Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 122 (2013)

PCPs via the low-degree long code and hardness for constrained hypergraph coloring*

Irit Dinur† Venkatesan Guruswami‡

Abstract

We develop new techniques to incorporate the recently proposed "short code" (a low-degree version of the long code) into the construction and analysis of PCPs in the classical "LABEL-COVER + Fourier Analysis" framework. As a result, we obtain more size-efficient PCPs that yield improved hardness results for approximating CSPs and certain coloring-type problems.

In particular, we show a hardness for a variant of hypergraph coloring (with hyperedges of size 6), with a gap between 2 and $\exp(2^{\Omega(\sqrt{\log\log N})})$ number of colors where N is the number of vertices. This is the first hardness result to go beyond the $O(\log N)$ barrier for a coloring-type problem. Our hardness bound is a doubly exponential improvement over the previously known $O(\log\log N)$-coloring hardness for 2-colorable hypergraphs, and an exponential improvement over the $(\log N)^{\Omega(1)}$-coloring hardness for $O(1)$-colorable hypergraphs. Stated in terms of "covering complexity," we show that for 6-ary Boolean CSPs, it is hard to decide if a given instance is perfectly satisfiable or if it requires more than $2^{\Omega(\sqrt{\log\log N})}$ assignments for covering all of the constraints.

While our methods do not yield a result for conventional hypergraph coloring due to some technical reasons, we also prove hardness of $(\log N)^{\Omega(1)}$-coloring 2-colorable 8-uniform hypergraphs (this result relies just on the long code).

A key algebraic result driving our analysis concerns a very low-soundness error testing method for Reed-Muller codes. We prove that if a function $\beta : \mathbb{F}_2^m \to \mathbb{F}_2$ is $2^{\Omega(d)}$ far in absolute distance from polynomials of degree $m-d$, then the probability that $\deg(\beta g) \leqslant m - 3d/4$ for a random degree $d/4$ polynomial g is *doubly exponentially* small in d.

*An extended abstract of this work was presented at the *54th Annual Symposium on Foundations of Computer Science (FOCS)*, October 2013 [8].

†Department of Applied Math and Computer Science, The Weizmann Institute of Science, Rehovot, Israel. Email: ⟨irit.dinur@weizmann.ac.il⟩. Research supported by US-Israel BSF grant number 2008293 and ERC grant number 239985.

‡Computer Science Department, Carnegie Mellon University, Pittsburgh, USA. Email: ⟨guruswami@cmu.edu⟩. Research supported in part by US-Israel BSF grant number 2008293 and the US National Science Foundation under Grant No. CCF-1115525.

ISSN 1433-8092

Contents

1 Introduction

Hardness of approximating constraint satisfaction problems is an area that has seen a great deal of progress in recent years. Following the pioneering works [3, 13], the standard framework for proving inapproximability has been via a combination of LABEL-COVER (or special cases such as Unique Games [18]) and the long code. For proving constant gap inapproximability, the relative inefficiency of the long code is insignificant. However, it becomes a serious bottleneck for non-constant parameter settings, most obviously, for proving hardness of approximate coloring. For this set of problems, there is an exponential or doubly exponential gap between the best known approximation algorithms (which require $n^{\Omega(1)}$ colors for n-vertex (hyper)graphs) and the best known hardness results (which at best only rule out efficient $o(\log n)$-coloring)

A very intriguing object called the "short code" was introduced and studied in [2]. This is a puncturing of the long code to locations indexed by low-degree polynomials, and to better reflect this, in this work we refer to the short code as the *low-degree long code*. This code was introduced in [2] as a "derandomization" of the long code, where it was used it to establish exponentially stronger integrality gaps for Unique Games, construct small set expanders whose Laplacians have many small eigenvalues, and obtain a more efficient version of the KKMO alphabet reduction [19] for Unique Games. The short code was used in conjunction with a pseudorandom generator for Lipschitz functions of polynomials to show an integrality gap of $\exp(\Omega(\sqrt{\log \log n}))$ for the Goemans-Linial semidefinite program for Uniform Sparsest Cut [15].

In this work we develop new techniques to use the low-degree long code in reductions from LABEL-COVER and obtain the following (quasi-)NP-hardness results. Our main results are

- A hardness for a *variant* of approximate hypergraph coloring, with a gap between 2 and $\exp(2^{\Omega(\sqrt{\log \log N})})$ number of colors (where N is the number of vertices). This is the *first* inapproximability result to go beyond the logarithmic barrier for a coloring-type problem.
- A hardness for gap$(1, \frac{15}{16} + \varepsilon)$-4SAT for $\varepsilon = \exp(-2^{\Omega(\sqrt{\log \log N})})$. This improves upon Håstad's result [13] where $\varepsilon = 1/(\log N)^c$ for some constant $c > 0$.
- A hardness for approximate hypergraph coloring, with a gap between 2 and $(\log N)^{\Omega(1)}$ colors.

Adapting a long-code test into the low-degree long code setting turns out to be non-trivial, and there seems to be no general recipe (as of yet) for doing so. For instance, while it is straightforward to import Håstad's classic gap$(1 - \varepsilon, 1/2 + \delta)$-3LIN result to the low-degree long code setting (we discuss this in Appendix A as a "warm-up"), the above results require a more carefully tailor-made construction. For certain PCPs in Håstad's work, such as 3SAT and 4-set splitting, we do not yet know how to adapt them to work with the low-degree long code. We comment that invariance-principle based analysis [23] is very powerful for analyzing dictatorship tests, and was used by [2] for analyzing their constructions. Nevertheless, for obtaining strong parameters we find that working directly with the Fourier expressions gives us a better handle on the kind of noise analysis that is needed.

For proving these results, we develop a "folding" mechanism for the low-degree long code that works with available LABEL-COVER constraints. The folding ensures that non-zero *low-weight* Fourier coefficients are supported only on assignments satisfying the associated constraints, which enables decoding a valid assignment from any such Fourier coefficient. One of the important components of any long-code test is the noise, which becomes especially subtle when aiming for perfect completeness. The degree restriction in the low-degree long code makes it harder to control the correlations between various functions via appropriately chosen noise. Finally, to analyze some of the noise expressions in our tests, and especially to be able to get stronger parameters, we prove some new results on local testing Reed Muller codes, which we discuss next.

1.1 Local testing of Reed Muller codes over $\mathbb{F}_2$

One of the key insights in [2] was a connection between the analysis of the low-degree long code and Reed-Muller testing. Let us denote by $P(m,r)$ the functions $\mathbb{F}_2^m \to \mathbb{F}_2$ that have degree $\leqslant r$. For functions $\beta, g : \mathbb{F}_2^m \to \mathbb{F}_2$, denote $\chi_\beta(g) = (-1)^{\sum_{x\in\mathbb{F}_2^m} \beta(x)g(x)}$. Specifically, given a β that is far from $P(m, m-d-1)$ polynomials, they noted that one can bound the expectation $|\mathbb{E}_\mu[\chi_\beta(\mu)]|$ for a random *low-weight* μ by appealing to a powerful result of [5] about testing Reed-Muller codes. This is formally stated in Proposition 16. Using such a noise μ enables attenuating the contribution of large weight Fourier coefficients; however, it causes the test to have imperfect completeness. To obtain our low-degree long code based constructions with perfect completeness, we prove a new result concerning testing Reed-Muller codes, stated below.

Theorem 1. *Let d be a multiple of 4. Let $\beta : \mathbb{F}_2^m \to \mathbb{F}_2$ be $2^{d/2}$-far from $P(m, m-d-1)$. Then for uniformly random polynomials $g \in P(m, d/4)$ and $h \in P(m, 3d/4)$, we have*

$$\mathop{\mathbb{E}}_{g}\Big[\ \big|\mathop{\mathbb{E}}_{h}[\chi_\beta(gh)]\big|\Big] \leqslant 2^{-4\cdot 2^{d/4}}\ .$$

The key quantitative aspect of the above result is the *doubly exponential* decay in d. To obtain such a bound, we observe that the set of "bad" choices of g, for which βg has degree $m - 3d/4 - 1$ (i.e., one lower than what one expects), is a *subspace* of $P(m, d/4)$. We then lower bound the co-dimension of this subspace by $2^{\Omega(d)}$. We do this via a recursive approach to pass to *two* similar problems in one less variable dimension $(m-1)$, by making use of the main technical ingredient in [5] which argues the abundance of hyperplanes A such that $\beta_{|A}$ is $2^{d/2-2}$-far from polynomials of degree $m-d-1$ in one less variable.

We note that a "robust" version of the above theorem, which argues that βg is also likely *far* from $P(m, m-3d/4-1)$, would be nice to have (as an interesting algebraic statement in itself). One can deduce such a claim from the above-mentioned result of [5] which proves such a robust version for $g \in P(m,1)$, but this only yields an upper bound of $2^{-O(d)}$ on the desired probability.

1.2 Inapproximability Results

To describe our results let us first briefly recall the notion of covering CSPs from [9]. A q-ary φ-CSP is given by a q-uniform hypergraph where each hyperedge is associated with a constraint φ. The *covering number* of a CSP is the minimal number of assignments to the vertices so that each hyperedge is covered by at least one assignment, see also Definition 1. If one views a hypergraph coloring instance as a not-all-equal CSP, then the covering number is exactly the ceiling of the base-2 logarithm of the chromatic number. This was the motivation of [12] and later [9] for studying the notion of covering numbers of CSPs.

In light of the lack of progress on hardness of approximate coloring for both graphs and hypergraphs, [9] suggested studying the hardness of gap covering problem, in the hope of approaching a potentially optimal gap-covering hardness result of 1 vs. $\Omega(\log N)$, which corresponds to a hardness gap of $O(1)$ vs. a polynomial number of colors. Given the current state of the art, they mentioned that even obtaining a gap of $O(1)$ vs. $\omega(\log\log N)$ would be interesting.

Theorem 2. *Given a 6-ary CSP of size N, no polynomial-time algorithm can decide if it is perfectly satisfiable, or if its covering number is at least $2^{\Omega(\sqrt{\log\log N})}$, unless* NP $\subseteq DTIME(n^{2^{O(\sqrt{\log\log n})}})$ *(note that this is contained in $DTIME(n^{o(\log n)})$).*

Prior to this work the best known gap-covering hardness was $O(1)$ vs. $O(\log\log N)$ (implicit in [16]) and 1 vs. $O(\log\log\log N)$ (implicit in [12]). Both these results in [12, 16] in fact applied to coloring (4-uniform) hypergraphs. It remains to be seen if a result similar to Theorem 2 can be obtained for hypergraph coloring. This would be a major quantitative jump, breaking the barrier of $O(\log N)$ colors.

We remark that the result above is obtained for a 6-ary constraint that is the disjunction of three inequality constraints. Since inequality makes sense over any alphabet size, one can think of this problem directly as a coloring-type problem, instead of a covering problem. This is always possible when the constraints are so-called "equality constrained languages" [6], and we give an alternative formulation of this theorem as a coloring problem in Theorem 28.

Along the way to proving Theorem 2, we establish the following inapproximability result for 4SAT with perfect completeness. We present this result first to illustrate our techniques in the basic setting of 4SAT, before applying them to a covering 6-CSP to deduce Theorem 2.

Theorem 3. *Given an instance of 4SAT of size N, then, assuming that $NP \not\subseteq DTIME(n^{O(\log n)})$, there is no polynomial time algorithm to distinguish between the following two cases:*

- *The instance is satisfiable.*
- *Every assignment satisfies at most a fraction $\frac{15}{16} + 2^{-2^{\Omega(\sqrt{\log\log N})}}$ of the clauses.*

We remark a similar result but without the perfect completeness would have been significantly easier to prove. We show in Appendix A a proof for the quasi-NP-hardness of 3LIN with similar sub-constant parameters that is a direct adaptation of Håstad's 3LIN proof. A direct adaptation of the perfect completeness tests seems less forthcoming due to the limitation on the noise imposed by working with the short code. It is worth mentioning that even for long code based constructions, perfect completeness tends to be significantly more difficult to ensure, often requiring additional technical elements, such as smoothness of LABEL-COVER projections [17], and/or picking functions whose bias itself is sampled from carefully chosen distributions as in [13, Sections 6,7], [14].

