Electronic Colloquium on Computational Complexity, Report No. 141 (2017)

A Family of Dictatorship Tests with Perfect Completeness for 2–to–2 Label Cover

Joshua Brakensiek* Venkatesan Guruswami†

Abstract

We give a family of dictatorship tests with perfect completeness and low-soundness for 2–to–2 constraints. The associated 2–to–2 conjecture has been the basis of some previous inapproximability results with perfect completeness. However, evidence towards the conjecture in the form of integrality gaps even against weak semidefinite programs has been elusive. Our result provides some indication of the expressiveness and non-triviality of 2–to–2 constraints, given the close connections between dictatorship tests and satisfiability and approximability of CSPs.

1 Introduction

In the study of constraint satisfaction problems and optimization, dictatorship tests are a surrogate for understanding the problems' computational complexity. For a given optimization problem over a domain D, a dictatorship test for a function $f : D^n \to D$ is a randomized query who distribution is based on the constraints from the optimization problem. For example, for NAE-3-SAT, $D = \{0, 1\}$, a query would pick uniformly at random three vectors $x, y, z \in \{0, 1\}^n$ such that $x_i \vee y_i \vee z_i = 1$ and $x_i \wedge y_i \wedge z_i = 0$ for all $i \in \{1, \ldots, n\}$ and then test whether $f(x) \vee f(y) \vee f(z) = 1$ and $f(x) \wedge f(y) \wedge f(z) = 0$. For a given function f, there is an associated probability $p \in [0, 1]$ for which it satisfies the dictatorship test. Typically some threshold $\tau > 0$ is assigned for which a function f 'passes' the test if $p \geq \tau$. As their name suggests, these tests are usually designed so that certain *dictator functions* (also known as projections), functions which depend on only a single coordinate, pass with probability 1 (or $1 - \epsilon$, depending on the application).

For constraint satisfaction problems (CSPs), dictatorship tests with a passing probability of $\tau = 1$ have been used to understand the computational complexity of the associated satisfiability problems. In the universal algebraic approach to understanding CSPs, functions which pass such dictatorship tests are known as *polymorphisms*. (See [Che09] for an excellent survey on the algebraic study of CSPs.) For a long time (e.g., [BJK05]), it has been known that the computational complexity is completely characterized by which functions pass the dictatorship tests. The deep understanding of such connections has very recently culminated in two proofs of the Algebraic Dichotomy Conjecture [Bul17, Zhu17]. Although the exact statement of the conjecture is quite technical, at a moral level it says that a CSP has a polynomial time algorithm if and only if one is able to satisfy the canonical dictatorship test with arbitrarily complex polymorphisms. (Formally, such functions are known as "weak near unanimity" operators.)

In the context of optimization, where the goal is to satisfy as many constraints of the CSP as possible, one often seeks to prove that for some threshold $\tau \in (0, 1)$, if a function f passes the dictatorship test

*Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA. Email: jbrakens@andrew.cmu.edu. Research supported in part by an REU supplement to NSF CCF-1526092.

†Computer Science Department, Carnegie Mellon University, Pittsburgh, PA 15213. Email: guruswami@cmu.edu. Research supported in part by NSF grants CCF-1422045 and CCF-1526092.

ISSN 1433-8092

with probability at least τ, then f essentially depends on a small number of coordinates. In this setting, a similar close connection between approximability and the passing of dictatorship tests is also known. In particular, Raghavendra [Rag08] showed that the threshold τ for which dictatorship tests can be passed by highly non-dictatorial functions (those with no influential coordinates), is exactly the approximation ratio which can be found by the Basic SDP relaxation. Further, under Khot's Unique Games Conjecture [Kho02], no polynomial time algorithm can deliver an approximation guarantee better than τ [KKMO07].

The complexity of approximate optimization of every CSP is thus settled under the Unique Games Conjecture (UGC). The UGC itself asserts the inapproximability of a particular CSP called Unique Games. The fascinating aspect of this landscape is that the hardness of the single Unique Games problem reduces the complexity-theoretic task of showing inapproximability of every other CSP to the combinatorial/analytic task of designing a good dictatorship test that works with checks allowed by the CSP. While we do not seem to be close to a proof of the UGC, there is at least some evidence of the difficulty of the Unique Games problem, in the form of integrality gaps against strong semidefinite programming relaxations (from the so-called SA+ hierarchy) [KV15, RS09, KS09].

The Unique Games problem is a special case of a more general family of CSPs called Label Cover (see Section 2.1 for formal definitions relating to Label Cover, but briefly these are arity two CSPs whose constraint relations are unions of disjoint bipartite graphs). As the decision version of Unique Games is trivially polynomial time solvable, the UGC is useless in understanding the complexity of approximating satisfiable instances of CSPs (which is an important goal in itself, and also crucial for problems such as graph and hypergraph coloring). The most general form of Label Cover has also led to numerous strong inapproximability results, but without some restriction on the structure of the allowed constraints, it is not flexible enough to give a broad understanding similar to what the UGC provides.

Towards this end, Khot himself in his original UGC paper [Kho02] proposed a restricted form of Label Cover called 2-to-2 Label cover as a possible UG surrogate in some settings where inapproximability for satisfiable CSPs is desired. Unfortunately, for 2–to–2 Label Cover with perfect completeness (see a formal definition in Section 2.1), we do not know integrality gaps even for the standard SDP relaxation; the only gap known is for a weaker form where even simple non-negativity of inner products is not enforced [GKO$^+$10]. In this paper, we give a dictatorship test with perfect completeness and arbitrarily small soundness for 2–to–2 Label Cover (a closely related version to 2–to–1 Label Cover that is more symmetric, see [DMR09, Thm A.3] for a reduction).

