_{i\in U} \boldsymbol{x}_i\right).$$

For every fixed $\boldsymbol{x}_r$, the magnitude of the inner expectation (over $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{r-1}$) can be bounded by Proposition E.2. To this end, we bound the influence and balance of the $h_U^x$'s.

By Proposition E.3 and Fact E.4, for any $x \in \mathbb{Z}_2^{dR}$, $t \in [R]$ and non-empty $U \subseteq [r]$,

$$\mathrm{Inf}_t[g_U^x] \leqslant 2(\mathrm{Inf}_t[\chi_U(F_U)] + \mathrm{Inf}_t[\chi_{U'}(F_{U'})]).$$

By (5) and Appendix B, $\mathrm{Inf}_t[\chi_U(F_U)] \leqslant 2\,\mathrm{Inf}_t[F_U]$. Therefore $\mathrm{Inf}_t[g_U^x] \leqslant 8\tau$.

[9] See (2) in Appendix B for the definition of the multilinearized version of a character.

As argued earlier, we may assume the non-trivial character $\chi$ is $T_0$-relevant for some non-empty $T_0 \subseteq [r]$ different from $\{r\}$. Let $T = T_0 \cap [r-1]$, so that $T \neq \emptyset$. We will show the following bound

$$\mathop{\mathbb{E}}_{\boldsymbol{x}_r} \left| \mathop{\mathbb{E}}_{\boldsymbol{z}^T} [g_T^{x_r}(\boldsymbol{z}^T)] \right| \leqslant 2\sqrt{\tau c_\eta}, \tag{7}$$

where $c_\eta$ is from Fact A.2. Assuming (7), we bound the bias by

$$\begin{aligned}
\mathrm{Bias}_{T,\chi}(f) &\leqslant \mathop{\mathbb{E}}_{\boldsymbol{x}_r} \left| \mathop{\mathbb{E}}_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{r-1}} \prod_{U \subseteq [r-1]} g_U^{\boldsymbol{x}_r} \left( \sum_{i \in U} \boldsymbol{x}_i \right) \right| \\
&\leqslant \mathop{\mathbb{E}}_{\boldsymbol{x}_r} \left[ \left| \mathop{\mathbb{E}}_{\boldsymbol{z}^T} [g_T^{\boldsymbol{x}_r}(\boldsymbol{z}^T)] \right| + k\sqrt{2\tau} \right] \quad \text{(Proposition E.2)} \\
&\overset{(7)}{\leqslant} 2(\sqrt{c_\eta} + k)\sqrt{\tau},
\end{aligned}$$

proving our theorem.

It remains to prove (7). Let

$$h(x) = \mathop{\mathbb{E}}_{\boldsymbol{z} \in \mathbb{Z}_2^{dR}} g_T^x(\boldsymbol{z}) = \mathop{\mathbb{E}}_{\boldsymbol{z}} \chi_T(F_T(\boldsymbol{z}))\chi_{T'}(F_{T'}(x + \boldsymbol{z})),$$

where $T' = T \cup \{r\}$. Expand $h$ in its Fourier series[10]

$$h(x) = \sum_{\alpha \in \mathbb{Z}_2^{dR}} \hat{h}(\alpha)\chi_\alpha(x),$$

where $\chi_\alpha(x) = (-1)^{\sum_{s \in [dR]} x_s \alpha_s}$ are Fourier characters and are unrelated to the given non-trivial character $\chi$. Then the square of the LHS of (7) equals

$$\begin{aligned}
\left( \mathop{\mathbb{E}}_{\boldsymbol{x}_r} |h(\boldsymbol{x}_r)| \right)^2 &\leqslant \mathop{\mathbb{E}}_{\boldsymbol{x}_r} [h(\boldsymbol{x}_r)^2] \qquad \text{(Cauchy–Schwarz)} \\
&= \sum_{\alpha,\beta} \hat{h}(\alpha)\hat{h}(\beta) \mathop{\mathbb{E}}_{\boldsymbol{x}_r} [\chi_\alpha(\boldsymbol{x}_r)\chi_\beta(\boldsymbol{x}_r)] \\
&= \sum_{\alpha,\beta : \pi_2(\alpha) = \pi_2(\beta)} \hat{h}(\alpha)\hat{h}(\beta), \qquad (8)
\end{aligned}$$

where $\pi_2 : \mathbb{Z}_2^{dR} \to \mathbb{Z}_2^R$ denotes the map $\pi_2(x)_t = \sum_{s \in B(t)} x_s$ for all $t \in [R]$. The last inequality uses the fact that $\boldsymbol{x}_r$ is constant on each block, so

$$\mathbb{E}[\chi_\alpha(\boldsymbol{x}_r)\chi_\beta(\boldsymbol{x}_r)] = \mathop{\mathbb{E}}_{\boldsymbol{x} \in \mathbb{Z}_2^R} [\chi_{\pi_2(\alpha)}(\boldsymbol{x})\chi_{\pi_2(\beta)}(\boldsymbol{x})],$$

which is 1 if $\pi_2(\alpha) = \pi_2(\beta)$ and is 0 otherwise.