Fortunately, for 4SAT one can establish hardness avoiding the more complicated technical elements [13, Thm. 6.2] (this would yield an inapproximability factor $\frac{15}{16} + \frac{1}{(\log N)^c}$ for some small absolute constant $c > 0$). Even so, adapting this to the low-degree long code setting involves some careful design choices, as multiplying two functions, which seems like an essential component when perfect completeness is desired, increases the degree. This necessitates restricting certain functions in the test to be of smaller degree. In order to ensure that this does not bias the query pattern to a small portion of the low-degree long code, we query the smaller degree functions in a *separate* low-degree long code of smaller degree. This "multipartite" structural restriction is what precludes us from extending our result for covering 6-CSP (Theorem 2) to a result about hypergraph coloring. (Clearly, if the variables of every constraint straddles two or more parts, then the associated hypergraph is trivially 2-colorable.)

Finally, we also include a result on the hardness of hypergraph coloring. This result does not rely on the low-degree long code and is just based on techniques in Håstad's 1997 paper [13]. However, as the result statement is not explicit in the literature, we include it here along with a proof in Appendix B. (Also, this test indirectly paved the way for the result with the low-degree long code stated in Theorem 2.)

Theorem 4. *There is an absolute constant $c > 0$ such that the following holds. Given a 8-uniform hypergraph on N vertices, then, unless $NP \subseteq DTIME(n^{O(\log\log n)})$, there is no polynomial time algorithm to distinguish between the following two cases:*

- *The hypergraph can be colored with* 2 *colors so that every hyperedge is bichromatic.*
- *The hypergraph does not have an independent set with* $N/(\log N)^c$ *vertices, and in particular any coloring of the vertices with* $(\log N)^c$ *colors leads to a monochromatic hyperedge.*

Khot obtained a similar result using the "split code" for coloring 7-colorable 4-uniform hypergraphs [16]. The above statement is incomparable as it applies to 2-colorable hypergraphs, albeit of larger uniformity. For 3-uniform hypergraphs, hardness of $O(\sqrt[3]{\log\log N})$-coloring 2-colorable hypergraphs is shown in [10], and quasi-NP-hardness of $(\log\log N)^{1/9}$-coloring for the 3-colorable case is shown in [17]. A recent result [21] shows that for *almost* 2-colorable 4-uniform hypergraphs, where the hypergraph becomes 2-colorable upon removal of an ε fraction of vertices (and all incident hyperedges), it is quasi-NP-hard to find an independent set of size $N/2^{(\log N)^{1-\gamma}}$, for arbitrary constants $\varepsilon, \gamma > 0$.

We note that $(\log N)^{\Omega(1)}$ colors (achieved by Theorem 4 above) was the strongest quantitative bound on hardness for hypergraph coloring at the time of publication of conference version of this paper [8] (that version incorrectly claimed hardness for 6-uniform case, which has been weakened to 8-uniform case here). Following our work there was fairly rapid progress on hypergraph coloring, as briefly described below.

Subsequent work on hypergraph coloring. Using the low-degree long code, hardness of hypergraph coloring with $\exp(2^{\sqrt{\log\log N}})$ colors was shown in [11] for the case of 2-colorable 8-uniform hypergraphs (and also 4-colorable 4-uniform hypergraphs). Around the same time, Saket [24] proved the hardness of $(\log N)^{\Omega(1)}$ coloring 2-colorable 4-uniform hypergraphs using the original long code, thus improving our Theorem 4. Khot and Saket [20], using a new outer verifier based in part on our Theorem 1 on Reed-Muller testing, and the degree-2 long code at the inner level, have recently proved the hardness of $2^{(\log N)^{\Omega(1)}}$-coloring 2-colorable 12-uniform hypergraphs.

1.3 Organization

We begin in Section 2 with background information on LABEL-COVER and CSPs, the low-degree long code and its connection to Reed-Muller testing, and describe our folding mechanism for the low-degree long code. Our new algebraic result on testing Reed-Muller codes (Theorem 1) is proved in Section 3. In Section 4, we prove Theorem 3 on the hardness of approximating satisfiable instances of 4SAT. We prove the result for covering 6-CSP (Theorem 2) in Section 5. We present the extension of Håstad's 3LIN result to the low-degree long code setting in Appendix A. Finally, Theorem 4 on hardness of hypergraph coloring, which does not rely on the low-degree long code, is proved in Appendix B.

2 Preliminaries

2.1 LABEL-COVER and its hardness

A LABEL-COVER instance is given by a bipartite graph $G = (U, V, E)$, two alphabets Σ_U and Σ_V and a projection constraint $\pi_{uv} : \Sigma_U \to \Sigma_V$ per edge $uv \in E$. The goal is to assign labels to the vertices in a way that maximizes the number of satisfied constraints.

We next state a theorem about the NP-hardness of LABEL-COVER, where the LABEL-COVER has a concrete structure that is convenient for use with the low-degree long code.

Theorem 5 (Hardness of LABEL-COVER)**.** *Let $\ell \in \mathbb{N}$ be a parameter. There is a polynomial-time reduction from a 3SAT instance of size n to a* LABEL-COVER *instance of size $n^{O(\ell)}$ that is specified by*

- *A constraint graph $G = (U, V, E)$, $\Sigma_U = \mathbb{F}_2^{3\ell}$ and $\Sigma_V = \mathbb{F}_2^{\ell}$.*
- *Every $u \in U$ carries ℓ functions $f_1^{(u)}, \ldots, f_\ell^{(u)} : \mathbb{F}_2^3 \to \mathbb{F}_2$.*
- *Every edge $uv \in E$ carries a projection mapping defined by a subset $\pi_{uv} \subset [3\ell]$, $|\pi_{uv}| = \ell$, that contains exactly one element in each triple of indices $(3i+1, 3i+2, 3i+3)$, for $i = 0, \ldots, \ell-1$. The constraint on an edge is said to be satisfied by $a \in (\mathbb{F}_2^3)^\ell$ and $b \in \mathbb{F}_2^\ell$ if*

$$f_1^{(u)}(a_1) = \ldots = f_\ell^{(u)}(a_\ell) = 0 \qquad \textit{and} \qquad \pi_{uv}(a) = b\,.$$

The LABEL-COVER *instance has the following completeness and soundness conditions:*

- *If the 3SAT instance is satisfiable, then there is an assignment for the* LABEL-COVER *instance satisfying every constraint.*
- *If the 3SAT instance is unsatisfiable, then every assignment for the* LABEL-COVER *instance satisfies at most a $2^{-\Omega(\ell)}$ fraction of the constraints.*

This theorem is obtained from standard techniques: start with an NP-hard instance of gap-3SAT, and then perform ℓ-parallel repetition [1]. The functions $f_1^{(u)}, \ldots, f_\ell^{(u)}$ associated with an ℓ-tuple u of clauses check that the clauses are satisfied.

2.2 CSPs, Covering CSPs, and coloring problems

Let $X = \{x_1, ..., x_n\}$ be a set of n boolean variables and $\varphi : \{0,1\}^q \to \{0,1\}$ be a predicate. A *φ-constraint* over X is an equation of the form $\varphi(x_{i_1}, \ldots, x_{i_t}) = 1$, for some $i_1, \ldots, i_t \in [n]$, where $[n]$ denotes $\{1, 2, \ldots, n\}$. A *φ-CSP instance* C is a set of φ-constraints over X.

It is standard to denote by 4SAT the CSP where each constraint is defined by a disjunction of four variables or their negations, and by 3LIN the CSP where each constraint is defined by a linear equation over three variables modulo 2.

Let $A_1, \ldots, A_k \in \{0,1\}^n$ be a set of assignments for X. We say that $A_1, \ldots, A_k$ *cover* the instance C if for every constraint in C, there exists $i \in [k]$ such that A_i satisfies the constraint. The *covering number* of C, denoted $\nu(C)$, is smallest number k of assignments for X such that each constraint is satisfied by at least one of the assignments. We denote by cover-φ the problem of finding the covering number of a given CSP. The gap problem is defined as follows

Definition 1 (gap-cover-φ)**.** *Let $c < s \in \mathbb{N}$, and let φ be a predicate. Given a φ-CSP instance C, decide between*

- ***Yes case:*** *$\nu(C) \leq c$. I.e., there exists a set of at most c assignments that covers C.*
- ***No case:*** *$\nu(C) \geq s$. I.e., no set of at most s assignments covers C.*

2.3 The low-degree long code

Notation. We denote the field with two elements by $\mathbb{F}_2$. For a positive integer m, we denote by $\mathcal{F}_m$ the $\mathbb{F}_2$-vector space of functions $\mathbb{F}_2^m \to \mathbb{F}_2$. We can equip $\mathcal{F}_m$ with the Hamming metric by defining for $g, h \in \mathcal{F}_m$, their distance $\Delta(g,h)$ to be the number of $x \in \mathbb{F}_2^m$ such that $g(x) \neq h(x)$. For a subset $A \subseteq \mathbb{F}_2^m$, we denote by $g_{|A}$ be the function g restricted to A. The distance between $g_{|A}$ and $h_{|A}$, $\Delta(g_{|A}, h_{|A})$, is the number of $\mathbf{x} \in A$ such that $g(\mathbf{x}) \neq h(\mathbf{x})$.

For $g \in \mathcal{F}_m$ and $\mathcal{H} \subseteq \mathcal{F}_m$, we define $\Delta(g, \mathcal{H}) = \min_{h \in \mathcal{H}} \Delta(g,h)$. We say g is Δ-far from a subset $\mathcal{H} \subseteq \mathcal{F}_m$ is $\Delta(g,\mathcal{H}) > \Delta$; otherwise we say g is Δ-close to $\mathcal{H}$.

Every function $f \in \mathcal{F}_m$ can be uniquely expressed as a multilinear polynomial over $\mathbb{F}_2$ of degree at most m. We are interested in those functions which have much lower degree.

Definition 2 (Reed-Muller code)**.** *We denote by $P(m,d)$ the space of all functions $f : \mathbb{F}_2^m \to \mathbb{F}_2$ that have degree at most d. The evaluations of the polynomials in $P(m,d)$ at all points in $\mathbb{F}_2^m$ gives the binary m-variate Reed-Muller code of degree d, usually denoted as $\mathsf{RM}(m,d)$.*

Note that $P(m,d)$ is a subspace of $\mathcal{F}_m$. It is well-known and easy to see that the dual subspace of $P(m,d)$, denoted $P(m,d)^\perp$, is the subspace $P(m, m-d-1)$ of $\mathcal{F}_m$ consisting of polynomials of degree less than $m-d$.

We now define the low-degree long code first introduced in [2], where it is called the "short code."

Definition 3. *Let $m \geqslant d$ be positive integers, and let $\mathbf{a} \in \mathbb{F}_2^m$. For integers m, d, the (m-variate degree-d) low-degree long code of $\mathbf{a}$, denoted $\mathsf{SC}_{m,d}(\mathbf{a})$, is a function from $P(m,d)$ to $\mathbb{F}_2$ defined by*

$$\mathsf{SC}_{m,d}(a)(g) = (-1)^{g(a)} \qquad \text{for } g \in P(m,d)\ .$$

When m, d are clear from context, we will refer to the low-degree long code as $\mathsf{SC}(a)$.

For $\beta : \mathbb{F}_2^m \to \mathbb{F}_2$, its *support*, denoted $\mathsf{supp}(\beta)$, is the set $\{\mathbf{x} \in \mathbb{F}_2^m \mid \beta(\mathbf{x}) = 1\}$. The *weight* of β, denoted $\mathsf{wt}(\beta)$, equals $|\mathsf{supp}(\beta)|$. Note that $\mathsf{wt}(\beta) = \Delta(\beta, \mathbf{0})$ is the distance of β from the zero polynomial.

Definition 4 (Character set)**.** *For positive integers $m \geqslant d$, we define by $\Lambda(m,d)$ the set of functions $\beta : \mathbb{F}_2^m \to \mathbb{F}_2$ which are the minimum weight functions (ties broken arbitrarily) in the cosets of $P(m, m-d-1)$ in $\mathcal{F}_m$.*[1]

By definition, for each $\beta \in \Lambda(m,d)$, the closest polynomial (in Hamming distance) of degree at most $m-d-1$ to β is the zero polynomial. The functions in $\Lambda(m,d)$ correspond to the "Voronoi cell" of the zero polynomial for the set of points $P(m, m-d-1)$, under the metric $\Delta(\cdot,\cdot)$.
For functions $\beta, g : \mathbb{F}_2^m \to \mathbb{F}_2$, we define the "character mapping" $\chi_\beta(g)$ by

$$\chi_\beta(g) = (-1)^{\sum_{\mathbf{x} \in \mathbb{F}_2^m} \beta(\mathbf{x}) g(\mathbf{x})}\ .$$

The following are easy consequences of $P(m,d)^\perp$ being equal to $P(m, m-d-1)$.