The main theorem we prove is informally as follows. The formal version appears as Theorem 3.5.

Theorem 1.1 (Informal)**.** *There is a a family of dictatorship tests with* 2*–to–*2 *constraints, which have perfect completeness (i.e., dictators pass the test with probability* 1*), and if a function f passes one of these tests with noticeable probability, then f has a coordinate that has noticeable low-degree influence.*

While dictatorship tests are probably not the combinatorial/analytic core underlying inapproximability in the perfect completeness land, nevertheless our result can be viewed as some indication of the expressiveness and non-triviality of 2–to–1 constraints.

Related work. Other results proved using 2–to–1 or 2–to–2 conjecture include hardness results on finding independent sets in 3-uniform 2-colorable hypergraphs [KS14] and max k-coloring [GS13]. For the Max-k-CSP problem, dictatorship tests are known to exist within a constant factor of known polynomial time approximation algorithms [TY10, BKT17], although NP-hardness results have an exponential gap. In fact, only very recently has a Unique Games-surrogate been found which closes this gap up to a polynomial factor [BG17].

In the case of the Not-Two (NTW) predicate, defined as the elements of $\{0,1\}^3$ which have Hamming weight not equal to 2, a series of papers [OW09a, OW09b, Hå14] went from a dictatorship test to conditional

hardness (d–to–1 conjecture) to finally unconditional hardness. The authors hope this family of dictatorship tests likewise provides valuable insights on the difficulty of the 2–to–2 conjecture, and will perhaps aid in attempts to resolve the 2–to–2 conjecture. Note that very recently there has been other progress [KMS17, DKK+16] toward the 2–to–2 conjecture, although they consider a variant without perfect completeness.

Note that nontrivial approximation algorithms are known for the 2–to–1 Label Cover, although not strong enough to refute the 2–to–1 conjecture. In particular, for all $\epsilon > 0$, [CMM06] found a $(L)^{-\frac{\sqrt{2}-1+\epsilon}{\sqrt{2}+1-\epsilon}}$ approximation algorithm for 2–to–1 Label Cover (where L is the label size of the instance) using an SDP relaxation.

2 Preliminaries

We first formally define what the 2–to–2 Label Cover problem and its corresponding conjecture are. Subsequently we review the Fourier analytical tools which shall be used to understand the dictatorship tests.

2.1 The d_1–to–d_2 Label Cover problems

For a positive integer L, we let $[L]$ denote the set $\{1, 2, \ldots, L\}$.

Let d_1, d_2, L_1, and L_2 be positive integers. For any relation $R \subseteq [L_1] \times [L_2]$ we may consider the bipartite graph whose two sides are identified with $[L_1]$ and $[L_2]$ respectively, and two vertices $u \in [L_1]$ and $v \in [L_2]$ are connected by an edge if $(u, v) \in R$. We say that R is an *d_1–to–d_2* constraint if all connected components of this bipartite graph are complete bipartite and have at most d_1 components on the $[L_1]$ side and at most d_2 vertices on the $[L_2]$ side. Note in particular that a d_1'–to–d_2' relation is a d_1–to–d_2 relation as long as $d_1' \leq d_1$ and $d_2' \leq d_2$.

An instance of d_1–to–d_2 Label Cover consists of a bipartite graph $G = (U, V, E)$, positive integers L_1 and L_2, and constraints d_1–to–d_2 constraints $\pi_{(u,v)} \subseteq [L_1] \times [L_2]$ or all $(u, v) \in E$. The 'goal' is to find labelings $\sigma : U \to [L_1]$ and $\psi : V \to [L_2]$ such that for all $(u, v) \in E$, $(\sigma(u), \psi(v)) \in \pi_{(u,v)}$. Any edge which satisfies this property is said to be *satisfied* the labelings σ and ψ.

Conjecture 2.1 (d_1–to–d_2 conjectures [Kho02])**.** *Let $d_1, d_2 \geq 1$ be positive integers and $\epsilon > 0$ be a constant.*

The following promise decision problem is NP-hard. Let $\Psi = (U, V, E, L, \pi_{(u,v)})$ be an instance of d_1–to–d_2 Label Cover. Distinguish between the following cases.

- ***Completeness.*** *There exist labelings $\sigma : U \to [L_1]$ and $\psi : V \to [L_2]$ which satisfy all of the constraints if $(d_1, d_2) \neq (1, 1)$ or at least $1 - \epsilon$ fraction of the constraints if $d_1 = d_2 = 1$.*
- ***Soundness.*** *For all labelings at most ϵ fraction of the constraints are satisfied.*

The 1–to–1 conjecture is more well-known as the Unique Games Conjecture.

2.2 Fourier Analysis over $\mathbb{Z}_p$

Before we describe the dictatorship tests, we first remind the reader of some the basics of Fourier Analysis over $\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$, with addition modulo p, where p is an odd prime.