Note that $h$ is the convolution of $\chi_T(F_T)$ and $\chi_{T'}(F_{T'})$. Since convolution becomes multiplication in frequency domain, if $\chi_T(F_T)$ and $\chi_{T'}(F_{T'})$ have Fourier series

$$\chi_T(F_T(x)) = \sum_{\alpha \in \mathbb{Z}_2^{dR}} \hat{f}(\alpha)\chi_\alpha(x) \quad \text{and} \quad \chi_{T'}(F_{T'}(x)) = \sum_{\alpha \in \mathbb{Z}_2^{dR}} \hat{f}'(\alpha)\chi_\alpha(x),$$

then

$$\hat{h}(\alpha) = \hat{f}(\alpha)\hat{f}'(\alpha) \tag{9}$$

[10] We assume the reader is familiar with Fourier analysis. See e.g. [Hås01, Section 2.4].

for all $\alpha \in \mathbb{Z}_2^{dR}$. Suppose $T_0 = T$ (the other case $T_0 = T \cup \{r\}$ is analogous). Then $f_T$, and hence $\chi_T(F_T)$,[11] is folded, thus $\hat{f}(\alpha) = 0$ whenever $\sum_s \alpha_s = 0$. Since $\pi_2(\alpha) = 0$ implies $\sum_s \alpha_s = 0$,

$$\hat{h}(\alpha) = \hat{f}(\alpha)\hat{f}'(\alpha) = 0$$

whenever $\pi_2(\alpha) = 0$.

As a result, the sum in (8) runs over $\alpha, \beta$ such that $\gamma \triangleq \pi_2(\alpha) = \pi_2(\beta)$ is *non-zero*. Thus

$$(8) = \sum_{\gamma \neq 0} \left( \sum_{\alpha : \pi_2(\alpha) = \gamma} \hat{h}(\alpha) \right)^2 \leqslant \sum_{t \in [R]} \sum_{\gamma_t \neq 0} \left( \sum_{\alpha : \pi_2(\alpha) = \gamma} \hat{h}(\alpha) \right)^2 .$$

For every $t$, the inner sum equals

$$\begin{aligned}
\sum_{\gamma_t \neq 0} \left( \sum_{\alpha : \pi_2(\alpha) = \gamma} \hat{h}(\alpha) \right)^2 &\overset{(9)}{=} \sum_{\gamma_t \neq 0} \left( \sum_{\alpha : \pi_2(\alpha) = \gamma} \hat{f}(\alpha)\hat{f}'(\alpha) \right)^2 \\
&\leqslant \sum_{\gamma_t \neq 0} \left( \sum_{\alpha : \pi_2(\alpha) = \gamma} \hat{f}(\alpha)^2 \right) \left( \sum_{\alpha : \pi_2(\alpha) = \gamma} \hat{f}'(\alpha)^2 \right) \quad \text{(Cauchy–Schwarz)} \\
&\leqslant \left( \sum_{\alpha : \pi_2(\alpha)_t \neq 0} \hat{f}(\alpha)^2 \right) \left( \sum_{\alpha : \pi_2(\alpha)_t \neq 0} \hat{f}'(\alpha)^2 \right) \\
&\leqslant \mathrm{Inf}_{B(t)}[\chi_T(F_T)] \cdot \mathrm{Inf}_{B(t)}[\chi_{T'}(F_{T'})]. \qquad (10)
\end{aligned}$$

The last inequality holds because $\pi_2(\alpha)_t \neq 0$ implies $\alpha_s = 1$ for some $s \in B(t)$. Again we have $\mathrm{Inf}_{B(t)}[\chi_T(F_T)] \leqslant 2\,\mathrm{Inf}_{B(t)}[F_T]$. Plugging (10) into (8), we get

$$\left( \mathop{\mathbb{E}}_{\boldsymbol{x}_r} |h(\boldsymbol{x}_r)| \right)^2 \leqslant 4 \sum_{t \in [R]} \mathrm{Inf}_{B(t)}[F_T]\,\mathrm{Inf}_{B(t)}[F_{T'}] \leqslant 4\tau \sum_{t \in [R]} \mathrm{Inf}_{B(t)}[F_{T'}] \leqslant 4\tau c_\eta,$$

where the second inequality is our influence assumption on $f_T$ and the last inequality is Fact A.2. This gives (7). □

This stronger bound propagates through Theorems A.1, 6.9 and C.2, leading to the bound

$$\delta = \mathrm{poly}(k/\eta) \cdot \sigma^{\Omega(1)}$$

for Theorem C.2, when $C$ is a hypergraph predicate.

## Appendix F. Håstad predicates

We describe a predicate due to Johan Håstad and announced in [MM12]. This predicate is used in Corollary 1.7.

Let $k \leqslant 2^t$, $q = 2^s$, and suppose $t \geqslant s$. A Håstad predicate is over $G = \mathbb{Z}_2^s$. We pick a random tuple $\boldsymbol{c} \in G^k$ as follows. Pick random $\boldsymbol{a} \in \mathbb{F}_{2^t}$ and $\boldsymbol{b} \in \mathbb{Z}_2^s$, and set

$$\boldsymbol{c}_i = \pi(\boldsymbol{a} \cdot \bar{i}) + \boldsymbol{b},$$

where $\bar{i}$ denotes the $i$-th element from $\mathbb{F}_{2^t}$, and $\pi : \mathbb{F}_{2^t} \to \mathbb{Z}_2^s$ is any surjective group homomorphism (e.g. $\pi$ takes the first $s$ bits in some vector space representation of $\mathbb{F}_{2^t}$ over $\mathbb{F}_2$).

[11] This is the only step that uses the $T_0$-relevant property of $\chi$.