Fact 6. *Suppose $\beta_1, \beta_2 : \mathbb{F}_2^m \to \mathbb{F}_2$ satisfy $\beta_1 + \beta_2 \in P(m, m-d-1)$. Then for all $g \in P(m,d)$, $\chi_{\beta_1}(g) = \chi_{\beta_2}(g)$.*

[1] Since $P(m,d)^\perp = P(m, m-d-1)$, one has $|\Lambda(m,d)| = |\mathcal{F}_m| / |P(m, m-d-1)| = |P(m,d)| = 2^{\sum_{j=0}^{d} \binom{m}{j}}$.

Fact 7. *For $\beta : \mathbb{F}_2^m \to \mathbb{F}_2$, we have*

$$\mathop{\mathbb{E}}_{g}[\chi_\beta(g)] = \begin{cases} 1 & \text{if } \beta \in P(m, m-d-1) \\ 0 & \text{otherwise} \end{cases}$$

where the expectation is taken over a random $g \in P(m,d)$.

Fact 8. *For $\beta_1, \beta_2 \in \Lambda(m,d)$, we have*

$$\mathop{\mathbb{E}}_{g}[\chi_{\beta_1}(g)\chi_{\beta_2}(g)] = \begin{cases} 1 & \text{if } \beta_1 = \beta_2 \\ 0 & \text{otherwise} \end{cases}$$

where the expectation is taken over a random $g \in P(m,d)$.

By well-known facts from the character theory of finite abelian groups, we have:

Fact 9. *Every function $A : P(m,d) \to \mathbb{R}$ admits the "Fourier" expansion*

$$A(g) = \sum_{\beta \in \Lambda(m,d)} \widehat{A}(\beta)\chi_\beta(g) \,,$$

where the Fourier coefficients are given by the inversion formula

$$\widehat{A}(\beta) = \mathop{\mathbb{E}}_{g}\big[A(g)\chi_\beta(g)\big] \,,$$

with the expectation taken over a uniformly random $g \in P(m,d)$.

Finally, we consider two functions over different-dimension domains, $A : P(m,d) \to \{-1,1\}$ and $B : P(\ell,d) \to \{-1,1\}$ where $m > \ell$. Suppose we have a projection $\pi : \mathbb{F}_2^m \to \mathbb{F}_2^\ell$ defined by $\pi(x_1,\ldots,x_m) = (x_{i_1},\ldots,x_{i_\ell})$ for some indices $1 \leqslant i_1 < \cdots < i_\ell \leqslant m$. The projection π allows us to lift a polynomial $f \in P(\ell,d)$ to the larger domain without changing its degree, defining $f \circ \pi \in P(m,d)$ by $f \circ \pi(x) = f(\pi(x))$. Now, for $\beta : \mathbb{F}_2^\ell \to \mathbb{F}_2$,

$$\begin{aligned} \chi_\beta(f \circ \pi) &= (-1)^{\sum_{x \in \mathbb{F}_2^m} (f\circ\pi)(x)\cdot\beta(x)} = (-1)^{\sum_{y \in \mathbb{F}_2^\ell} f(y)\cdot\sum_{x\in\pi^{-1}(y)}\beta(x)} \\ &= (-1)^{\sum_{y\in\mathbb{F}_2^\ell} f(y)\cdot\pi_2(\beta)(y)} = \chi_{\pi_2(\beta)}(f) \end{aligned} \tag{1}$$

where we define $\pi_2(\beta) : \mathbb{F}_2^\ell \to \mathbb{F}_2$ by $\pi_2(\beta)(y) = \sum_{x\in\pi^{-1}(y)} \beta(x) \mod 2$.

Fact 10. *Let $\beta \in \Lambda(m,d)$ and let $\alpha \in \Lambda(\ell,d)$. Then*

$$\mathop{\mathbb{E}}_{f\in P(\ell,d)}\big[\chi_\beta(f\circ\pi)\chi_\alpha(f)\big] = \begin{cases} 1 & \text{if } \alpha = \pi_2(\beta) \\ 0 & \text{otherwise .} \end{cases}$$

2.4 Folding properties of low-degree long code

Folding over constraints. Let $p_1,\ldots,p_k \in P(m,3)$ be given. Let

$$I = \langle p_1,\ldots,p_k\rangle = \left\{ \sum_{i=1}^k p_i q_i \;\middle|\; q_i \in P(m,d-3) \right\},$$

clearly a linear space. We define $P(m,d)/I$ to be the collection of cosets of I in $P(m,d)$, and we denote by $p+I$ the coset of $p \in P(m,d)$.

Definition 5 (Folding). *A function $A : P(m,d) \to \mathbb{R}$ is folded over $I = \langle p_1, \ldots, p_k \rangle$ if*

$$\forall p, p' \in P(m,d), \qquad p - p' \in I \quad \Rightarrow \quad A(p) = A(p').$$

A is folded over $\{-1,1\}$ if $A(g) = -A(1+g)$ for all $g \in P(m,d)$.

Fact 11. *Let $\mathbf{a} \in \mathbb{F}_2^m$. If $A = \mathsf{SC}(\mathbf{a})$ and $p_i(\mathbf{a}) = 0$ for all $i \in [k]$, then A is folded over $\langle p_1, \ldots, p_k \rangle$ and over $\{-1,1\}$.*

We next show that a function folded over I cannot have weight on small Fourier coefficients that are non-zero on I.

Claim 12. *Let $\beta : \mathbb{F}_2^m \to \mathbb{F}_2$ have $\mathsf{wt}(\beta) < 2^{d-3}$, and suppose there is an element $x \in \mathbb{F}_2^m$ with $\beta(x) = 1$ for which there is some p_i such that $p_i(x) \neq 0$. Then if $A : P(m,d) \to \mathbb{R}$ is folded over I then*

$$\widehat{A}(\beta) = \mathop{\mathbb{E}}_{g}[\chi_\beta(g) A(g)] = 0 \, .$$

Proof. Let $X = \{x \in \mathbb{F}_2^m \mid \ \beta(x) = 1 \text{ and } \exists i, p_i(x) \neq 0\}$. Choose some $a \in X$ and let i be such that $p_i(a) = 1$. Let $p = qp_i \in I$ where q is a polynomial that vanishes on all points of X except a. As $|X| < 2^{d-3}$, we can pick such a q that has degree at most $d-3$. This is because the dual space of $P(m,r)$ in $\mathcal{F}_m$ is $P(m, m-r-1)$, and any nonzero polynomial in $P(m, m-r-1)$ is nonzero on at least 2^{r+1} points (these are standard coding-theoretic facts about the distance property of binary Reed-Muller codes). Therefore, there are no linear dependencies amongst the evaluations of a degree r polynomial at fewer than 2^{r+1} values. Taking $r = d-3$, we can interpolate a degree $d-3$ polynomial q to take on any desired values at a set of less than 2^{d-3} distinct points in $\mathbb{F}_2^m$.

Now pair each function $g \in P(m,d)$ with $g + p$. By folding, $A(g) = A(g+p)$, but $\chi_\beta(g+p) = \chi_\beta(g)\chi_\beta(p) = -\chi_\beta(g)$, so $\widehat{A}(\beta) = 0$. □

Folding over "true". Let us denote by $P'(m,d)$ the set obtained by choosing exactly one function out of each pair $g, 1+g \in P(m,d)$. Similarly, denote by $P'(m,d)/I$ the set obtained by choosing exactly one coset out of each pair $g+I, 1+g+I \in P(m,d)/I$.

Given a function $A' : P'(m,d) \to \{-1,1\}$ it can be naturally extended to $A : P(m,d) \to \{-1,1\}$ by setting $A(1+g) = -A'(g)$. A function $A : P(m,d) \to \{-1,1\}$ is said to be folded over $\{-1,1\}$ if $A(g) = -A(1+g)$ for all g. The following are useful easy facts about folded functions.

Fact 13. *Given a function $\tilde{A} : P'(m,d)/I \to \mathbb{R}$, there is a unique function $A : P(m,d) \to \mathbb{R}$ that is folded over $\{-1,1\}$ and folded over I and for all $g \in P(m,d)$, $\tilde{A}(g+I) = A(g)$.*

Fact 14. *If A is folded over $\{-1,1\}$ then for any $\beta \in \Lambda(m,d)$ with even $\mathsf{wt}(\beta)$, $\widehat{A}(\beta) = 0$. In particular, $\widehat{A}(0) = 0$.*

When there is a tie in the choice of representative $\beta \in \Lambda(m,d)$ (as per Definition 4), all minimum weight functions in the coset have the same weight, and hence the parity of the weight does not depend on the choice of representative.

2.5 Reduction from LABEL-COVER using the low-degree long code

All of our inapproximability results follow the same general framework [3, 13], and combine LABEL-COVER with the long code adapted to the low-degree variant in the following way. Start from a LABEL-COVER instance G as in Theorem 5. For each $v \in V$ place a block of variables corresponding to $P(\ell, d)$. For each $u \in U$, let $I^{(u)} = \langle f_1^{(u)}, \ldots, f_\ell^{(u)} \rangle$ where $f_1^{(u)}, \ldots, f_\ell^{(u)}$ are the degree-3 functions that are associated with u. For each u place a block of variables corresponding to $P(3\ell, d)/I^{(u)}$.

Note that an assignment to these variables is equivalent to a collection of functions

$$\forall u, v, \qquad A^{(v)} : P(\ell, d) \to \{-1, 1\} \qquad \text{and} \qquad B^{(u)} : P(3\ell, d) \to \{-1, 1\}$$

such that for each $u \in U$, $B^{(u)}$ is folded over $I^{(u)}$. Sometimes we also require the tables to be folded over $\{-1, 1\}$, in which case the block of variables (from which we extend $A^{(v)}$ to all of $P(\ell, d)$) are restricted to $P'(\ell, d)$, and similarly $B^{(u)}$ is extended to $P(3\ell, d)$ from $P'(3\ell, d)/I^{(u)}$.

Our reductions, as usual, are described by a PCP verifier that randomly queries the functions $A^{(v)}$ and $B^{(u)}$. If φ is the acceptance predicate of the PCP verifier, then together with the query pattern this describes a φ-CSP. To analyze the reduction, one writes Fourier expressions that describe the probability of acceptance. The following lemma is an adaptation to the low-degree long code of Håstad's technique for converting certain Fourier expressions into a LABEL-COVER strategy. One subtle point below is that we need $\mathsf{wt}(\beta)$ to be bounded to ensure that every element in the support of β is a valid assignment to u, i.e., one that satisfies $f_1^{(u)}, \ldots, f_\ell^{(u)}$.

Lemma 15. *If $K \leqslant 2^{d-3}$ and*

$$\mathop{\mathbb{E}}_{uv}\Bigg[\sum_{\substack{\beta : \mathsf{wt}(\beta) < K \\ \pi_2(\beta) \neq 0}} \widehat{A^{(v)}}(\pi_2(\beta))^2 \widehat{B^{(u)}}(\beta)^2 \Bigg] \geqslant \delta, \tag{2}$$

then there is an assignment for the LABEL-COVER *satisfying at least δ/K of the constraints.*

Proof. Define a randomized assignment as follows. For each $u \in U$ choose a random $\beta \in \Lambda(3\ell, d)$ with probability proportional to $\widehat{B^{(u)}}(\beta)^2$ and then assign u with a random element $b \in \beta^{-1}(1)$. Similarly, for each $v \in V$, choose a random $\alpha \in \Lambda(\ell, d)$ with probability proportional to $\widehat{A}(\alpha)^2$ and then assign u with a random element $a \in \alpha^{-1}(1)$. Since $\sum_\beta \widehat{B}(\beta)^2 \leqslant 1$, the probability of picking a certain β is at least $\widehat{B}(\beta)^2$, and similarly for α.