Let $L^2(\mathbb{Z}_p^n; \mathbb{C})$ be the vector space of functions $f : \mathbb{Z}_p^n \to \mathbb{C}$ with inner product

$$\langle f, g \rangle = \frac{1}{p^n} \sum_{x \in \mathbb{Z}_p^n} f(x)\overline{g(x)}. \tag{1}$$

The Fourier basis we pick for $\mathbb{Z}_p$ is the set of characters $\{\chi_0, \chi_1, \ldots, \chi_{p-1}\}$, where $\chi_i(j) = \zeta^{ij}$, in which ζ is the primitive pth root of unity for all $i, j \in \mathbb{Z}_p$. We extend this basis to $\mathbb{Z}_p^n$ multiplicatively. For all $\alpha, x \in \mathbb{Z}_p^n$, let

$$\chi_\alpha(x) = \prod_{i=1}^{n} \chi_{\alpha_i}(x_i).$$

It is well-known that the χ_α's form an orthonormal basis with respect to the inner product (1). For a function $f : \mathbb{Z}_p^n \to \mathbb{C}$, its *Fourier coefficients* $\hat{f} : \mathbb{Z}_p^n \to \mathbb{C}$ are the unique function which satisfies

$$f = \sum_{\alpha \in \mathbb{Z}_p^n} \hat{f}(\alpha)\chi_\alpha.$$

For $\alpha \in \mathbb{Z}_p^n$, we let $|\alpha| = \{i \in [n] : \alpha_i \neq 0\}$. We can now state/define few basic properties of a $f : \mathbb{Z}_p^n \to \mathbb{C}$ (e.g., from [O'D14]):

$$\begin{aligned}
\mathbb{E}[f] &= \hat{f}(0, \ldots, 0) \\
\mathrm{Var}[f] &= \sum_{\alpha \in \mathbb{Z}_p^n, \alpha \neq (0,\ldots,0)} |\hat{f}(\alpha)|^2 \\
\mathrm{Inf}_i(f) &= \sum_{\alpha \in \mathbb{Z}_p^n, \alpha_i \neq 0} |\hat{f}(\alpha)|^2 \\
\mathrm{Inf}_i^{\leq d}(f) &= \sum_{\alpha \in \mathbb{Z}_p^n, \alpha_i \neq 0, |\alpha| \leq d} |\hat{f}(\alpha)|^2.
\end{aligned}$$

In our analysis of the dictatorship tests, we need a few results from [MOO10, KKMO07] on the *noise stability* of our function. There are many definitions, but we only need the Fourier analytic one (e.g., Proposition 8.28 of [O'D14]).

Definition 2.1. For $f \in \mathbb{Z}_p^n \to \mathbb{R}$ and $\eta \in [-1, 1]$,

$$\mathbb{S}_\eta(f) = \sum_{\alpha \in \mathbb{Z}_p^n} \eta^{|\alpha|} |\hat{f}(\alpha)|^2,$$

where $0^0 = 1$.

We now cite a variant of the MOO theorem [MOO10] as stated in [KKMO07]

Theorem 2.2 (Low-degree MOO theorem [MOO10] from [KKMO07])**.** *Consider any $\eta \in [0, 1)$, $\epsilon > 0$, and $f : \mathbb{Z}_p^n \to [0, 1]$. If $\max_{i \in [n]} \mathrm{Inf}_i^{\leq d}(f) \leq \delta$, where $\delta(\epsilon, \eta, p) > 0$ and $d(\epsilon, \eta, p) \in \mathbb{N}$, then*

$$\mathbb{S}_\eta(f) \leq \Lambda_\eta(\mathbb{E}[f]) + \epsilon,$$

where $\Lambda_\eta(\mu)$ is the probability that two η-correlated Gaussians with mean 0 and variance 1 are in the upper μ fraction of their CDFs.

The only fact we need about Λ_η is the following upper bound.

Lemma 2.3 (Corollary 3 of [KKMO07])**.** *If $\eta \in (0, 1)$ and $q \geq 2$, then*

$$\Lambda_\eta\left(\frac{1}{q}\right) \leq \left(\frac{1}{q}\right)^{\frac{2}{1+\eta}}.$$

3 A Family of Dictatorship Tests for 2–to–2 Label Cover

In this section, we construct the dictatorship tests for 2–to–2 constraints, and then analyze this construction using Fourier analysis.

3.1 The dictatorship tests

Consider the following 2–to–2 constraint. Let $S := \{0, 1, \ldots, (p-1)/2\} \subseteq \mathbb{Z}_p$, where p is an odd prime. Then our constraints are

$$\pi_{a,b} := \bigcup_{x \in S} \{(x+a, x+b), (-x+a, x+b), (x+a, -x+b), (-x+a, -x+b)\} \subseteq \mathbb{Z}_p^2.$$

In fact, we can replace $x \in S$ with $x \in \mathbb{Z}_p$ in the above definition. When we identify $\pi_{a,b}$ as a probability distribution, we give double weight to the 1–to–1 constraint (a, b) so that the marginal distributions are uniform.

Our dictatorship test is then the following. Let $n \geq 1$ be a positive integer. The test is on a set of p^n variables, whose labeling is represented by a function $F : \mathbb{Z}_p^n \to \mathbb{Z}_p$. The test proceeds by the following algorithm.

- Draw $a, b \in \mathbb{Z}_p$ uniformly and independently at random.
- Draw $x, y \in \mathbb{Z}_p^n$ uniformly at random from $\pi_{a,b}^{\otimes n}$.
- Test $(F(x), F(y)) \in \pi_{a,b}$.