Let $C$ be the collection of random tuples $\boldsymbol{c}$ generated as above. Then $C$ has size at most $qk$. Further, $C$ is balanced pairwise independent, because for every $i \neq j \in [k]$, the difference

$$\boldsymbol{c}_i - \boldsymbol{c}_j = \pi(\boldsymbol{a} \cdot \bar{i}) - \pi(\boldsymbol{a} \cdot \bar{j}) = \pi(\boldsymbol{a} \cdot (\bar{i} - \bar{j}))$$

is uniformly random over $\mathbb{Z}_2^s$, for any fixed $\boldsymbol{b}$.

Håstad predicates require $q$ to be a prime power. To obtain Corollary 1.7 where $q$ is arbitrary, pick the smallest power of two $q' \geqslant q$, and apply Makarychev's randomized reduction [AM09, Proposition B.1] from domain size $q'$ to domain size $q$.

## Appendix G. Lasserre integrality gaps

In this section, we observe that Schoenebeck's Lasserre gap construction for $k$-XOR [Sch08] also works for the predicates in Theorem 1.1, drawing a pleasing parallel between Lasserre gap construction and NP-hardness results.[12] Previously, Tulsiani [Tul09] extended Schoenebeck's construction to any predicate that is a linear code of dual distance at least 3 over a prime field. Later Schoenebeck simplified his own proof of $k$-XOR using Fourier analysis [Sch08]. Not surprisingly, his new proof can be further generalized to arbitrary abelian group using Pontryagin duality, as shown below. For intuition about the construction, see [Sch08].

Given an abelian group $G$, its dual group $\hat{G}$ is the abelian group of characters on $G$, under pointwise multiplication. The inverse of $\chi \in \hat{G}$ is therefore $\overline{\chi}$. Pontryagin duality says that $G$ is naturally isomorphic to the dual of $\hat{G}$ (i.e. double dual of $G$), via the "evaluation map"

$$g \in G \quad \mapsto \quad \{\chi \in \hat{G} \mapsto \chi(g)\}.$$

Given a subgroup $H$ of $G$, denote by $H^\perp = \{\chi \in \hat{G} \mid \chi(h) = 1 \ \forall h \in H\}$ the annihilator of $H$.[13] The following fact is well known.

**Proposition G.1** ([HR94, Theorems 23.25, 24.10]). *Let $\Lambda$ be a subgroup of a finite abelian group $\Gamma$. Then (a) $\widehat{\Gamma/\Lambda} \cong \Lambda^\perp$ and (b) $(\Lambda^\perp)^\perp = \Lambda$.*

A (linear) equation is a pair $(\chi, z) \in \widehat{G^V} \times \mathbb{T}$, encoding the constraint $\chi(f) = z$ for an assignment $f : V \to G$. Since $\widehat{G^V}$ is isomorphic to $\hat{G}^V$, we write $\hat{G}^V$ in place of $\widehat{G^V}$ for better typography. The support of $\chi \in \hat{G}^V$ is $\mathrm{supp}(\chi) \triangleq \{v \in V \mid \chi \text{ is } v\text{-relevant}\}$, and the degree of $\chi$ is the size of its support. Denote by $\Omega_t$ the collection of $\chi$ of degree at most $t$.

**Definition G.2.** Given a collection $R$ of equations, its width-$t$ resolution $\Pi_t(R) \subseteq \hat{G}^V \times \mathbb{T}$ contains all equations in $R$ and those derived via the resolution step

$$(\chi, z), (\psi, y) \in \Pi_t(R) \text{ and } \chi\overline{\psi} \in \Omega_t \quad \Longrightarrow \quad (\chi\overline{\psi}, z\overline{y}) \in \Pi_t(R).$$

The resolution has no contradiction if $(\mathbf{1}, z) \in \Pi_t(R)$ implies $z = 1$.

In this section, an Additive-CSP($C$) instance $M = (V, \boldsymbol{Q})$ will not be $k$-partite, so all variables $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ of the $k$-tuple $\boldsymbol{v}$ in a question $\boldsymbol{Q} = (\boldsymbol{v}, \boldsymbol{b})$ come from the same variable set $V$. Let $R_M$ be the set of equations from constraints in $M$, defined as

$$R_M \triangleq \{(\chi, \chi(\boldsymbol{b})) \mid (\boldsymbol{v}, \boldsymbol{b}) \in M, \chi \in C^\perp \subseteq \hat{G}^{\boldsymbol{v}}\}.$$

We say that $M$ has resolution width at least $t$ if $\Pi_t(R_M)$ has no contradiction.

Our definition of Lasserre solution is a rephrasing of the one in [Tul09].

[12] Even without the result in the section, Theorem 1.1 implies a Lasserre gap via reduction, but the number of rounds of the Lasserre solution will not be linear, due to the blow-up in size from direct sum.

[13] Annihilator is only defined with respect to an ambient group $G$, which will always be clear from the context.

**Definition G.3.** A $t$-round Lasserre solution $U$ for a CSP instance $M = (V, \boldsymbol{Q})$ over domain $\Sigma$ is a collection $\{U_f \mid f \in \Sigma^S, S \subseteq V \text{ s.t. } |S| \leqslant t\}$ of vectors, one for each partial assignment $f : S \to \Sigma$ on a subset $S$ of size at most $t$.