The left hand side of (2) lower bounds the probability that u was assigned through β, and v was assigned through $\alpha = \pi_2(\beta)$. If that happened, then for each choice of $a \in \alpha^{-1}(1)$ there is at least one matching $b \in \beta^{-1}(1)$, which is chosen with probability at least $1/K$. It remains to observe the key fact that b is a valid assignment for u because of Claim 12 and the fact that $\mathsf{wt}(\beta) < K \leqslant 2^{d-3}$. □

2.6 Local testing of Reed-Muller codes

From Fact 8 we have, for $\beta \in \Lambda(m, d)$,

$$\mathop{\mathbb{E}}_{g}[\chi_\beta(g)] = \begin{cases} 1 & \text{if } \beta = 0 \\ 0 & \text{if } \beta \in \Lambda(m, d) \setminus \{0\} \end{cases} .$$

when the expectation is taken over a random $g \in P(m,d)$. Thus, orthogonality (over $\mathbb{F}_2$) with a random degree d polynomial $g \in P(m,d)$ serves as a perfect test for whether $\beta \in \Lambda(m,d)$ is the zero polynomial or not (or equivalently, if $\beta \in \mathcal{F}_m$ belongs to $P(m, m-d-1)$ or not). The next result, which follows from [5], shows that when $\beta \in \Lambda(m,d)$ has large weight (or equivalently, if $\beta \in \mathcal{F}_m$ is far from $P(m, m-d-1)$), the above expectation is bounded away from 1 even when g is chosen *pseudorandomly*, corresponding to the minimum weight codewords of $\mathsf{RM}(m,d)$ (i.e., products of d linearly independent affine forms). Specifically, let $L(m,d) \subseteq P(m,d)$ be the subset of degree d polynomials which are the product of exactly d linearly independent affine forms. Then we have the following claim which we will use in our warm-up 3LIN PCP (but not for any other PCP construction).

Proposition 16. *There exists an absolute constant $\rho_0 < 1$ such that for all $\beta \in \Lambda(m,d)$,*

$$\mathop{\mathbb{E}}_{\mu \in L(m,d)} \Big[\chi_\beta(\mu) \Big] \leqslant \rho = \max \left\{ 1 - \frac{\mathsf{wt}(\beta)}{2^d}, \rho_0 \right\} . \tag{3}$$

Moreover, if we choose $\mu_1, \ldots, \mu_t$ independently at random from $L(m,d)$ then

$$\mathop{\mathbb{E}}_{\mu_1,\ldots,\mu_t \in L(m,d)} \Big[\chi_\beta(\mu_1 + \cdots + \mu_t) \Big] \leqslant \rho^t , \tag{4}$$

Proof. Consider the test for membership of β in $P(m, m-d-1)$ that proceeds by picking a random $\mu \in L(m,d)$ and checking that $\sum_{\mathbf{x} \in \mathbb{F}_2^m} \beta(\mathbf{x})\mu(\mathbf{x}) = 0$. Then $\mathbb{E}_{\mu \in L(m,d)}[\chi_\beta(\mu)] = 1 - 2\mathsf{Rej}(\beta)$ where $\mathsf{Rej}(\beta)$ is probability that the test rejects β. Theorem 1 in [5], applied for m variables and degree $m-d-1$, implies that $\mathsf{Rej}(\beta) \geqslant \min\{\frac{\mathsf{wt}(\beta)}{2^{d+1}}, \epsilon_1\}$ for some absolute constant $\epsilon_1 > 0$. The bound (3) follows by setting $\rho_0 = 1 - 2\epsilon_1$. The bound (4) follows by noting that $\mathbb{E}[\chi_\beta(\mu_1 + \cdots + \mu_t)] = \mathbb{E}[\chi_\beta(\mu_1)] \cdots \mathbb{E}[\chi_\beta(\mu_t)]$. □

3 A new low-error tester for Reed-Muller codes

In this section, our goal is to prove the following result, which will be used in the analysis of our low-degree long code based PCPs to show that the "high frequency" terms in the Fourier expansion make a negligible contribution.

Theorem 17. *Let d be a multiple of 4. Let $\beta \in \mathcal{F}_m$ be $2^{d/2}$-far from $P(m, m-d-1)$, and let $g \in P(m, d/4)$ and $h \in P(m, 3d/4)$ be uniformly random polynomials from their respective domains. Then*

$$\mathop{\mathbb{E}}_{g} \Big[\big| \mathop{\mathbb{E}}_{h} [\chi_\beta(gh)] \big| \Big]] \leqslant 2^{-4 \cdot 2^{d/4}} . \tag{5}$$

Remark 18. Our proof yields a similar statement when the degrees of g, h are picked to be θd and $(1-\theta)d$ for $\theta \in (0, 1/2]$: We will be able to effectively test that β is $4^{\theta d}$-far from $P(m, m-d-1)$, and the upper bound in (5) will be $\exp(-\Omega(2^{\theta d}))$. The specific choice of $\theta = 1/4$ is made with an eye towards our applications to the PCPs in Sections 4 and 5. We did not attempt to optimize the constants in the above statement, and expect that by a more careful argument the base 4 in the $4^{\theta d}$-farness assumption can be replaced with any absolute constant bigger than 2, with a corresponding degradation in the constant multiplying $2^{\theta d}$ in the exponent of the upper bound (5).

Lte us now turn to the proof of Theorem 17. Fix a $\beta \in \mathcal{F}_m$. Appealing to Fact 7 we know

$$\mathop{\mathbb{E}}_{h \in P(m,3d/4)}[\chi_\beta(gh)] = \mathop{\mathbb{E}}_{h \in P(m,3d/4)}[\chi_{\beta \cdot g}(h)] = \begin{cases} 1 & \text{if } \beta g \in P(m, m-3d/4-1) \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

Therefore, the expectation in (5) equals

$$\mathop{\mathbb{E}}_{g}\Big[\ \big|\mathop{\mathbb{E}}_{h}[\chi_\beta(gh)]\big|\Big] = \Pr_{g \in P(m,d/4)}[\beta g \in P(m, m-3d/4-1)]\ . \tag{7}$$

The following simple observation shows that estimating the above probability is really a linear-algebraic problem of bounding the dimension of a certain subspace. This is the subspace of polynomials g for which the degree of βg is strictly smaller than the sum of the degrees.

Observation 19. *Fix any $\beta : \mathbb{F}_2^m \to \mathbb{F}_2$. For an integer $k \leqslant d$, the set*

$$B^{(m)}_{d,k}(\beta) \stackrel{\text{def}}{=} \{g \in P(m,k) \mid \beta g \in P(m, m-d-1+k)\} \tag{8}$$

is a subspace of $P(m,k)$.

Combining the above with Equation (7), we see that the expectation in (5) is given by

$$\mathop{\mathbb{E}}_{g}\Big[\ \big|\mathop{\mathbb{E}}_{h}[\chi_\beta(gh)]\big|\Big] = 2^{\dim(B^{(m)}_{d,d/4}(\beta)) - \dim(P(m,d/4))}\ .$$

Theorem 17 now follows from the following result.

Theorem 20. *For all positive integers m, d, k satisfying $m > d$ and $4|d$, the following holds. If $\beta : \mathbb{F}_2^m \to \mathbb{F}_2$ has distance more than $2^{d/2}$ from $P(m, m-d-1)$, then the subspace $B^{(m)}_{d,d/4}(\beta)$ defined in* (8) *has co-dimension (as a subspace of $P(m, d/4)$) at least $2^{d/4-2}$.*

The rest of the section is devoted to proving Theorem 20. For positive integers d, k, let us define the function $\Phi_{d,k} : \mathbb{N} \to \mathbb{N}$ as follows. If $d < k$, then $\Phi_{d,k}$ is identically 0. Otherwise, for $d \geqslant k$,

$$\Phi_{d,k}(D) = \min_{\substack{m > d \\ \beta \in \mathcal{F}_m;\ \Delta(\beta, P(m,m-d-1)) \geqslant D}} \Big\{\ \dim(P(m,k)) - \dim(B^{(m)}_{d,k}(\beta))\ \Big\}\ , \tag{9}$$

where $B^{(m)}_{d,k}(\beta)$ is as defined in (8).

We note that Theorem 20 follows if we prove that

$$\Phi_{d,d/4}(2^{d/2}) \geqslant 2^{d/4-2}\ . \tag{10}$$

We begin with the following claim which gives us the base case showing a lower bound when the distance $D \geqslant 1$.

Claim 21. *For $d \geqslant k$, and $D \geqslant 1$, $\Phi_{d,k}(D) \geqslant 1$.*

Proof. The claim can be restated as follows: If $\beta \notin P(m, m-d-1)$, then $B^{(m)}_{d,k}(\beta)$ is a proper subspace of $P(m,k)$, or in other words there exists $\nu \in P(m,k)$ such that $\beta\nu \notin P(m, m-d-1+k)$. We now prove this fact. As the dual space of $P(m, m-d-1)$ in $\mathcal{F}_m$ is $P(m,d)$, when $\beta \notin P(m, m-d-1)$, there must exist $\xi \in P(m,d)$ such that $\sum_{x\in\mathbb{F}_2^m} \beta(x)\xi(x) = 1$, or equivalently $\beta\xi \notin P(m, m-1)$. We may assume that ξ is a monomial $\xi = x_{i_1}x_{i_2}\cdots x_{i_l}$ with $l \leqslant d$ as such monomials form a basis of $P(m,d)$. If $l \leqslant k$, then ξ itself serves as the witness ν such that $\beta\nu \notin P(m, m-d-1+k)$. Otherwise, we can take $\nu = x_{i_1}x_{i_2}\cdots x_{i_k}$ and $\beta\nu$ can't have degree at most $m-d-1+k$ as that would imply $\beta\xi$ has degree at most $m-d-1+l \leqslant m-1$, a contradiction. □

The following lemma will be used in the recursive step when proving Theorem 20. It is based on a similar statement proved in [5].

Lemma 22. *Let $m > d$ be integers, and let $40 < D < 2^d$. If $\beta : \mathbb{F}_2^m \to \mathbb{F}_2$, which we think of as a polynomial in variables $x_1, x_2, \ldots, x_m$, is D-far from $P(m, m-d-1)$, then there exists a nonzero linear form $L = L(x_1, \ldots, x_m) \in P(m,1)$ such that $\beta_{|L=0}$ and $\beta_{|L=1}$ are* both *$D/4$-far from polynomials of degree $m-d-1$.*

Proof. Lemma 10 in [5], applied with $f = \beta$ and degree $m-d-1$, implies that if there are $K > 2^{m-d}$ *affine* forms $A_1, \ldots, A_K \in P(m,1)$ such that $\beta_{|A_i=0}$ is D'-close to a degree $m-d-1$ polynomial for $D' < 2^{d-2}$, then

$$\Delta(\beta, P(m, m-d-1)) \leqslant 3D' + 9\cdot 2^m/K \tag{11}$$

(after we adjust for the fact that we use unnormalized distance rather than fractional Hamming distance). If for every nonzero linear form L, at least one of $\beta_{|L=0}$ or $\beta_{|L=1}$ is $D/4$-close to a degree $m-d-1$ polynomial, applying (11) with $D' = D/4$ and $K = 2^m - 1$ we get $\Delta(\beta, P(m, m-d-1)) \leqslant 3D/4 + 10 < D$. This contradicts the hypothesis that β is D-far from $P(m, m-d-1)$. □

Proof. (of Theorem 20) Our goal is to establish the lower bound (10) on $\Phi_{d,d/4}(D)$ for $D = 2^{d/2}$. Recall that we are assuming $4|d$, and Claim 21 implies (10) when $d \leqslant 8$, so we may assume $d \geqslant 12$. Let $\beta \in \mathcal{F}_m$ be a polynomial in $x_1, x_2, \ldots, x_m$ that is D-far from $P(m, m-d-1)$. We need to prove

$$\dim(B^{(m)}_{d,d/4}(\beta)) \leqslant P(m, d/4) - 2^{d/4-2}\,.$$

By Lemma 22, we may assume, after applying a linear transformation on the coordinates, that $\beta_{x_m=0}$ and $\beta_{x_m=1}$ are both $D/4$-far from $P(m-1, m-d-1)$. Let us write the polynomial β in the form $\beta = x_m a(x_1, \ldots, x_{m-1}) + b(x_1, \ldots, x_{m-1})$. In other words, $\beta_{x_m=0} = b$ and $\beta_{x_m=1} = a + b$ where a, b are polynomials in $x_1, \ldots, x_{m-1}$. We know

$$\Delta(b, P(m-1, m-d-1)) \geqslant D/4 \qquad \text{and} \qquad \Delta(a+b, P(m-1, m-d-1)) \geqslant D/4\,. \tag{12}$$

Define $r = m-d-1$. We need to understand when $\nu \in P(m,k)$ is such that $\beta\nu \in P(m, r+k)$. Let us write the polynomial $\nu \in P(m,k)$ as $\nu = x_m p + q$ where $p \in P(m-1, k-1)$ and $q \in P(m-1, k)$ are polynomials in $x_1, \ldots, x_{m-1}$ of degree at most $k-1$ and k respectively. We have the following claim.