By a standard trick, we may assume that F is *folded*. That is, for all $a \in \mathbb{Z}_p$ and $x \in \mathbb{Z}_p^n$, $F(x+(a, \ldots, a)) = F(x) + a$. As usual, to enforce this, when querying $F(x)$, we instead query $F(x - (x_1, \ldots, x_1))$ and add x_1 to the result. (For convenience, we write $1_n = (1, \ldots, 1)$.) This technicality can be moved within the CSP, by adjusting checking $(F(x), F(y)) \in \pi_{a,b}$ to $(F(x - x_1 1_n), F(y - y_1 1_n)) \in \pi_{a-x_1, b-y_1}$.

Note that the dictator functions $F(x) = x_i$ for some $i \in [n]$ pass this test with probability 1. We now seek to show that if some F passes this test with probability at least ϵ, then F is 'close' to one of these dictator functions. Like O'Donnell and Wu [OW09a] we measure closeness in terms of having a coordinate with a significant low-degree influence.

3.2 Analysis of the Dictatorship Tests

Identify F with the function $f : \mathbb{Z}_p^n \to \mathbb{C}$, where $f(x) = 1$ iff $F(x) = 0$ and $f(x) = 0$ otherwise. Since F is folded, we must have that $\mathbb{E}[f] = \frac{1}{p}$. The success probability is

$$\begin{aligned}
\epsilon &\leq \mathop{\mathbb{E}}_{a,b}\left[\mathop{\mathbb{E}}_{(x,y)\sim\pi_{a,b}^{\otimes n}}[1[(F(x), F(y)) \in \pi_{a,b}]]\right] \\
&\leq \mathop{\mathbb{E}}_{a,b}\left[\mathop{\mathbb{E}}_{(x,y)\sim\pi_{a,b}^{\otimes n}}\left[2p \mathop{\mathbb{E}}_{(i,j)\sim\pi_{a,b}}[f(x - i1_n)\overline{f(y - j1_n)}]\right]\right] \text{ (double counting when } (i,j) = (a,b)) \\
&= 2p \mathop{\mathbb{E}}_{a,b}\left[\mathop{\mathbb{E}}_{(x,y)\sim\pi_{a,b}^{\otimes n}}\left[\mathop{\mathbb{E}}_{(i,j)\sim\pi_{a,b}}\left[\sum_{\alpha,\beta\in\mathbb{Z}_p^n} \hat{f}(\alpha)\hat{f}(\beta)\chi_\alpha(x - i1_n)\overline{\chi_\beta(y - j1_n)}\right]\right]\right]
\end{aligned}$$

$$= 2p \mathop{\mathbb{E}}_{a,b}\left[\mathop{\mathbb{E}}_{(i,j)\sim\pi_{a,b}}\left[\sum_{\alpha,\beta\in\mathbb{Z}_p^n}\hat{f}(\alpha)\hat{f}(\beta)\prod_{k=1}^{n}\mathop{\mathbb{E}}_{(x_k,y_k)\sim\pi_{a,b}}[\chi_{\alpha_k}(x_k-i)\overline{\chi_{\beta_k}(y_k-j)}]\right]\right]$$

$$= 2p \mathop{\mathbb{E}}_{a,b}\left[\mathop{\mathbb{E}}_{(i,j)\sim\pi_{a,b}}\left[\sum_{\alpha,\beta\in\mathbb{Z}_p^n}\hat{f}(\alpha)\hat{f}(\beta)\prod_{k=1}^{n}\mathop{\mathbb{E}}_{x_k\sim\mathbb{Z}_p}\left[\frac{\chi_{\alpha_k}(x_k-i+a)(\overline{\chi_{\beta_k}(x_k-j+b)}+\overline{\chi_{\beta_k}(-x_k-j+b)})}{2}\right]\right]\right]$$

$$= 2p \mathop{\mathbb{E}}_{a,b}\left[\mathop{\mathbb{E}}_{(i,j)\sim\pi_{a,b}}\left[\sum_{\alpha,\beta\in\mathbb{Z}_p^n}\hat{f}(\alpha)\hat{f}(\beta)\chi_\alpha((a-i)1_n)\overline{\chi_\beta((b-j)1_n)}\prod_{k=1}^{n}\mathop{\mathbb{E}}_{x_k\sim\mathbb{Z}_p}\left[\frac{\chi_{\alpha_k+\beta_k}(x_k)+\chi_{\alpha_k-\beta_k}(x_k)}{2}\right]\right]\right]$$

Note that

$$\mathop{\mathbb{E}}_{x_k\sim\mathbb{Z}_p}\left[\frac{\chi_{\alpha_k+\beta_k}(x_k)+\chi_{\alpha_k-\beta_k}(x_k)}{2}\right] = \begin{cases}1 & (\alpha_k,\beta_k)=(0,0)\\ 1/2 & \alpha_k^2=\beta_k^2\neq 0\\ 0 & \text{otherwise}\end{cases}$$

If $\alpha,\beta\in\mathbb{Z}_p^n$, define $\alpha\beta$ to be the pointwise product. Furthermore, define

$$C_\alpha = \{\alpha z\in\mathbb{Z}_p^n \mid z\in\{1,p-1\}^n\subseteq\mathbb{Z}_p^n\}.$$

Then, it is easy to see that the above probability equals

$$\epsilon \le 2p \mathop{\mathbb{E}}_{a,b}\left[\mathop{\mathbb{E}}_{(i,j)\sim\pi_{a,b}}\left[\sum_{\alpha\in\mathbb{Z}_p^n}\sum_{\beta\in C_\alpha}\hat{f}(\alpha)\hat{f}(\beta)\chi_\alpha((a-i)1_n)\chi_\beta((b-j)1_n)2^{-|\alpha|}\right]\right]$$

$$= 2p\sum_{\alpha\in\mathbb{Z}_p^n}\sum_{\beta\in C_\alpha}\hat{f}(\alpha)\hat{f}(\beta)2^{-|\alpha|}\mathop{\mathbb{E}}_{a,b}\left[\mathop{\mathbb{E}}_{(i,j)\sim\pi_{a,b}}\left[\chi_\alpha((a-i)1_n)\chi_\beta((b-j)1_n)\right]\right].$$