The Lasserre solution induces a collection of distributions $\{\mu_W \in \triangle_{\Sigma^W} \mid W \subseteq V \text{ s.t. } |W| \leqslant 2t\}$ over partial assignments, subject to the following condition: For any two partial assignments $f \in \Sigma^S$ and $g \in \Sigma^T$ with $|S|, |T| \leqslant t$, we have

$$\langle U_f, U_g \rangle = \mathop{\mathbb{P}}_{\boldsymbol{h} \sim \mu_{S \cup T}} [\boldsymbol{h} \restriction_S = f \text{ and } \boldsymbol{h} \restriction_T = g]. \tag{11}$$

The value of the Lasserre solution is $\mathrm{val}(M, U) = \mathbb{E}_{\boldsymbol{Q}} \mathbb{P}[\boldsymbol{Q} \text{ is satisfied under } \mu_{\langle \boldsymbol{Q} \rangle}]$, where $\langle \boldsymbol{Q} \rangle \subseteq V$ denotes the set of variables that $\boldsymbol{Q}$ depends on.

A key step will be the following generalization of [Tul09, Theorem B.1].

**Theorem G.4.** *Let $G$ be an abelian group, and $C$ a subgroup of $G^k$. If an* Additive-CSP$(C)$ *instance $M$ has resolution width at least $2t$, then there is a $t$-round Lasserre solution to $M$ of value 1.*

Given the resolution proof $\Pi = \Pi_{2t}(R_M)$, denote by $\Lambda = \{\chi \mid (\chi, z) \in \Pi\}$ the collection of $\chi$'s appearing in an equation. If $\Pi$ has no contradiction, then for every $\chi \in \Lambda$, there is a unique $z(\chi) \in \mathbb{T}$ such that $(\chi, z(\chi)) \in \Pi$. Otherwise the existence of distinct $(\chi, z), (\chi, y)$ in $\Pi$ implies $(\mathbf{1}, 1) \neq (\mathbf{1}, z\overline{y}) \in \Pi$, a contradiction (pun intended). By definition of the resolution step, if $\chi, \psi, \chi\psi \in \Lambda$, then

$$z(\chi\psi) = z(\chi)z(\psi), \tag{12}$$

so $z : \Lambda \to \mathbb{T}$ is a homomorphism wherever it is defined.

The key observation is that if $\chi \notin \Lambda$, then $\chi$ does not enforce any constraint on partial assignments. We make this precise in (13) below. For $W \subseteq V$, let $\Lambda_W = \{\chi \in \Lambda \mid \mathrm{supp}(\chi) \subseteq W\}$, which will be considered as a subgroup of $\hat{G}^W$. Let $H_W$ be the set of partial assignments on $W$ that satisfy all the constraints contained in $W$,

$$H_W = \{h \in G^W \mid \forall \chi \in \Lambda_W,\ \chi(h) = z(\chi)\}.$$

We now show that for every $W$ of size at most $2t$ and every $\chi \in \hat{G}^W \setminus \Lambda_W$,

$$\mathop{\mathbb{E}}_{\boldsymbol{h} \in H_W} \chi(\boldsymbol{h}) = 0. \tag{13}$$

Indeed, $H_W$ is a coset of $\Lambda_W^\perp$, so (13) follows from Proposition G.5 with $\Lambda := \Lambda_W, \Gamma := \hat{G}_W, H := H_W$.

**Proposition G.5.** *Let $\Lambda$ be a subgroup of an abelian group $\Gamma$, and $H \subseteq \hat{\Gamma}$ be a coset of $\Lambda^\perp$. Then for any $\chi \in \Gamma$,*

$$\chi \in \Lambda \quad \Longleftrightarrow \quad \mathop{\mathbb{E}}_{\boldsymbol{h} \in H} \boldsymbol{h}(\chi) \neq 0.$$

*Proof.* Let $H = h\Lambda^\perp$. We have

$$\mathop{\mathbb{E}}_{\boldsymbol{h} \in H} \boldsymbol{h}(\chi) = h(\chi) \cdot \mathop{\mathbb{E}}_{\boldsymbol{h} \in \Lambda^\perp} \boldsymbol{h}(\chi) = h(\chi) \cdot \mathop{\mathbb{E}}_{\boldsymbol{z} \in \chi(\Lambda^\perp)} \boldsymbol{z},$$

where second equality uses the fact that $\chi$ is a homomorphism from $\hat{\Gamma}$ to $\mathbb{T}$, by Pontryagin duality. Now the RHS is non-zero if and only if $\chi(\Lambda^\perp)$ contains only one element, that is $\chi(\Lambda^\perp)$ is the trivial subgroup $\{1\}$ of $\mathbb{T}$. The latter condition is equivalent to $\chi \in (\Lambda^\perp)^\perp$, and the result follows by Proposition G.1(b). □

Partition $\Omega_t$ into equivalence classes $[\chi]$'s so that $[\chi] = [\psi]$ if $\chi\overline{\psi} \in \Lambda$. It is easily checked that the latter condition is indeed an equivalence relation. Also fix an arbitrary representative $\chi'$ for

each equivalence class $[\chi]$. In the Lasserre vector construction, there will be an orthonormal set of vectors $e_{[\chi]}$'s, one for each equivalent class.

Our goal is Lasserre vectors $U_f$ for partial assignments $f : S \to G$, and to this end we first construct Lasserre vectors $U_A$ for any $t$-junta $A$, which is a function $A : G^V \to \mathbb{C}$ depending on at most $t$ variables. Formally, let $\mathrm{supp}(A)$ be the smallset subset $S \subseteq V$ on which there is $B : S \to G$ satisfying $A(h) = B(h \restriction_S)$ for all $h \in G^V$. Then $A$ is a $t$-junta if $\mathrm{supp}(A)$ has size at most $t$. Since any $t$-junta $A$ is a linear combination of characters of degree at most $t$, i.e.

$$A = \sum_{\chi \in \hat{G}^S} \hat{A}(\chi)\chi$$

where $S = \mathrm{supp}(A)$, it suffices to define the Lasserre vector

$$U_\chi = z(\chi\overline{\chi'})e_{[\chi]}$$

for $\chi \in \Omega_t$ and extend the definition to an arbitrary $t$-junta $A$ by linearity.