Claim 23. *If $\nu \in B^{(m)}_{d,k}(\beta)$, then $q \in B^{(m-1)}_{d-1,k}(b)$, i.e.,*

$$qb \in P(m-1, r+k), \quad \textit{and} \quad p(a+b) \in qa + P(m-1, r+k-1)\,.$$

Proof. (of Claim) Indeed, $\beta\nu = qb + x_m((a+b)p + qa)$. The terms in qb, which is a polynomial in $x_1, \ldots, x_{m-1}$, cannot be canceled by any terms in $x_m((a+b)p + qb)$. So if $\beta\nu$ has degree at most $r+k$, qb must also have degree at most $r+k$. Also, if $\beta\nu$ has degree at most $r+k$, the polynomial $p(a+b) + qa$ must have degree at most $r+k-1$, which is the same thing as $p(a+b) \in qa + P(m-1, r+k-1)$. □

By the above claim, the choice of ν in the subspace $B^{(m)}_{d,k}(\beta)$ amounts to picking an arbitrary q in the subspace $B^{(m-1)}_{d-1,k}(b)$ of $P(m-1,k)$, and then p from a coset of the subspace

$$B^{(m-1)}_{d-1,k-1}(a+b) = \{\tilde{\nu} \in P(m-1,k-1) \mid (a+b)\tilde{\nu} \in P(m-1, r+k-1)\}$$

of $P(m-1,k-1)$. Therefore,

$$\dim(B^{(m)}_{d,k}(\beta)) \leqslant \dim(B^{(m-1)}_{d-1,k}(b)) + \dim(B^{(m-1)}_{d-1,k-1}(a+b)) \,. \tag{13}$$

Combining (9), (12), (13), and the equality

$$\dim(P(m,k)) = \dim(P(m-1,k)) + \dim(P(m-1,k-1)) \,,$$

we can conclude the following for all $d \geqslant k$ and $D < 2^d$:

$$\Phi_{d,k}(D) \geqslant \Phi_{d-1,k}(D/4) + \Phi_{d-1,k-1}(D/4) \,. \tag{14}$$

When $D = 2^{d/2} = 4^{d/4}$ and $k = d/4$, recursively applying the above for a depth of $d/4 - 2$ (to reduce D geometrically from $4^{d/4}$ to 16), and using Claim 21, we can lower bound $\Phi_{d,d/4}(2^{d/2}) \geqslant 2^{d/4-2}$, giving us (10), as desired. □

4 PCP checking 4SAT using the low-degree long code

In this section, our goal is to give a low-degree long code based PCP that has perfect completeness. The smallest number of queries for which we are able to do so is 4 queries. The predicate tested by the PCP is 4SAT (actually we can test a slightly stronger arity 4 predicate $x \vee y \vee (z \neq w)$). As a result we establish Theorem 3 on the inapproximability of 4SAT stated in the introduction. Our construction is inspired by Håstad's tight inapproximability result for satisfiable instances of 4SAT [13, Theorem 6.2]. The analysis here is more subtle due to the restriction of using the low-degree long code. Our main motivation here is to illustrate these techniques in the simple setting of 4SAT, before applying them to show hardness for covering CSP later on.

As explained in Section 2.5, we describe the PCP verifier as a randomized test that checks if a LABEL-COVER instance is satisfiable, or highly unsatisfiable, in the sense of Theorem 5. The verifier has access to tables $A^{(v)}$ and $B^{(u)}$ of purported low-degree long codes of the labels of the nodes $u \in U$ and $v \in V$ of the LABEL-COVER instance.

However, there are some key differences in the setting here. First, the table for the "smaller" side is a low-degree long code for smaller degree ($3d/4$ as opposed to d). Second, there are *two* tables for the nodes on the "larger" side, with one being a low-degree long code of smaller degree. This structure seems technically necessary as we need to restrict the degree of some of the functions to be smaller than d, and in this case the analysis necessitates making them from a separate low-degree long code so that they are well-distributed amongst the coordinates of that low-degree long code. Let us proceed with the formal description of the PCP construction.

Let $G = (U, V, E)$ be a LABEL-COVER instance with parameter ℓ as promised in Theorem 5. The integer d is a degree parameter to be chosen later (it will be a multiple of 4).

For each $v \in V$ we add a block of variables corresponding to $P'(\ell, 3d/4)$ (recall that $P'(\ell, 3d/4)$ contains for each $g \in P(\ell, 3d/4)$ exactly one of g and $1 + g$). For each $u \in U$, we add *two* blocks of variables, one corresponding to $P'(3\ell, d/4)$ and another corresponding to $P'(3\ell, d)/I^{(u)}$ (where $I^{(u)}$ denotes the ideal corresponding to node u described in Section 2.5).

Let us denote $m = 3\ell$. An assignment for the variables is described by a collection of functions $A^{(v)} : P'(\ell, 3d/4) \to \{-1, 1\}$ for each $v \in V$, and functions $C^{(u)} : P'(m, d/4) \to \{-1, 1\}$, $B^{(u)} : P'(m, d)/I^{(u)} \to \{-1, 1\}$ for each $u \in U$. We can extend the functions in the natural way to assume we have access to functions $A^{(v)} : P(\ell, 3d/4) \to \{-1, 1\}$, $C^{(u)} : P(m, d/4) \to \{-1, 1\}$ that are folded over $\{-1, 1\}$, and a function $B^{(u)} : P(m, d) \to \{-1, 1\}$ that is folded over $\{-1, 1\}$ and $I^{(u)}$.

We now describe our PCP, which we call 4SAT-PCP :

1. Choose a random edge (u, v) in the LABEL-COVER instance, and let $\pi_{uv} : \mathbb{F}_2^m \to \mathbb{F}_2^\ell$ be the associated projection.

   For notational simplicity, we denote $\pi = \pi_{uv}$, $A = A^{(v)}$, $B = B^{(u)}$ and $C = C^{(u)}$.
2. Sample functions $f \in P(\ell, 3d/4)$, $g \in P(m, d/4)$, $\tilde{g} \in P(m, d)$ and $h \in P(m, 3d/4)$, where each function is chosen independently at random from its respective domain.
3. Denote $g' = \tilde{g} + gh + (1 + g)(1 + f \circ \pi)$ and note that $g' \in P(m, d)$.

   Accept iff at least one of $A(f)$, $C(g)$, $B(\tilde{g})$, and $B(g')$ equals -1.

Before proceeding to the analysis of the above PCP, let us pause to make the following remark concerning the choice of degree parameters in the test.

Remark 24. In the above test, we pick the degree of g to be $d/4$ and the degree of h, f to be $3d/4$. Admittedly, it will be aesthetically nicer to pick them to have degrees $d/2$ each. We made this choice for a somewhat technical reason which we hint at now. The soundness analysis of the test is based on killing the contribution of the "high frequency" Fourier coefficients $\widehat{B}(\beta)$ corresponding to β that is Δ-far from $P(m, m - d - 1)$ for some proximity parameter Δ. Theorem 17 (see Remark 18) would only allow us to handle $\Delta \gg 2^{d/2}$ if we picked g to have degree $d/2$. So we have to handle β with $\mathsf{wt}(\beta) = 2^{d/2}$ in the "low-frequency" part of the analysis (eg. (19) in the ensuing soundness analysis). But as there are polynomials of degree $m - d/2 - 1$ with weight at least $2^{d/2}$, this weight bound is consistent with $\beta(1 + g)$ being a nonzero polynomial in $P(m, m - d/2 - 1)$ in which case its representative in $\Lambda(m, d/2)$ equals 0. This causes trouble in the rest of the analysis (steps (20) and beyond).

Remark 25. It turns out that one can make a simpler test where we pick $f \in P(\ell, d)$, and set $g' = \tilde{g} + gh + (1 + f \circ \pi)$. This was realized in the follow-up work [11]. The completeness of the test (Lemma 26 below) is clear in this case also. With this modification, we can in fact take g, h to have degree $d/2$ and the soundness analysis goes through, thus removing the degree asymmetry mentioned in Remark 24 above. The same holds for the test in Section 5.

4.1 Completeness

We first establish the perfect completeness of the test which also explains the logic behind the test.

Lemma 26. *If G is satisfiable, then there are tables $A^{(v)}$, $B^{(u)}$, and $C^{(u)}$ for which the test* 4SAT-PCP *accepts with probability* 1*. In particular, there are tables so that the four bits read by the verifier are never all equal to* 1*.*

Proof. Given a perfectly satisfying assignment for G, let us assign each $A^{(v)}$ to be $\mathsf{SC}_{\ell,3d/4}(a)$, the degree-$3d/4$ low-degree long code of a, where $a \in \mathbb{F}_2^\ell$ is the label for v. Similarly, define $B^{(u)} = \mathsf{SC}_{m,d}(b)$ and $C^{(u)} = \mathsf{SC}_{m,d/4}(b)$ where b is the label for u. For the choice of edge (u, v), the condition checked by the test amounts to

$$f(a) = 1 \ \vee \ g(b) = 1 \ \vee \ \tilde{g}(b) = 1 \ \vee \ g'(b) = 1 \, . \tag{15}$$

To prove (15) holds, let us assume $f(a) = g(b) = 0$ and then argue that in this case $\tilde{g}(b) \neq g'(b)$ (which in particular means one either $\tilde{g}(b)$ or $g'(b)$ equals 1). Indeed

$$\tilde{g}(b) + g'(b) = g(b)h(b) + (1 + g(b))(1 + f(a)) = 1$$

when $f(a) = g(b) = 0$. Note that we have shown the more stringent condition $A(f) = -1$ or $C(g) = -1$ or $B(\tilde{g}) \neq B(g')$ always holds in the completeness case. □

4.2 Soundness

We now turn to the soundness analysis. We prove that if the original LABEL-COVER instance G is highly unsatisfiable, then the test does not accept any proof with probability noticeably larger than $15/16$ (which is the probability with which completely random tables are accepted). The formal theorem follows.

Theorem 27. *If every assignment for G satisfies at most a fraction $2^{-\Omega(\ell)}$ of the edges, and $d = 4\lceil \log_2 \ell \rceil$, then the test* 4SAT-PCP *accepts with probability at most $\frac{15}{16} + 2^{-\Omega(\ell)}$.*

Proof. The probability that the test 4SAT-PCP accepts equals

$$1 - \mathop{\mathbb{E}}_{u,v,f,g,h,\tilde{g}} \left[\left(\frac{1 + A(f)}{2} \right) \left(\frac{1 + C(g)}{2} \right) \left(\frac{1 + B(\tilde{g})}{2} \right) \left(\frac{1 + B(g')}{2} \right) \right]$$

where we use the shorthand $\overline{g}$ to denote $1 + g$, $g' = \tilde{g} + gh + \overline{g}\ \overline{f \circ \pi}$, and as in the test A denotes $A^{(v)}$, $B = B^{(u)}$ and $C = C^{(u)}$.

Let us now fix a choice of edge $(u, v) \in E$ and focus on the inner expectation over just $f, g, h, \tilde{g}$. Expanding out the product, by the mutual independence of triples $(f, g, \tilde{g})$ and (f, g, g'), and the fact that A, B, C are all folded over $\{-1, 1\}$, all of the product terms which don't include both $B(\tilde{g})$ and $B(g')$ equal 0. The distribution of $(g, \tilde{g}, g')$ is identical to that of $(g, \tilde{g}, \overline{g'})$, as can be seen by replacing f, h by the identically distributed $\overline{f}, \overline{h}$. This together with $B(\overline{g'}) = -B(g')$ implies that $\mathbb{E}_{f,g,h,\tilde{g}}\big[C(g)B(\tilde{g})B(g')\big] = 0$. The distribution of $(f, \tilde{g}, g')$ can also be seen to be identical to that of $(\overline{f}, \overline{\tilde{g}}, \overline{g'})$, which implies

$$\mathop{\mathbb{E}}_{f,g,h,\tilde{g}}\big[A(f)B(\tilde{g})B(g')\big] = 0 \, .$$

After these simplifications, conditioned on picking $(u, v) \in E$, the probability $p_{(u,v)}$ that the test accepts is given by

$$p_{(u,v)} = \frac{15}{16} - \frac{1}{16} \underbrace{\mathop{\mathbb{E}}_{f,g,h,\tilde{g}}\Big[A(f)C(g)B(\tilde{g})B(\tilde{g} + gh + \overline{g}\ \overline{f \circ \pi})\Big]}_{\Theta_{(u,v)}} \, . \tag{16}$$

Writing the Fourier expansions of B as given by Fact 9, we can expand the inner expectation as

$$\Theta_{(u,v)} = \sum_{\beta_1,\beta_2} \widehat{B}(\beta_1)\widehat{B}(\beta_2) \mathop{\mathbb{E}}_{f,g,h}\Big[A(f)C(g)\chi_\gamma(g)\chi_{\beta_2}(gh + \overline{g}\,\overline{f\circ\pi})\Big] \mathop{\mathbb{E}}_{\tilde{g}}\Big[\chi_{\beta_1}(\tilde{g})\chi_{\beta_2}(\tilde{g})\Big] \tag{17}$$

summed over $\beta_1, \beta_2 \in \Lambda(m,d)$. The expectation over $\tilde{g}$ is 0 unless $\beta_1 = \beta_2$ by Fact 8, in which case it equals 1.