Note that $(i,j)\in\pi_{a,b}$ if and only if $a-i=\pm(b-j)$. In fact, we can rewrite the above expression as

$$\epsilon \le 2p\sum_{\alpha\in\mathbb{Z}_p^n}\sum_{\beta\in C_\alpha}\hat{f}(\alpha)\hat{f}(\beta)2^{-|\alpha|}\mathop{\mathbb{E}}_{c\sim\mathbb{Z}_p}\left[\chi_\alpha(c1_n)\frac{\chi_\beta(c1_n)+\chi_\beta(-c1_n)}{2}\right]$$

$$= p\sum_{\alpha\in\mathbb{Z}_p^n}\sum_{\beta\in C_\alpha}\hat{f}(\alpha)\hat{f}(\beta)2^{-|\alpha|}(1[\chi_{\alpha+\beta}(1_n)=1]+1[\chi_{\alpha-\beta}(1_n)=1])$$

Note that the C_α's are equivalence classes of size $2^{|\alpha|}$. In particular, $\beta\in C_\alpha$ iff $\alpha\in C_\beta$. Let $\mathcal{C}=\{C_\alpha:\alpha\in\mathbb{Z}_p^n\}$. Thus, we then have that

$$\epsilon \le p\sum_{C_\gamma\in\mathcal{C}}\frac{1}{|C_\gamma|}\sum_{\alpha,\beta\in C_\gamma}\hat{f}(\alpha)\hat{f}(\beta)(1[\chi_{\alpha+\beta}(1_n)=1]+1[\chi_{\alpha-\beta}(1_n)=1])$$

Let $\hat{f}(C_\gamma)$ be the vector of length $|C_\gamma|$ whose entries are $\hat{f}(\alpha)$. Let M^γ be the $|C_\gamma|\times|C_\gamma|$ matrix with entries

$$M^\gamma_{\alpha,\beta} = \frac{1[\chi_{\alpha+\beta}(1_n)=1]+1[\chi_{\alpha-\beta}(1_n)=1]}{|C_\gamma|}.$$

One can verify that M^γ is a real, symmetric, doubly stochastic matrix (modulo scaling of the entries). Thus, the largest eigenvalue of M^γ, λ^γ is this common row/column sum.[1] Then, we can see that[2]

$$\epsilon \le p\sum_{C_\gamma\in\mathcal{C}}\hat{f}(C_\gamma)^* M^\gamma \hat{f}(C_\gamma)$$

[1]For example, see [hs11].

[2]This deduction non-obviously uses the facts that $\bar{\hat{f}}(\alpha)=\hat{f}(-\alpha)$, $-\alpha\in C_\alpha$, and $M^\gamma_{\alpha,\beta}=M^\gamma_{\alpha,-\beta}$

$$\leq p \sum_{C_\gamma \in C} \lambda^\gamma \|\hat{f}(C_\gamma)\|^2$$
$$= p \sum_{C_\gamma \in C} \sum_{\alpha \in C_\gamma} |\hat{f}(\alpha)|^2 \lambda^\gamma. \tag{2}$$

We next seek to show that λ^γ is bounded by a particular decreasing function of $|\gamma|$.

Lemma 3.1. *There exists a universal constant $C > 0$ such that for all $n \in \mathbb{N}$, p odd prime, and $\gamma \in \mathbb{Z}_p^n$:*

$$\lambda^\gamma \leq \frac{2}{p} + \frac{C}{|\gamma|^{1/4} + 1}.$$

Before we can do that, we need the following result.

Lemma 3.2. *There exists a universal constant $C > 0$ such that for all $n \in \mathbb{N}$, p prime, and $\beta \in \mathbb{Z}_p^n$:*

$$\Pr_{z \sim \{0,1\}^n}[\chi_\beta(z) = 1] = \Pr_{z \sim \{0,1\}^n}\left[\sum_{i=1}^n \beta_i z_i = 0\right] \leq \frac{1}{p} + \frac{C}{|\beta|^{1/4} + 1}$$

Remark. This lemma can be considered a finite-field analogue of a result of Littlewood and Offord [LO43].

The the exponent of $1/4$ can in fact be improved to $1/2 - o(1)$, but the former is all we need.

Proof. Let $k = |\beta|$. Assume $k \geq 1$. Assume without loss of generality that the nonzero terms of β are $\beta_1, \ldots, \beta_k$.Then

$$\Pr_{z \sim \{0,1\}^n}\left[\sum_{i=1}^n \beta_i z_i = 0\right] = \mathop{\mathbb{E}}_{z \sim \{0,1\}^n}\left[\mathop{\mathbb{E}}_{j \sim \mathbb{Z}_p}\left[\zeta^{j \sum_{i=1}^n \beta_i z_i}\right]\right]$$
$$= \mathop{\mathbb{E}}_{j \sim \mathbb{Z}_p}\left[\prod_{i=1}^k \frac{1 + \zeta^{j\beta_i}}{2}\right]$$
$$\leq \frac{1}{p} \sum_{j=0}^{p-1} \prod_{i=1}^k \left|\frac{1 + \zeta^{j\beta_i}}{2}\right|.$$