The following proposition highlights the main property.

**Proposition G.6.** *For any $t$-juntas $A, B : G^V \to \mathbb{C}$, let $W = \mathrm{supp}(A) \cup \mathrm{supp}(B)$. Then*

$$\langle U_A, U_B \rangle = \mathop{\mathbb{E}}_{\boldsymbol{h} \in H_W} [A(\boldsymbol{h})\overline{B}(\boldsymbol{h})].$$

*Proof.* By linearity, it suffices to show that for any $\chi, \psi \in \Omega_t$, if $W = \mathrm{supp}(\chi) \cup \mathrm{supp}(\psi)$ (which has size at most $2t$), then

$$\langle U_\chi, U_\psi \rangle = \mathop{\mathbb{E}}_{\boldsymbol{h} \in H_W} [\chi(\boldsymbol{h})\overline{\psi}(\boldsymbol{h})] = \mathop{\mathbb{E}}_{\boldsymbol{h} \in H_W} [\chi\overline{\psi}(\boldsymbol{h})].$$

When $[\chi] \neq [\psi]$, the LHS is zero because $e_{[\chi]}$ and $e_{[\psi]}$ are orthogonal, and the RHS is also zero by (13).

When $[\chi] = [\psi]$, the LHS is $z(\chi\overline{\chi'})\overline{z}(\psi\overline{\chi'}) = z(\chi\overline{\psi})$ by (12), and the RHS is also $z(\chi\overline{\psi})$ by definition of $H_W$ and the fact that $\chi\overline{\psi} \in \Lambda$. □

*Proof of Theorem G.4.* We will consider the indicator function $A : G^V \to \mathbb{R}$ for a partial assignment $f : S \to G$, defined as

$$A(h) = \mathbb{I}(h \restriction_S = f).$$

Then $A$ is a $t$-junta. We then define $U_f$ as $U_A$.

For any partial assignments $f \in G^S, g \in G^T$,

$$\langle U_f, U_g \rangle = \mathop{\mathbb{E}}_{\boldsymbol{h} \in H_W} [\mathbb{I}(\boldsymbol{h} \restriction_S = f)\mathbb{I}(\boldsymbol{h} \restriction_T = g)]$$

by Proposition G.6. Taking $\mu_W$ as the uniform distribution over $H_W$, the vectors $U_f$'s satisfy the Lasserre constraints (11).

The Lasserre solution has value 1, because every constraint $\boldsymbol{Q} \in M$ is satisfied by every $f \in H_{\langle \boldsymbol{Q} \rangle}$. Indeed, since $\boldsymbol{Q} = (\boldsymbol{v}, \boldsymbol{b})$ induces linear equations $\{(\chi, \chi(\boldsymbol{b})) \mid \chi \in C^\perp \subseteq \hat{G}^{\boldsymbol{v}}\}$ in $\Pi$, we have

$$f \in H_W \quad \Longrightarrow \quad \chi(f - \boldsymbol{b}) = 1 \ \forall \chi \in C^\perp \quad \Longleftrightarrow \quad f - \boldsymbol{b} \in (C^\perp)^\perp = C,$$

where the equivalence is Pontryagin duality and the last equality is Proposition G.1(b).

The vectors $U_f$ may have complex entries, but equivalent *real* vectors exist. Indeed, the Gram matrix $[\langle U_f, U_g \rangle]_{f,g}$ has only real entries and is positive semidefinite over $\mathbb{C}$, and hence over $\mathbb{R}$. □

As usual, a random Additive-CSP($C$) instance $M$ will be a Lasserre gap instance. To be precise, the $m$ constraints of $M$ are chosen independently (with replacement), where each constraint $\boldsymbol{Q} = (\boldsymbol{v}, \boldsymbol{b})$ is uniformly random in $V^k \times G^k$.

**Theorem G.7.** *Let $G$ be a finite abelian group, and $C$ a balanced pairwise independent subgroup of $G^k$ for some $k \geqslant 3$. Let $M$ be a random instance of* Additive-CSP$(C)$ *with $m = \gamma n$ constraints and $n$ variables. Then $M$ has resolution width $n/\gamma^{O(1)}$ with probability $1 - o_{n;\gamma}(1)$.*

*Proof sketch.* This follows by Tulsiani's proof [Tul09, Theorem 4.3]. As in [Tul09, Theorem 4.3], we need $M$ to be expanding (i.e. every set of $s \leqslant \Omega(1/\gamma)^{25}n$ constraints contains at least $(k - 6/5)s$ variables); the expansion property is guaranteed by [Tul09, Lemma A.1(2)]. In our setting, the number of variables involved in an equation $(\chi, z)$ is simply the degree of $\chi$.