Simplifying (17) using this, we get

$$\Theta_{(u,v)} = \sum_{\beta\in\Lambda(m,d)} \widehat{B}(\beta)^2 \mathop{\mathbb{E}}_{g}\Big[C(g) \mathop{\mathbb{E}}_{h}\big[\chi_\beta(gh)\big] \mathop{\mathbb{E}}_{f}\big[A(f)\chi_\beta(\overline{g}\,\overline{f\circ\pi})\big]\Big] \tag{18}$$

For terms with $\mathsf{wt}(\beta) \geqslant 2^{d/2}$, we have the absolute value of the expectation over g in (18) is at most

$$\mathop{\mathbb{E}}_{g}\Big[|C(g)|\,\big|\mathop{\mathbb{E}}_{h}[\chi_\beta(gh)]\big|\,\big|\mathop{\mathbb{E}}_{f}[A(f)\chi_\alpha(f)\chi_\beta(\overline{g}\,\overline{f\circ\pi})]\big|\Big] \leqslant \mathop{\mathbb{E}}_{g}[\big|\mathop{\mathbb{E}}_{h}[\chi_\beta(gh)]\big|] \leqslant 2^{-2^{d/4}}$$

using Theorem 17. Since $\sum_\beta \widehat{B}(\beta)^2 \leqslant 1$, we can conclude

$$\Theta_{(u,v)} \geqslant \Bigg(\sum_{\beta:\mathsf{wt}(\beta)<2^{d/2}} \widehat{B}(\beta)^2 \mathop{\mathbb{E}}_{g}\Big[C(g) \mathop{\mathbb{E}}_{h}\big[\chi_\beta(gh)\big] \mathop{\mathbb{E}}_{f}\big[A(f)\chi_\beta(\overline{g}\,\overline{f\circ\pi})\big]\Big]\Bigg) - 2^{-2^{d/4}}\,. \tag{19}$$

When $\mathsf{wt}(\beta) < 2^{d/2}$, we have $\mathsf{wt}(\beta\overline{g}) < 2^{d/2}$ as well. This means the closest polynomial of degree $m - 3d/4 - 1$ to $\beta\overline{g}$ is 0, and so $\beta\overline{g} \in \Lambda(m, 3d/4)$. Writing the Fourier expansion of A as $A(f) = \sum_{\alpha\in\Lambda(\ell,3d/4)} \widehat{A}(\alpha)\chi_\alpha(f)$, we can simplify

$$\mathop{\mathbb{E}}_{f}\big[A(f)\chi_\beta(\overline{g}\,\overline{f\circ\pi})\big] = \sum_{\alpha} \widehat{A}(\alpha) \mathop{\mathbb{E}}_{f}\big[\chi_\alpha(f)\chi_{\beta\overline{g}}(1+f\circ\pi)\big] = \widehat{A}(\pi_2(\beta\overline{g}))(-1)^{\mathsf{wt}(\pi_2(\beta\overline{g}))} \tag{20}$$

using Fact 10.

Likewise $\chi_\beta(gh) = \chi_{\beta g}(h)$, and so $\mathbb{E}_h\big[\chi_\beta(gh)\big] = 1$ if $\beta g = 0$, and 0 otherwise. Putting together this fact and (20), the expectation over g in (19) equals

$$\begin{aligned}
&\mathop{\mathbb{E}}_{g}\Big[C(g)(-1)^{\mathsf{wt}(\pi_2(\beta\overline{g}))}\widehat{A}(\pi_2(\beta\overline{g}))\,\mathbf{1}(\beta g = 0)\Big] \\
&= \mathop{\mathbb{E}}_{g}\Big[C(g)(-1)^{\mathsf{wt}(\pi_2(\beta))}\widehat{A}(\pi_2(\beta))\,\mathbf{1}(\beta g = 0)\Big] \geqslant -|\widehat{A}(\pi_2(\beta))|
\end{aligned}$$

where we use $\mathbf{1}(E)$ for the indicator of an event E. Plugging this into (19), we get

$$\Theta_{(u,v)} \geqslant -\Bigg(\sum_{\beta:\mathsf{wt}(\beta)<2^{d/2}} |\widehat{A}(\pi_2(\beta))|\,\widehat{B}(\beta)^2\Bigg) - 2^{-2^{d/4}} \tag{21}$$

Since B is folded over $\{-1,1\}$, Fact 14 implies that $\widehat{B}(\beta) = 0$ when $\mathsf{wt}(\beta)$ is even. Combining (21) and (16), the probability that the test accepts is

$$\begin{aligned}
\mathop{\mathbb{E}}_{(u,v)\in E}\big[p_{(u,v)}\big] &\leqslant \frac{15}{16} + \frac{1}{16}\cdot 2^{-2^{d/4}} + \frac{1}{16}\cdot \mathop{\mathbb{E}}_{(u,v)}\Bigg[\sum_{\substack{\beta:\mathsf{wt}(\beta)<2^{d/2}\\ \mathsf{wt}(\beta)\text{ odd}}} |\widehat{A}(\pi_2(\beta))|\,\widehat{B}(\beta)^2\Bigg] \\
&\leqslant \frac{15}{16} + 2^{-2^{d/4}} + \sqrt{\mathop{\mathbb{E}}_{(u,v)}\Bigg[\sum_{\substack{\beta:\mathsf{wt}(\beta)<2^{d/2}\\ \mathsf{wt}(\beta)\text{ odd}}} \widehat{A}(\pi_2(\beta))^2\widehat{B}(\beta)^2\Bigg]}
\end{aligned} \tag{22}$$

where in the second step we used Cauchy-Schwarz inequality and $\sum_\beta \widehat{B}(\beta)^2 \leqslant 1$. As $\mathsf{wt}(\pi_2(\beta))$ and $\mathsf{wt}(\beta)$ have the same parity, when $\mathsf{wt}(\beta)$ is odd, $\pi_2(\beta) \neq 0$. Appealing to Lemma 15, the quantity inside the square root in (22), divided by 2^d, gives a lower bound on the optimum fraction of edges that can be satisfied in the LABEL-COVER instance G. As the latter is at most $2^{-\Omega(\ell)}$, we conclude that

$$\mathop{\mathbb{E}}_{(u,v)\in E}\left[p_{(u,v)}\right] \leqslant \frac{15}{16} + 2^{-2^{d/4}} + 2^{d-\Omega(\ell)} \; .$$

Therefore for $d = 4\lceil \log_2 \ell \rceil$, the test accepts with probability at most $\frac{15}{16} + 2^{-\Omega(\ell)}$. □

Picking $\ell = 2^{\lfloor \sqrt{\log\log n} \rfloor/4}$ and $d = \lfloor \sqrt{\log\log n} \rfloor$, the size of the 4SAT instance produced is at most polynomial in $N \leqslant n^{3\ell} 2^{(3\ell)^d} \leqslant n^{2^{O(\sqrt{\log\log n})}}$, and the reduction runs in $N^{O(1)}$ time. As a function of N, we have $\ell \geqslant 2^{\Omega(\sqrt{\log\log N})}$. Combining the completeness Lemma 26 and the soundness Theorem 27, we can conclude Theorem 3 showing a 1 vs. $\frac{15}{16} + 2^{-2^{\Omega(\sqrt{\log\log N})}}$ gap for 4SAT. In comparison, Håstad's result using the long code [13] can establish an inapproximability gap of 1 vs. $\frac{15}{16} + 1/(\log N)^c$ for some small absolute constant $c > 0$.

5 6-query covering PCP using low-degree long code

In this section, we prove Theorem 2, showing it is hard to decide if a given instance of a φ-CSP has covering number 1 or at least $k = 2^{O(\sqrt{\log\log n})}$, where the predicate φ is defined by

$$\varphi(a,b,c,d,e,f) = (a \neq b) \vee (c \neq d) \vee (e \neq f) \; .$$

Before moving to the proof, let us mention that since this predicate involves monotone Boolean operations over inequality constraints on the variables, it makes sense to assign variables with any number of colors (rather than Boolean values only). Given a φ-CSP instance over variables X, such that the variables occur without negations, we say it is c-colorable if there is a coloring of the variables $\psi : X \to \{1, 2, \ldots, c\}$ such that every constraint is satisfied. It is easy to generalize the connection in [12] to this case, showing that the logarithm of this version of the chromatic number is equal to its covering number. Thus, an equivalent statement of Theorem 2 is the following.

Theorem 28. *Given a φ-CSP instance with N vertices, then, assuming that $NP \not\subseteq DTIME(n^{2^{O(\sqrt{\log\log n})}})$, there is no polynomial time algorithm to distinguish between the following two cases:*

- *The instance can be colored with $c = 2$ colors.*
- *The instance cannot be colored even with $2^{2^{\Omega(\sqrt{\log\log N})}}$ colors.*

Every 4SAT instance is always trivially covered by any pair of satisfying assignment and its complement, so the covering number of 4SAT is always at most 2. So 4SAT-PCP from the previous section cannot give the desired coloring hardness. However, we now show that a small change to the test gives us the desired PCP with a total of 6 queries. Specifically, we will replace the condition $A(f) = -1$ with the check $A(f_1) \neq A(f_1 + f)$, and the condition $C(g) = -1$ with the check $C(g_1) \neq C(g_1 + g)$, for independent uniformly random functions f_1, g_1.

As in Section 4 we begin with a LABEL-COVER instance $G = (U, V, E)$, and place low-degree long code tables for the vertices of G. Namely, for each $v \in V$, a table $A^{(v)} : P(\ell, 3d/4) \to \{-1, 1\}$, and for each

$u \in U$, two tables $C^{(u)} : P(m, d/4) \to \{-1, 1\}$ and $B^{(u)} : P(m, d)/I^{(u)} \to \{-1, 1\}$ (where $m = 3\ell$). The fact that we are working with a constrained coloring predicate prevents us from folding the tables over $\{-1, 1\}$. Once again, we extend $B^{(u)}$ to all of $P(m, d)$ by defining $B(h) = B(h + I^{(u)})$, and assume that $B^{(u)} : P(m, d) \to \{-1, 1\}$ is folded over $I^{(u)}$.

We now describe our PCP, which we call 6-NE-PCP

1. Choose a random edge (u, v) in the LABEL-COVER instance, and let $\pi_{uv} : \mathbb{F}_2^m \to \mathbb{F}_2^\ell$ be the associated projection.

   For notational simplicity, we denote $\pi = \pi_{uv}$, $A = A^{(v)}$, $B = B^{(u)}$ and $C = C^{(u)}$.

2. Sample functions $f, f_1 \in P(\ell, 3d/4)$, $g, g_1 \in P(m, d/4)$, $\tilde{g} \in P(m, d)$ and $h \in P(m, 3d/4)$, where each function is chosen independently at random from its respective domain. Denote $g' = \tilde{g} + gh + (1 + g)(1 + f \circ \pi)$ and note that $g' \in P(m, d)$.