Since the β_i are nonzero for $i \in [k]$, we have that for a fixed i, $1 + \zeta^{j\beta_i}$ is a permutation of $1 + \zeta^0, \ldots, 1 + \zeta^{p-1}$. By the rearrangement inequality, the above product is maximized when all the permutations are consistent. That is,

$$\frac{1}{p} \sum_{i=0}^{p-1} \prod_{j=1}^k \left|\frac{1 + \zeta^{i\beta_j}}{2}\right| \leq \frac{1}{p} \sum_{j=0}^{p-1} \left|\frac{1 + \zeta^j}{2}\right|^k$$
$$= \frac{1}{p} \sum_{j=0}^{p-1} \left|\frac{1 + \cos(2\pi j/p) + i \sin(2\pi j/p)}{2}\right|^k$$
$$\leq \frac{1}{p} \sum_{j=0}^{p-1} |\cos(\pi j/p)|^k$$
$$= \frac{1}{p} + \frac{2}{p} \sum_{j=1}^{(p-1)/2} \cos^k(\pi j/p).$$

To further bound, we need the following elementary inequality

Claim 3.3. *For all $\theta \in [0, \pi/2]$,*

$$\cos\theta \le 1 - \frac{\theta^2}{4}.$$

Proof. Note that $\sin\eta \ge \eta/2$ for all $\eta \in [0, \pi/2]$ since $\sin\eta - \eta/2$ is a concave function in $[0, \pi/2]$ which is nonnegative at its endpoints. We then integrate this from 0 to θ to obtain the above inequality. □

Thus, for all $j \in \{1, \dots, (p-1)/2\}$,

$$\cos^k(\pi j/p) \le \left(1 - \frac{\pi^2 j^2}{4p^2}\right)^k$$

Let $\ell = \lfloor 2p/(k^{1/4}) \rfloor$. Then, we have that

$$\begin{aligned}
\frac{1}{p} + \frac{2}{p}\sum_{j=1}^{(p-1)/2} \cos^k(\pi j/p) &\le \frac{1}{p} + \frac{2}{p}\sum_{j=1}^{(p-1)/2} \left(1 - \frac{\pi^2 j^2}{4p^2}\right)^k \\
&\le \frac{1}{p} + \frac{2\ell}{p} + \max\left(\frac{2((p-1)/2 - \ell)}{p}\left(1 - \frac{\pi^2(\ell+1)^2}{4p^2}\right)^k, 0\right) \\
&\le \frac{1}{p} + \frac{4}{k^{1/4}} + \max\left(1 - \frac{\pi^2}{k^{1/2}}, 0\right)^k \\
&\le \frac{1}{p} + \frac{4}{k^{1/4}} + e^{-\pi^2\sqrt{k}} \\
&\le \frac{1}{p} + \frac{C'}{k^{1/4}+1},
\end{aligned}$$

for some universal constant $C' > 0$. To account for the case $k = 0$, we may set $C = \max(C', 1)$. □

Proof of Lemma 3.1. By the previously mentioned symmetries of M^γ,

$$\begin{aligned}
\lambda^\gamma &= \mathop{\mathbb{E}}_{\alpha\sim C_\gamma}\left[1[\chi_{\alpha+\gamma}(1_n) = 1] + 1[\chi_{\alpha-\gamma}(1_n) = 1]\right] \\
&= 2\mathop{\mathbb{E}}_{\alpha\sim C_\gamma}\left[1[\chi_{\alpha-\gamma}(1_n) = 1]\right] \\
&= 2\mathop{\mathbb{E}}_{z\sim\{-1,1\}^n}\left[1[\chi_{\gamma z-\gamma}(1_n) = 1]\right] \\
&= 2\mathop{\mathbb{E}}_{z\sim\{-1,1\}^n}\left[1[\chi_{\gamma}(z - 1_n) = 1]\right] \\
&= 2\mathop{\mathbb{E}}_{z'\sim\{0,1\}^n}\left[1[\chi_{\gamma}((1_n - 2z') - 1_n) = 1]\right] \\
&= 2\mathop{\mathbb{E}}_{z'\sim\{0,1\}^n}\left[1[\chi_{\gamma}(z') = 1]\right] \qquad \text{(since } p \text{ odd prime)} \\
&\le \frac{2}{p} + \frac{2C}{|\gamma|^{1/4}+1},
\end{aligned}$$

where the last step follows from Lemma 3.2. □

We also need to invoke the following analytical lemma.

Lemma 3.4. *For all $\epsilon > 0$ and $L \in \mathbb{N}$, there exists $P(\epsilon, L) \in \mathbb{N}$, such that for all primes $p \geq P$, there exists $\delta(\epsilon, L, p) > 0$ and $d(\epsilon, L, p) \in \mathbb{N}$ such that for all $f : \mathbb{Z}_p^n \to \{0, 1\}$ with $\mathbb{E}[f] = 1/p$ the following holds. If $\mathrm{Inf}_i^{\leq d}(f) \leq \delta$ for all $i \in [n]$ then*

$$\sum_{\alpha \in \mathbb{Z}_p^n, |\alpha| \leq L} |\hat{f}(\alpha)|^2 \leq \epsilon \, \mathbb{E}[f] = \frac{\epsilon}{p}.$$

Remark. The LHS does not fundamentally depend on the choice of Fourier basis for $\mathbb{Z}_p$. For example, the Efron-Stein decomposition [ES81] works just as well.