Also, a subgroup $C \subseteq G^k$ has dual distance at least 3 (i.e. non-trivial characters in $C^\perp$ have degree at least 3) if and only if $C$ is balanced pairwise independent. To see this, for any $i \neq j \in [k]$, let $C^{ij} \triangleq \{(c_i, c_j) \mid c \in C\} \subseteq G^{\{i\}} \times G^{\{j\}} \cong G^2$ be the projection of $C$ to $i$ and $j$ coordinates. Balanced pairwise independence of $C$ means for all $i \neq j \in [k]$, we have $C^{ij} \cong G^2$, which is equivalent to $(C^{ij})^\perp = \{\mathbf{1}\} \subseteq \mathbb{T}$ by Proposition G.1(a) and the isomorphism $\hat{\Gamma} \cong \Gamma$ for any finite abelian group $\Gamma$. Now the condition $(C^{ij})^\perp = \{\mathbf{1}\}\ \forall i \neq j \in [k]$ is the same as non-trivial characters in $C^\perp$ having degree at least 3.

One can check that Tulsiani's proof goes through. We omit details. □

We summarize the result of this section in the next theorem, which follows by combining Theorem G.4, Theorem G.7 and [Tul09, Lemma A.1(1)], and choosing $\gamma = O(|G|^k/\varepsilon^2)$. This is a generalization of [Tul09, Theorem 4.6] and a Lasserre gap analogue of Theorem 1.1.

**Theorem G.8.** *Let $G$ be a finite abelian group, and $C$ be a balanced pairwise independent subgroup of $G^k$ for some $k \geqslant 3$. For any $\varepsilon > 0$, some* Additive-CSP$(C)$ *instance $M$ on $n$ variables has a* (poly$(\varepsilon/|G|^k) \cdot n$)*-round Lasserre solution of value 1 and satisfies* val$(M) \leqslant |C|/|G|^k + \varepsilon$.

## References


[ABH+05] Sanjeev Arora, Eli Berger, Elad Hazan, Guy Kindler, and Muli Safra. On non-approximability for quadratic programs. In *FOCS*, pages 206–215, 2005. 4

[ABS10] Sanjeev Arora, Boaz Barak, and David Steurer. Subexponential algorithms for unique games and related problems. In *Proceedings of the 2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, FOCS '10, pages 563–572, Washington, DC, USA, 2010. IEEE Computer Society. 5

[AFWZ95] Noga Alon, Uriel Feige, Avi Wigderson, and David Zuckerman. Derandomized graph products. *Computational Complexity*, 5(1):60–75, 1995. 2

[AKS11] Per Austrin, Subhash Khot, and Muli Safra. Inapproximability of vertex cover and independent set in bounded degree graphs. *Theory of Computing*, 7(1):27–43, 2011. 3

[AM09] Per Austrin and Elchanan Mossel. Approximation resistant predicates from pairwise independence. *Computational Complexity*, 18(2):249–271, 2009. 1, 2, 3, 4, 6, 24

[BBCR10] Boaz Barak, Mark Braverman, Xi Chen, and Anup Rao. How to compress interactive communication. In *Proceedings of the 42nd ACM symposium on Theory of computing*, STOC '10, pages 67–76, New York, NY, USA, 2010. ACM. 4

[BBLV09] Jop Briët, Harry Buhrman, Troy Lee, and Thomas Vidick. Multiplayer XOR games and quantum communication complexity with clique-wise entanglement. Manuscript at `http://arxiv.org/abs/0911.4007`, 2009. 6

[BGMT12] Siavosh Benabbas, Konstantinos Georgiou, Avner Magen, and Madhur Tulsiani. SDP gaps from pairwise independence. *Theory of Computing*, 8(1):269–289, 2012. 6

[BGS98] Mihir Bellare, Oded Goldreich, and Madhu Sudan. Free bits, PCPs, and nonapproximability — towards tight results. *SIAM Journal of Computing*, 27(3):804–915, June 1998. 3, 4

[BK09] Nikhil Bansal and Subhash Khot. Optimal long code test with one free bit. In *Proceedings of the 2009 50th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '09, pages 453–462, Washington, DC, USA, 2009. IEEE Computer Society. 3

[BK10] Nikhil Bansal and Subhash Khot. Inapproximability of hypergraph vertex cover and applications to scheduling problems. In *Proceedings of the 37th international colloquium conference on Automata, languages and programming*, ICALP'10, pages 250–261, Berlin, Heidelberg, 2010. Springer-Verlag. 3

[BV10] Andrej Bogdanov and Emanuele Viola. Pseudorandom bits for polynomials. *SIAM Journal on Computing*, 39(6):2464–2486, January 2010. 5

[CM08] Siu On Chan and Michael Molloy. A dichotomy theorem for the resolution complexity of random constraint satisfaction problems. In *Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '08, pages 634–643, Washington, DC, USA, 2008. IEEE Computer Society. 2

[CMM09] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Near-optimal algorithms for maximum constraint satisfaction problems. *ACM Transactions on Algorithms*, 5(3), 2009. 1, 3

[DH10] Irit Dinur and Prahladh Harsha. Property testing. chapter Composition of low-error 2-query PCPs using decodable PCPs, pages 280–288. Springer-Verlag, Berlin, Heidelberg, 2010. 17

[DKPS10] Irit Dinur, Subhash Khot, Will Perkins, and Muli Safra. Hardness of finding independent sets in almost 3-colorable graphs. In *FOCS*, pages 212–221. IEEE Computer Society, 2010. 2, 3, 10

[DS05] Irit Dinur and Shmuel Safra. On the hardness of approximating minimum vertex cover. *Annals of Mathematics*, 162(1):439–485, 2005. 18