3. Accept iff
$$(A(f_1) \neq A(f_1 + f)) \ \vee \ (C(g_1) \neq C(g_1 + g)) \ \vee \ (B(\tilde{g}) \neq B(g')) \ . \tag{23}$$

Perfect Completeness. Given a perfectly satisfying assignment for G, let us assign each $A^{(v)}$ to be $\mathsf{SC}_{\ell,3d/4}(a)$, the degree-$3d/4$ low-degree long code of a, where $a \in \mathbb{F}_2^\ell$ is the label for v. Similarly, define $B^{(u)} = \mathsf{SC}_{m,d}(b)$ and $C^{(u)} = \mathsf{SC}_{m,d/4}(b)$ where b is the label for u. For the choice of edge (u, v), the condition checked by the test amounts to

$$(f_1(a) \neq f_1(a) + f(a)) \ \vee \ (g_1(b) \neq g_1(b) + g(b)) \ \vee \ (\tilde{g}(b) \neq g'(b)) \ .$$

This equation clearly holds if $f(a) = 1$ or $g(b) = 1$. Otherwise, if $f(a) = g(b) = 0$ we claim that $\tilde{g}(b) \neq g'(b)$. Indeed

$$\tilde{g}(b) + g'(b) = g(b)h(b) + (1 + g(b))(1 + f(a)) = 0 + 1 = 1$$

when $f(a) = g(b) = 0$.

Soundness analysis. As in Section 4, it can be proved that when the LABEL-COVER instance G is highly unsatisfiable, no choice of tables can make the 6-NE-PCP test accept with probability more than $7/8$ (again random tables are accepted with this probability, so this bound is tight). Given our interest in covering soundness we now show that even a large number of proofs cannot cover every test made by the verifier. The formal statement follows.

Theorem 29. *If every assignment of labels to the* LABEL-COVER *instance G satisfies at most a fraction $2^{-\Omega(\ell)}$ of the edges, and $d = 4\lceil \log \ell \rceil$, then there exists $k = \Omega(\ell)$ such that for every set of k tables there is some check* (23) *that is violated by all of them.*

Proof. Suppose there are k proofs such that every check (23) accepts at least one of them. Let $\rho = 1/2^k$. Then, viewing these k proofs as a 2^k-coloring, we can choose a subset consisting of exactly a fraction ρ of the locations of each of the $A^{(v)}$-tables, and similarly for the $B^{(u)}$ and $C^{(u)}$-tables, such that no check (23) has all 6 queries amongst the chosen locations. (To see this simply take the most popular color class in each of the tables, and in each table choose arbitrarily a ρ-sized subset of this color class). To express this analytically, let $F^{(v)} : P(\ell, 3d/4) \to \{0, 1\}$ be the indicator function of this subset restricted to $A^{(v)}$, and similarly define

indicator functions $G^{(u)} : P(m, d/4) \to \{0,1\}$ and $H^{(u)} : P(m,d) \to \{0,1\}$ corresponding to the tables $C^{(u)}$ and $B^{(u)}$ respectively. Further, $H^{(u)}$ can be assumed to be folded over $I^{(u)}$. By construction, we have for every u, v

$$\mathop{\mathbb{E}}_{f}[F^{(v)}(f)] = \mathop{\mathbb{E}}_{g}[G^{(u)}(g)] = \mathop{\mathbb{E}}_{h}[H^{(u)}(h)] = \rho\,. \tag{24}$$

and

$$\delta \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{u,v}\Big[\mathbb{E}\big[F^{(v)}(f_1)F^{(v)}(f_1+f)G^{(u)}(g_1)G^{(u)}(g_1+g)H^{(u)}(\tilde{g})H^{(u)}(\tilde{g}+gh+\overline{g}\,\overline{f\circ\pi_{uv}})\big]\Big] = 0\,, \tag{25}$$

where the inner expectation is over the choice of all the functions $f, f_1, g, g_1, \tilde{g}, h$. Our goal is to prove that (24) and (25) imply $\rho \leqslant 2^{-\Omega(\ell)}$. We now analyze the inner expectation in (25) for a fixed (u,v), call it $\Gamma_{(u,v)}$. Let us use the shorthand

$$F = F^{(v)},\ G = G^{(u)},\ H = H^{(u)},\ \text{ and }\ \pi = \pi_{uv}\,.$$

Define the "self-corrected" versions $\tilde{F}$ and $\tilde{G}$ of the tables F and G as

$$\tilde{F}(f) = \mathop{\mathbb{E}}_{f_1}[F(f_1)F(f_1+f)] \qquad \text{and} \qquad \tilde{G}(g) = \mathop{\mathbb{E}}_{g_1}[G(g_1)G(g_1+g)]$$

respectively. Note that the tables $\tilde{F}$ and $\tilde{G}$ take values in the interval $[0,1]$.

As in the proof of Theorem 27, using Fourier expansion and eliminating some zero terms, the expectation $\Gamma_{(u,v)}$ can be written as the sum

$$\sum_{\beta\in\Lambda(m,d)} \widehat{H}(\beta)^2 \underbrace{\mathop{\mathbb{E}}_{g}\Big[\tilde{G}(g)\ \mathop{\mathbb{E}}_{h}\big[\chi_\beta(gh)\big]\ \mathop{\mathbb{E}}_{f}\big[\tilde{F}(f)\chi_\beta(\overline{g}\overline{f\circ\pi})\big]\Big]}_{\Upsilon_g}\,. \tag{26}$$

The $\beta = 0$ term equals

$$\widehat{H}(0)^2\ \mathop{\mathbb{E}}_{g}[\tilde{G}(g)]\ \mathop{\mathbb{E}}_{f}[\tilde{F}(f)] = \Big(\mathop{\mathbb{E}}_{h}[H(h)]\Big)^2\ \Big(\mathop{\mathbb{E}}_{g}[G(g)]\Big)^2\ \Big(\mathop{\mathbb{E}}_{f}[F(f)]\Big)^2 = \rho^6 \quad \text{using (24)}\,. \tag{27}$$

Our goal is to prove that the rest of the terms (for $\beta \neq 0$) in (26) have a very small contribution. To this end, we proceed similarly to the proof of Theorem 27. First, the terms in (26) with $\mathsf{wt}(\beta) \geqslant 2^{d/2}$ can be bounded in absolute value by $\mathbb{E}_g\big[|\mathbb{E}_h\,\chi_\beta(gh)|\big] \leqslant 2^{-2^{d/4}}$ due to Theorem 17. For terms with $\mathsf{wt}(\beta) < 2^{d/2}$, note that

$$\mathop{\mathbb{E}}_{h}[\chi_\beta(gh)] = \mathop{\mathbb{E}}_{h}[\chi_{\beta g}(h)] = 0$$

unless $\beta g = 0$. This follows from Fact 7 because $\mathsf{wt}(\beta g) < 2^{d/2}$ and so βg cannot be a nonzero polynomial of degree $P(m, m - 3d/4 - 1)$. Expanding $\tilde{F}(f) = \sum_\alpha \widehat{F}(\alpha)^2\chi_\alpha(f)$, we can simplify the expected value Υ_g in (26) as

$$\begin{aligned}
\Upsilon_g &= \mathop{\mathbb{E}}_{g}\Big[\tilde{G}(g)\ \mathbf{1}(\beta g = 0)\ \mathop{\mathbb{E}}_{f}\Big[\sum_\alpha \widehat{F}(\alpha)^2\chi_\alpha(\overline{f})\chi_{\beta\overline{g}}(f\circ\pi)\Big]\Big] \\
&= \mathop{\mathbb{E}}_{g}\Big[\tilde{G}(g)\ \mathbf{1}(\beta g = 0)\ (-1)^{\mathsf{wt}(\pi_2(\beta))}\widehat{F}(\pi_2(\beta))^2\Big] \quad \text{(using Fact 10)} \\
&\geqslant \begin{cases} 0 & \text{when } \mathsf{wt}(\beta) \text{ is even} \\ -\widehat{F}(\pi_2(\beta))^2 & \text{when } \mathsf{wt}(\beta) \text{ is odd} \end{cases}
\end{aligned} \tag{28}$$

where in the last step we use the fact that $\mathsf{wt}(\beta)$ and $\mathsf{wt}(\pi_2(\beta))$ have the same parity.

Combining (26), (27), and (28), we can lower bound δ from (25) as

$$\delta \geqslant \rho^6 - 2^{-2^{d/4}} - \sum_{\substack{\beta:\mathsf{wt}(\beta)<2^{d/2} \\ \mathsf{wt}(\beta)\text{ odd}}} \widehat{F}(\pi_2(\beta))^2 \widehat{H}(\beta)^2 \ .$$

Appealing to Lemma 15, the sum in the above expression is at most $2^{d-\Omega(\ell)}$ when the LABEL-COVER instance G is at most $2^{-\Omega(\ell)}$-satisfiable. Recalling $\delta = 0$ and $\rho = 1/2^k$, we conclude $k \geqslant \Omega(\ell)$ when $d = \Theta(\log \ell)$. □

Picking parameters as in Section 4.2 we get a proof of Theorem 2 (alternatively stated as Theorem 28 at the beginning of this section).

6 Concluding remarks

We have shown some new hardness of approximation results based on the low-degree long code, with an eye towards hardness of approximate hypergraph coloring. The elegant connection discovered in [2] between the strong analyis of Reed-Muller testing due to [5] and the soundness of low-degree long code constructions has also proved fruitful. Indeed, in a follow-up paper by the second author and coauthors [11] our results have been pushed further to show various hardness of approximate hypergraph coloring results (with quantitatively similar hardness factors), thereby answering the most immediate question left open by our work.

It is clear that replacing the long code by the low-degree version yields quantitative improvements but it is still not clear how far these can be pushed. In particular, can one derandomize the long code further and move closer to $N^{\Omega(1)}$ (or at least $2^{(\log N)^{\Omega(1)}}$) hardness for hypergraph coloring or related problems? Recently, Khot and Saket were able to use a different "outer verifier" together with the degree-2 long code to prove hardness of $2^{(\log N)^{\Omega(1)}}$-coloring 2-colorable 12-uniform hypergraphs [20].

In addition to the quantitative bottleneck currently given by the low-degree long code, our starting point LABEL-COVER being the result of parallel repetition caused the reduction to be (slightly) super-polynomial. It is interesting whether the parallel-repetition-based LABEL-COVER can be replaced by an outer PCP that makes more than 2 (but constant) number of queries, a la [7]. Such PCPs have nearly polynomial hardness gaps but *without* the super polynomial blow up in the reduction. Two difficulties that would need to be addressed are: first, these PCPs make $O(1)$ queries and not just two, so one needs to be able to simulate such constraints with long codes; and second, the constraint structure is not direct product so it might be harder to use the low-degree long code to fold over such constraints.

It seems inevitable, although we are not aware of a formal connection, that for reaching hardness of approximation to within polynomial factors we might have to first resolve the barrier of the sliding-scale conjecture, see [4, 22].

7 Acknowledgement

We are grateful to the anonymous referee for many thoughtful and helpful comments, and for catching a bug in the original soundness argument of Appendix B.

References

[1] Sanjeev Arora and Carsten Lund. *Approximation Algorithms for NP-hard Problems*, chapter Hardness of Approximations. PWS Publishing, 1996. 7

[2] Boaz Barak, Parikshit Gopalan, Johan Håstad, Raghu Meka, Prasad Raghavendra, and David Steurer. Making the long code shorter. In *Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science*, pages 370–379, 2012. 3, 4, 8, 22

[3] Mihir Bellare, Oded Goldreich, and Madhu Sudan. Free bits, PCPs, and nonapproximability: Towards tight results. *SIAM J. Comput.*, 27(3):804–915, 1998. 3, 11

[4] Mihir Bellare, Shafi Goldwasser, Carsten Lund, and Alexander Russell. Efficient probabilistic checkable proofs and applications to approximation. In *Proceedings of the 26th Annual ACM Symposium on Theory of Computing*, page 820, 1994. 22

[5] Arnab Bhattacharyya, Swastik Kopparty, Grant Schoenebeck, Madhu Sudan, and David Zuckerman. Optimal testing of Reed-Muller codes. In *Proceedings of the 51th Annual IEEE Symposium on Foundations of Computer Science*, pages 488–497, 2010. 4, 12, 14, 22

[6] Manuel Bodirsky and Jan Kára. The complexity of equality constraint languages. *Theory Comput. Syst.*, 43(2):136–158, 2008. 5

[7] Irit Dinur, Eldar Fischer, Guy Kindler, Ran Raz, and Shmuel Safra. PCP characterizations of NP: Toward a polynomially-small error-probability. *Computational Complexity*, 20(3):413–504, 2011. 22

[8] Irit Dinur and Venkatesan Guruswami. PCPs via low-degree long code and hardness for constrained hypergraph coloring. In *Proccedings of the 54th Annual IEEE Symposium on Foundations of Computer Science*, pages 340–349, 2013. 1, 6