Proof. First, we set the parameters

$$\begin{aligned}
\eta &= 1/2 \\
P &= \lceil (\epsilon \eta^L)^{-\frac{1+\eta}{1-\eta}} \rceil + 1 \\
\epsilon' &= \frac{\epsilon \eta^L}{p} - \left(\frac{1}{p}\right)^{\frac{2}{1+\eta}} \text{ (positive since } p \geq P)
\end{aligned}$$

Let d and δ be the parameters guaranteed by the modified MOO theorem (Theorem 2.2) where their (ϵ, η, p) is our (ϵ', η, p). If $\max_{i \in [n]} \mathrm{Inf}_i^{\leq d}(f) \leq \delta$, by a standard argument

$$\begin{aligned}
\sum_{\alpha \in \mathbb{Z}_p^n, |\alpha| \leq L} |\hat{f}(\alpha)|^2 &\leq \sum_{\alpha \in \mathbb{Z}_p^n} \eta^{|\alpha| - L} |\hat{f}(\alpha)|^2 \\
&\leq \eta^{-L} \, \mathbb{S}_\eta(f) \\
&\leq \eta^{-L} (\Lambda_\eta(1/p) + \epsilon') \text{ (modified MOO theorem, since } \max_i \mathrm{Inf}_i^{\leq d}(f) \leq \delta) \\
&\leq \eta^{-L} \left(\left(\frac{1}{p}\right)^{\frac{2}{1+\eta}} + \epsilon' \right) \text{ (Lemma 2.3)} \\
&= \eta^{-L} \frac{\epsilon \eta^L}{p} = \frac{\epsilon}{p},
\end{aligned}$$

as desired. □

Now, we may establish the main result.

Theorem 3.5. *For every $\epsilon > 0$, there exists $P \in \mathbb{N}$ such that for all primes $p \geq P$, there exists $\delta > 0$ and $d \in \mathbb{N}$ such that the following holds. If $F : \mathbb{Z}_p^n \to \mathbb{Z}_p$ is a folded function which passes the 2–to–2 Label Cover modulo p dictatorship test with probability at least ϵ, then*

$$\max_{i \in [n]} \mathrm{Inf}_i^{\leq d} f \geq \delta,$$

where $f(x) = 1[F(x) = 0]$.

Proof. First select the following parameters

$$\begin{aligned}
\epsilon' &= \epsilon/(8C) \\
L &= \left\lceil \left(\frac{1}{\epsilon'}\right)^4 \right\rceil.
\end{aligned}$$

Let P' be the P guaranteed by Lemma 3.4 when $\epsilon = \epsilon'/2$ and $L = L$. Let $P = \max(P', 4/\epsilon)$. Assume $p \geq P$. Let δ and d be the values (independent of n) guaranteed by Lemma 3.4 for our choice of p.

From our previous analysis (2), we know that

$$\begin{aligned}
\epsilon &\leq p \sum_{C_\gamma \in C} \sum_{\alpha \in C_\gamma} |\hat{f}(\alpha)|^2 \lambda^\gamma(2) \\
&\leq p \sum_{C_\gamma \in C} \sum_{\alpha \in C_\gamma} |\hat{f}(\alpha)|^2 \left(\frac{2}{p} + \frac{C}{|\alpha|^{1/4}+1} \right) \text{(Lemma 3.1)}. \\
&= 2 \sum_{\alpha \in \mathbb{Z}_p^n} |\hat{f}(\alpha)|^2 + Cp \sum_{\alpha \in \mathbb{Z}_p^n} \frac{|\hat{f}(\alpha)|^2}{|\alpha|^{1/4}+1}.
\end{aligned}$$

By Parseval's identity,

$$\sum_{\alpha \in \mathbb{Z}_p^n} |\hat{f}(\alpha)|^2 = \frac{1}{p^n} \sum_{x \in \mathbb{Z}_p^n} f(x)^2 = 1/p < \epsilon/4.$$

Thus, the first term contributes at most $\epsilon/2$ to the sum. For the next part, note that $\frac{C}{L^{1/4}+1} \leq \frac{\epsilon}{8}$. Then,

$$\begin{aligned}
\epsilon &\leq \epsilon/2 + pC \sum_{\beta \in \mathbb{Z}_p^n, |\beta| \leq L} \frac{|\hat{f}(\beta)|^2}{|\beta|^{1/4}+1} + p \sum_{\beta \in \mathbb{Z}_p^n, |\beta| > L} \frac{C|\hat{f}(\beta)|^2}{|\beta|^{1/4}+1} \\
&\leq \epsilon/2 + pC \sum_{\beta \in \mathbb{Z}_p^n, |\beta| \leq L} |\hat{f}(\beta)|^2 + p \sum_{\beta \in \mathbb{Z}_p^n, |\beta| > L} \frac{\epsilon}{8} |\hat{f}(\beta)|^2 \\
&\leq \epsilon/2 + pC \sum_{\beta \in \mathbb{Z}_p^n, |\beta| \leq L} |\hat{f}(\beta)|^2 + \epsilon/8.
\end{aligned}$$

Thus, we have that

$$\sum_{\beta \in \mathbb{Z}_p^n, |\beta| \leq L} |\hat{f}(\beta)|^2 \geq \frac{3\epsilon}{8pC} > \frac{\epsilon'}{2} \mathbb{E}[f].$$

By the contrapositive to Lemma 3.4, we have that $\max_{i \in [n]} \mathrm{Inf}_i^{\leq d} f \geq \delta$, as desired. □

Acknowledgments

We would like to thank an anonymous reviewer for pointing us to [LO43].