[EH00] Lars Engebretsen and Jonas Holmerin. Clique is hard to approximate within $n^{1-o(1)}$. In Ugo Montanari, José D. P. Rolim, and Emo Welzl, editors, *ICALP*, volume 1853 of *Lecture Notes in Computer Science*, pages 2–12. Springer, 2000. 1

[EH08] Lars Engebretsen and Jonas Holmerin. More efficient queries in PCPs for NP and improved approximation hardness of maximum CSP. *Random Structures and Algorithms*, 33(4):497–514, December 2008. 1, 2, 3, 4

[Eng05] Lars Engebretsen. The nonapproximability of non-boolean predicates. *SIAM Journal on Discrete Mathematics*, 18(1):114–129, January 2005. 2, 3

[FGL+96] Uriel Feige, Shafi Goldwasser, Laszlo Lovász, Shmuel Safra, and Mario Szegedy. Interactive proofs and the hardness of approximating cliques. *Journal of the ACM*, 43(2):268–292, March 1996. 10

[FGRW09] Vitaly Feldman, Venkatesan Guruswami, Prasad Raghavendra, and Yi Wu. Agnostic learning of monomials by halfspaces is hard. In *Proceedings of the 2009 50th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '09, pages 385–394, Washington, DC, USA, 2009. IEEE Computer Society. 4

[FK98] Uriel Feige and Joe Kilian. Zero knowledge and the chromatic number. *Journal of Computer and System Sciences*, 57(2):187–199, October 1998. 10

[GNW11] Oded Goldreich, Noam Nisan, and Avi Wigderson. On Yao's XOR-lemma. In Oded Goldreich, editor, *Studies in Complexity and Cryptography*, volume 6650 of *Lecture Notes in Computer Science*, pages 273–301. Springer, 2011. 4

[GR08] Venkatesan Guruswami and Prasad Raghavendra. Constraint satisfaction over a non-boolean domain: Approximation algorithms and unique-games hardness. In *APPROX-RANDOM*, pages 77–90, 2008. 3, 4, 6

[GRSW12] Venkatesan Guruswami, Prasad Raghavendra, Rishi Saket, and Yi Wu. Bypassing UGC from some optimal geometric inapproximability results. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '12, pages 699–717. SIAM, 2012. 4

[Hal98] Magnús M. Halldórsson. Approximations of independent sets in graphs. In *APPROX*, pages 1–13, 1998. 3

[Hal02] Eran Halperin. Improved approximation algorithms for the vertex cover problem in graphs and hypergraphs. *SIAM Journal of Computing*, 31(5):1608–1623, May 2002. 3

[Hås99] Johan Håstad. Clique is hard to approximate within $n^{1-\epsilon}$. *Acta Mathematica*, 182(1):105–142, March 1999. 4

[Hås01] Johan Håstad. Some optimal inapproximability results. *Journal of the ACM*, 48(4):798–859, July 2001. 1, 17, 20, 22

[Has05a] Gustav Hast. Approximating Max $k$CSP - outperforming a random assignment with almost a linear factor. In Luís Caires, Giuseppe F. Italiano, Luís Monteiro, Catuscia Palamidessi, and Moti Yung, editors, *ICALP*, volume 3580 of *Lecture Notes in Computer Science*, pages 956–968. Springer, 2005. 3

[Has05b] Gustav Hast. *Beating a Random Assignment*. PhD thesis, KTH, Stockholm, 2005. 1

[Hås09] Johan Håstad. On the approximation resistance of a random predicate. *Computational Complexity*, 18(3):413–434, October 2009. 2, 4, 12, 20

[Hol09] Thomas Holenstein. Parallel repetition: Simplification and the no-signaling case. *Theory of Computing*, 5(1):141–172, 2009. 2, 4

[HR94] Edwin Hewitt and Kenneth A. Ross. *Abstract Harmonic Analysis: Volume 1: Structure of Topological Groups. Integration Theory. Group Representations*. Abstract Harmonic Analysis. Springer, 2nd edition, 1994. 24

[IPZ01] Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. Which problems have strongly exponential complexity? *Journal of Computer and System Sciences*, 63(4):512–530, 2001. 5

[Kho01] Subhash Khot. Improved inapproximability results for maxclique, chromatic number and approximate graph coloring. In *Proceedings of the 42nd IEEE symposium on Foundations of Computer Science*, FOCS '01, pages 600–, Washington, DC, USA, 2001. IEEE Computer Society. 3, 12

[Kho02a] Subhash Khot. Hardness results for coloring 3-colorable 3-uniform hypergraphs. In *Proceedings of the 43rd Symposium on Foundations of Computer Science*, FOCS '02, pages 23–32, Washington, DC, USA, 2002. IEEE Computer Society. 4

[Kho02b] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, STOC '02, pages 767–775, New York, NY, USA, 2002. ACM. 1

[KM11] Subhash Khot and Dana Moshkovitz. NP-hardness of approximately solving linear equations over reals. In *Proceedings of the 43rd annual ACM symposium on Theory of computing*, STOC '11, pages 413–420, New York, NY, USA, 2011. ACM. 4

[KS11] Subhash Khot and Muli Safra. A two prover one round game with strong soundness. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 648–657. IEEE, October 2011. 2, 4

[KS12] Subhash Khot and Rishi Saket. Hardness of finding independent sets in almost $q$-colorable graphs. In *FOCS*, 2012. To appear. 2, 3, 10, 12

[Las01] Jean B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001. 2

[Lov08] Shachar Lovett. Lower bounds for adaptive linearity tests. In Susanne Albers and Pascal Weil, editors, *STACS*, volume 1 of *LIPIcs*, pages 515–526. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany, 2008. 12