References

[BG17] Joshua Brakensiek and Venkatesan Guruswami. The quest for strong inapproximability results with perfect completeness. In *20th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX)*, pages 4:1–4:20, 2017.

[BJK05] Andrei A. Bulatov, Peter Jeavons, and Andrei A. Krokhin. Classifying the complexity of constraints using finite algebras. *SIAM J. Comput.*, 34(3):720–742, 2005.

[BKT17] Amey Bhangale, Subhash Khot, and Devanathan Thiruvenkatachari. An improved dictatorship test with perfect completeness. *Electronic Colloquium on Computational Complexity (ECCC)*, 24:30, 2017.

[Bul17] Andrei A. Bulatov. A dichotomy theorem for nonuniform CSPs. *arXiv:1703.03021 [cs]*, March 2017. arXiv: 1703.03021.

[Che09] Hubie Chen. A rendezvous of logic, complexity, and algebra. *ACM Comput. Surv.*, 42(1):2:1–2:32, December 2009.

[CMM06] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Near-optimal algorithms for unique games. In *Proceedings of the Thirty-eighth Annual ACM Symposium on Theory of Computing*, STOC '06, pages 205–214, New York, NY, USA, 2006. ACM.

[DKK+16] Irit Dinur, Subhash Khot, Guy Kindler, Dor Minzer, and Muli Safra. Towards a Proof of the 2-to-1 Games Conjecture? *Electronic Colloquium on Computational Complexity (ECCC)*, 23:198, 2016.

[DMR09] Irit Dinur, Elchanan Mossel, and Oded Regev. Conditional hardness for approximate coloring. *SIAM Journal on Computing*, 39(3):843–873, 2009.

[ES81] B. Efron and C. Stein. The Jackknife Estimate of Variance. *The Annals of Statistics*, 9(3):586–596, 1981.

[GKO+10] Venkatesan Guruswami, Subhash Khot, Ryan O'Donnell, Preyas Popat, Madhur Tulsiani, and Yi Wu. SDP gaps for 2-to-1 and other label-cover variants. In *Proceedings of the 37th International Colloquium on Automata, Languages and Programming*, pages 617–628, 2010.

[GS13] Venkatesan Guruswami and Ali Kemal Sinop. Improved inapproximability results for maximum k-colorable subgraph. *Theory of Computing*, 9:413–435, 2013.

[hs11] Mike Spivey (http://math.stackexchange.com/users/2370/mike spivey). Proof that the largest eigenvalue of a stochastic matrix is 1. Mathematics Stack Exchange, 2011. URL:http://math.stackexchange.com/q/40396 (version: 2011-05-20).

[Hå14] Johan Håstad. On the np-hardness of max-not-2. *SIAM Journal on Computing*, 43(1):179–193, 2014.

[Kho02] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings on 34th Annual ACM Symposium on Theory of Computing*, pages 767–775, 2002.

[KKMO07] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for max-cut and other 2-variable csps? *SIAM Journal on Computing*, 37(1):319–357, 2007.

[KMS17] Subhash Khot, Dor Minzer, and Muli Safra. On independent sets, 2-to-2 games, and grassmann graphs. In *Proceedings of the 49th Annual ACM Symposium on Theory of Computing*, pages 576–589, 2017.

[KS09] Subhash Khot and Rishi Saket. SDP integrality gaps with local ℓ_1-embeddability. In *Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 565–574, 2009.

[KS14] Subhash Khot and Rishi Saket. Hardness of finding independent sets in 2-colorable and almost 2-colorable hypergraphs. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1607–1625, 2014.

[KV15] Subhash Khot and Nisheeth K. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative-type metrics into ℓ_1. *J. ACM*, 62(1):8:1–8:39, 2015.

[LO43] J. E. Littlewood and A. C. Offord. On the number of real roots of a random algebraic equation. III. *Rec. Math. [Mat. Sbornik] N.S.*, 12(54):277–286, 1943.

[MOO10] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. *Ann. of Math. (2)*, 171(1):295–341, 2010.

[O'D14] Ryan O'Donnell. *Analysis of Boolean functions*. Cambridge University Press, 2014.

[OW09a] Ryan O'Donnell and Yi Wu. 3-bit dictator testing: 1 vs. 5/8. In *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '09, pages 365–373, Philadelphia, PA, USA, 2009. Society for Industrial and Applied Mathematics.

[OW09b] Ryan O'Donnell and Yi Wu. Conditional hardness for satisfiable 3-CSPs. In *Proceedings of the Forty-first Annual ACM Symposium on Theory of Computing*, STOC '09, pages 493–502, New York, NY, USA, 2009. ACM.

[Rag08] Prasad Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, pages 245–254, 2008.

[RS09] Prasad Raghavendra and David Steurer. Integrality gaps for strong SDP relaxations of UNIQUE GAMES. In *Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 575–585, 2009.

[TY10] Suguru Tamaki and Yuichi Yoshida. A query efficient non-adaptive long code test with perfect completeness. In Maria Serna, Ronen Shaltiel, Klaus Jansen, and José Rolim, editors, *Proceedings of the 14th International Workshop on Randomization and Approximation Techniques in Computer Science*, pages 738–751, 2010.

[Zhu17] Dmitriy Zhuk. The Proof of CSP Dichotomy Conjecture. *arXiv:1704.01914 [cs]*, April 2017. arXiv: 1704.01914.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il