[MM12] Konstantin Makarychev and Yury Makarychev. Approximation algorithm for non-boolean MAX $k$-CSP. In Anupam Gupta, Klaus Jansen, José D. P. Rolim, and Rocco A. Servedio, editors, *APPROX-RANDOM*, volume 7408 of *Lecture Notes in Computer Science*, pages 254–265. Springer, 2012. 1, 3, 23

[MOO10] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: Invariance and optimality. *Annals of Mathematics*, 171(1), 2010. 12

[Mos10] Elchanan Mossel. Gaussian bounds for noise correlation of functions. *Geometric and Functional Analysis*, 19:1713–1756, 2010. 9, 12, 14

[MR10] Dana Moshkovitz and Ran Raz. Two-query PCP with subconstant error. *Journal of the ACM*, 57(5):29:1–29:29, June 2010. 17

[Ole03] Krzysztof Oleszkiewicz. On a nonsymmetric version of the Khinchine–Kahane inequality. In Evariste Giné, Christian Houdré, and Nualart David, editors, *Stochastic Inequalities and Applications*, volume 56, pages 157–168. Birkhäuser Basel, 2003. 14

[OW12] Ryan O'Donnell and John Wright. A new point of NP-hardness for unique games. In *Proceedings of the 44th symposium on Theory of Computing*, STOC '12, pages 289–306, New York, NY, USA, 2012. ACM. 4, 9, 12, 14, 17

[OWZ11] Ryan O'Donnell, Yi Wu, and Yuan Zhou. Hardness of Max-2Lin and Max-3Lin over integers, reals, and large cyclic groups. In *26th Annual Conference on Computational Complexity*, CCC, pages 23–33. IEEE Computer Society, 2011. 17

[Rag08] Prasad Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proceedings of the 40th annual ACM symposium on Theory of computing*, STOC '08, pages 245–254, New York, NY, USA, 2008. ACM. 1, 4

[Rao08] Anup Rao. Parallel repetition in projection games and a concentration bound. In *Proceedings of the 40th annual ACM symposium on Theory of computing*, STOC '08, pages 1–10, New York, NY, USA, 2008. ACM. 2, 4

[Raz98] Ran Raz. A parallel repetition theorem. *SIAM Journal of Computing*, 27(3):763–803, June 1998. 2, 4

[Sch08] Grant Schoenebeck. Linear level Lasserre lower bounds for certain $k$-CSPs. In *Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '08, pages 593–602, Washington, DC, USA, 2008. IEEE Computer Society. Newer version available at author's homepage. 2, 24

[She12] Alexander A. Sherstov. The multiparty communication complexity of set disjointness. In *Proceedings of the 44th symposium on Theory of Computing*, STOC '12, pages 525–548, New York, NY, USA, 2012. ACM. 4

[SSS95] Jeanette P. Schmidt, Alan Siegel, and Aravind Srinivasan. Chernoff-Hoeffding bounds for applications with limited independence. *SIAM Journal on Discrete Mathematics*, 8(2):223–250, 1995. 11

[ST98] Madhu Sudan and Luca Trevisan. Probabilistically checkable proofs with low amortized query complexity. In *FOCS*, pages 18–27. IEEE Computer Society, 1998. 4

[ST00] Alex Samorodnitsky and Luca Trevisan. A PCP characterization of NP with optimal amortized query complexity. In *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, STOC '00, pages 191–199, New York, NY, USA, 2000. ACM. 1, 2, 3, 4

[ST09] Alex Samorodnitsky and Luca Trevisan. Gowers uniformity, influence of variables, and PCPs. *SIAM Journal of Computing*, 39(1):323–360, 2009. 1, 2, 4, 6, 11, 12, 18, 20

[Tre98] Luca Trevisan. Recycling queries in PCPs and in linearity tests. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, STOC '98, pages 299–308, New York, NY, USA, 1998. ACM. Also avaliable at `http://eccc.hpi-web.de/report/1998/007/`. 4

[Tre01] Luca Trevisan. Non-approximability results for optimization problems on bounded degree instances. In *Proceedings of the thirty-third annual ACM symposium on Theory of computing*, STOC '01, pages 453–461, New York, NY, USA, 2001. ACM. 2, 3, 18, 19

[Tul09] Madhur Tulsiani. CSP gaps and reductions in the Lasserre hierarchy. In *Proceedings of the 41st annual ACM symposium on Theory of computing*, STOC '09, pages 303–312, New York, NY, USA, 2009. ACM. 2, 3, 6, 24, 25, 27

[Tul12] Madhur Tulsiani. Personal communication, May 2012. 6

[TW12] Madhur Tulsiani and Pratik Worah. LS+ lower bounds from pairwise independence. *Electronic Colloquium on Computational Complexity (ECCC)*, 2012. 6

[Wol07] Paweł Wolff. Hypercontractivity of simple random variables. *Studia Mathematica*, 180(3):219–236, 2007. 14

[Zwi98] Uri Zwick. Approximation algorithms for constraint satisfaction problems involving at most three variables per constraint. In *Proceedings of the ninth annual ACM-SIAM symposium on Discrete algorithms*, SODA '98, pages 201–210, Philadelphia, PA, USA, 1998. Society for Industrial and Applied Mathematics. 1

EECS, UC Berkeley
*E-mail address*: siuon@cs.berkeley.edu



ECCC ISSN 1433-8092
http://eccc.hpi-web.de