[9] Irit Dinur and Gillat Kol. Covering CSPs. In *Proceedings of the 28th Conference on Computational Complexity*, pages 207–218, 2013. Longer version appears as ECCC TR12-088 at `http://eccc.hpi-web.de/report/2012/088`. 4

[10] Irit Dinur, Oded Regev, and Clifford D. Smyth. The hardness of 3-uniform hypergraph coloring. *Combinatorica*, 25(5):519–535, 2005. 6

[11] Venkatesan Guruswami, Prahladh Harsha, Johan Håstad, Srikanth Srinivasan, and Girish Varma. Super-polylogarithmic hypergraph coloring hardness via low-degree long codes. In *Proceedings of the 46th ACM Symposium on Theory of Computing (STOC)*, 2014. 6, 16, 22

[12] Venkatesan Guruswami, Johan Håstad, and Madhu Sudan. Hardness of approximate hypergraph coloring. *SIAM J. Comput.*, 31(6):1663–1686, 2002. 4, 5, 19, 26

[13] Johan Håstad. Some optimal inapproximability results. *Journal of the ACM*, 48(4):798–859, 2001. 3, 5, 11, 15, 19, 25, 26

[14] Johan Håstad. On the NP-hardness of Max-Not-2. *SIAM J. Comput.*, 43(1):179–193, 2014. 5

[15] Daniel M. Kane and Raghu Meka. A PRG for lipschitz functions of polynomials with applications to sparsest cut. In *Proceedings of the ACM Symposium on Theory of Computing*, pages 1–10, 2013. Available as http://arxiv.org/abs/1211.1109. 3

[16] Subhash Khot. Hardness results for approximate hypergraph coloring. In *Proceedings on 34th Annual ACM Symposium on Theory of Computing*, pages 351–359, 2002. 5, 6

[17] Subhash Khot. Hardness results for coloring 3 -colorable 3 -uniform hypergraphs. In *Proceedings of 43rd Symposium on Foundations of Computer Science*, pages 23–32, 2002. 5, 6, 27

[18] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing*, pages 767–775, 2002. 3

[19] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM J. Comput.*, 37(1):319–357, 2007. 3

[20] Subhash Khot and Rishi Saket. Hardness of coloring 2-colorable 12-uniform hypergraphs with $2^{(\log n)^{\Omega(1)}}$ colors. *Electronic Colloquium on Computational Complexity (ECCC)*, 21:51, 2014. 6, 22

[21] Subhash Khot and Rishi Saket. Hardness of finding independent sets in 2-colorable and almost 2-colorable hypergraphs. In *Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1607–1625, 2014. 6

[22] Dana Moshkovitz. The projection games conjecture and the NP-hardness of ln n-approximating set-cover. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques - 15th International Workshop, APPROX 2012, and 16th International Workshop, RANDOM 2012*, pages 276–287, 2012. 22

[23] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. *Ann. Math.*, 171(1):295–341, 2010. 3

[24] Rishi Saket. Hardness of finding independent sets in 2-colorable hypergraphs and of satisfiable CSPs. In *Proceedings of the 29th annual IEEE Conference on Computational Complexity*, CCC '14, 2014. To appear; available as arXiv preprint arXiv:1312.2915. 6

A 3LIN PCP using low-degree long code

In this section we will use the low-degree long code code to prove the following theorem.

Theorem 30. *Given a system of linear equations over $\mathbb{F}_2$ with 3 variables per equation, of size N, it is quasi-NP-hard to distinguish between the following cases:*

- *There is an assignment satisfying at least $1 - 1/2^{\Omega(\sqrt{\log \log N})}$ fraction of the equations.*
- *Every assignment satisfies at most $1/2 + 1/2^{2^{\Omega(\sqrt{\log \log N})}}$ fraction of the equations.*

More than proving the theorem, our goal is to illustrate how to replace the long code by the low-degree long code code in the simplest of Håstad's PCP constructions.

We describe a reduction from an instance of LABEL-COVER to a system of linear equations over $\mathbb{F}_2$ with three variables per equation.

The Reduction. Let $G = (U, V, E)$ be a LABEL-COVER instance with parameter ℓ as promised in Theorem 5. The integer d will be a degree parameter that we will choose later.

For each $v \in V$ we add a block of variables corresponding to $P'(\ell, d)$ (recall that $P'(\ell, d)$ contains for each $g \in P(\ell, d)$ exactly one of g and $1 + g$). For each $u \in U$ let $I^{(u)}$ be the ideal spanned by $f_1^{(u)}, \ldots, f_\ell^{(u)}$ viewed as functions over $m = 3\ell$ bits such that $f_i^{(u)}$ only looks at the three relevant bits numbered $3i + 1, 3i + 2, 3i + 3$. For each u we add a block of variables corresponding to $P'(m, d)/I^{(u)}$.

An assignment for the variables is given by a collection of functions $A^{(v)} : P(\ell, d) \to \{-1, 1\}$ per v, and $B^{(u)} : P(m, d) \to \{-1, 1\}$ per u and, such that $B^{(u)}$ is folded over $I^{(u)}$ and over $\{-1, 1\}$, and $A^{(v)}$ is folded over $\{-1, 1\}$ (see Fact 13).

The equations are conveniently described by a randomized test. Recall that $L(m, d) \subseteq P(m, d)$ denotes the set of products of d linearly independent affine forms.

1. Choose a random edge (u, v) in the LABEL-COVER instance, and let $\pi_{uv} : F_2^{3\ell} \to \mathbb{F}_2^\ell$ be the associated projection,
2. Choose a random $g \in P(\ell, d)$, and a random $h \in P(m, d)$.
3. Choose t independently random functions $\xi_1, \ldots, \xi_t \in L(m, d)$, and let $\xi = \xi_1 + \cdots + \xi_t$.
4. Accept iff $A^{(v)}(g)B^{(u)}(h)B^{(u)}(h + \xi + g \circ \pi_{uv}) = 1$.

To analyze this reduction we follow Håstad's analysis of 3LIN using long codes [13], just replacing the analysis of the effect of the noise function in the soundness proof with Proposition 16.

Completeness. Given a perfectly satisfying assignment for the initial LABEL-COVER, let us assign each $B^{(u)} = \mathsf{SC}(b)$ where b is the label for u, and $A^{(v)} = \mathsf{SC}(a)$ where a is the label for v. In that case we get, since $g \circ \pi_{uv}(b) = g(a)$,

$$A^{(v)}(g)B^{(u)}(h)B^{(u)}(h + \xi + g \circ \pi_{uv}) = (-1)^{g(a)+h(b)+h(b)+\xi(b)+g\circ\pi_{uv}(b)} = (-1)^{\xi(\mathbf{b})}$$

which equals 1 with probability at least $1 - \frac{\mathsf{wt}(\xi)}{2^m} \geqslant 1 - t2^{-d} = 1 - 2^{-d/4}$.

Soundness. We next show that an assignment to the 3LIN system that satisfies $(1 + \varepsilon)/2$ of the equations, can be decoded into an assignment for the initial LABEL-COVER instance that satisfies $\mathrm{poly}(\varepsilon/2^d)$ fraction of the constraints. So assume a 3LIN assignment with

$$\varepsilon \leqslant \mathop{\mathbb{E}}_{u,v,g,h,\xi}[A(g)B(h)B(h + \xi + g \circ \pi_{uv})]. \tag{29}$$

(where we omit the dependence of A, B on u, v from the notation). Plugging in the Fourier expansion of A and B, for each v, u the expectation over g, h, ξ can be written as

$$\sum_{\alpha,\beta,\gamma} \widehat{A}(\alpha)\widehat{B}(\beta)\widehat{B}(\gamma) \mathop{\mathbb{E}}_{g,h,\xi}[\chi_\alpha(g)\chi_\beta(h)\chi_\gamma(h + \xi + g \circ \pi_{uv})],$$

summed over all $\alpha \in \Lambda(\ell, d)$, and $\beta, \gamma \in \Lambda(m, d)$. The expectation is zero unless $\alpha = \beta$ and $\gamma = \pi_2(\alpha)$ (see Fact 10). So (29) becomes

$$\varepsilon \leqslant \sum_\beta \widehat{B}(\beta)^2 \widehat{A}(\pi_2(\beta)) \mathop{\mathbb{E}}_{\xi}[\chi_\beta(\xi)] \leqslant \sum_\beta \widehat{B}(\beta)^2 \widehat{A}(\pi_2(\beta))\rho(\beta)^t. \tag{30}$$

where the last inequality follows from Proposition 16 with $\rho(\beta) = \max\left\{1 - \frac{\mathsf{wt}(\beta)}{2^d}, \rho_0\right\}$.

If $\mathsf{wt}(\beta) > 2^{d/2}$ for large enough d, then $\rho(\beta) \leqslant 1 - 1/2^{d/2}$ and $\rho(\beta)^t \leqslant \exp(-2^{d/4})$ for $t = 2^{3d/4}$. Using Cauchy-Schwarz inequality and Parseval's equality $\sum_\beta \widehat{B}(\beta)^2 \leqslant 1$, we can bound the sum of all terms for which $\mathsf{wt}(\beta) > 2^{d/2}$ by $\exp(-2^{d/8})$ so we are left with

$$\varepsilon - \exp(2^{-d/8}) \leqslant \sum_{\mathsf{wt}(\beta) \leqslant 2^{d/2}} \widehat{B}(\beta)^2 \widehat{A}(\pi_2(\beta)) \leqslant \sqrt{\sum_{\mathsf{wt}(\beta) \leqslant 2^{d/2}} \widehat{B}(\beta)^2 \widehat{A}(\pi_2(\beta))^2}\,, \tag{31}$$

where the last step we again used Cauchy-Schwarz inequality and Parseval's equality. As B is folded over $\{-1, 1\}$, the terms with $\mathsf{wt}(\beta)$ even in (31) are 0. Therefore we can restrict the summation to $\mathsf{wt}(\beta)$ (and therefore also $\mathsf{wt}(\pi_2(\beta))$) odd, which in particular means $\pi_2(\beta) \neq 0$. Appealing to Lemma 15, we can find a labeling satisfying $(\varepsilon - \exp(-2^{d/8}))^2/2^{d/2}$ fraction of the LABEL-COVER constraints.

Therefore we conclude that in the soundness case, every assignment to the 3LIN instance satisfies at most $\frac{1}{2} + 2^{d-\Omega(\ell)} + \exp(-2^{d/8})$ fraction of the constraints.

Parameters. Finally we pick parameters suitably to deduce Theorem 30. Let us pick $\ell = 2^{\lfloor\sqrt{\log\log n}\rfloor/8}$ and $d = \lfloor\sqrt{\log\log n}\rfloor$. The size of 3LIN instance produced will be at most polynomial in $N \leqslant n^{3\ell}2^{(3\ell)^d} \leqslant n^{2^{O(\sqrt{\log\log n})}}$, and the reduction will run in $N^{O(1)}$ time. As a function of N, we have $\ell \geqslant 2^{\Omega(\sqrt{\log\log N})}$. As the completeness is $1 - 2^{-\Omega(d)}$ and the soundness is $1/2 + 2^{-\Omega(\ell)} + 2^{-2^{\Omega(d)}}$, the bounds claimed in Theorem 30 follow.

B Hardness of hypergraph coloring, based on the long code

The result of Section 5 showed a covering CSP of arity 6 for which it is hard to tell if the instance is satisfiable or has covering number exceeding $\exp(\Omega(\sqrt{\log\log N}))$. The constraints of the CSP were for the form $(x_1 \neq x_2) \vee (y_1 \neq y_2) \vee (z_1 \neq z_2)$. In this section, our goal is show a similar super-constant hardness for the covering CSP "monotone Not-all-Equal-8SAT" whose constraints check that the 8 variables in their scope are not all equal. The motivation is that this directly corresponds to showing hardness results for coloring 2-colorable 8-hypergraphs. Note that the test made by the construction 6-NE-PCP has a "tripartite" structure with the queries of each check coming from all three parts. This means that the corresponding hypergraph is always trivially 2-colorable.

To get a result for hypergraph coloring, we need a test that makes all its queries on a single side. In this section, we describe and analyze such a test. However, this test needs access to the long code of the labels, and we have not been able to design a similar test using only the low-degree long code. As a result, we will only show hardness of distinguishing satisfiable instances from those with covering number $\Omega(\log\log N)$. In coloring terms, we show that $(\log N)^c$-coloring a 2-colorable 8-uniform hypergraph is hard for some absolute constant $c > 0$. The previous best result for 2-*colorable* hypergraphs showed hardness of coloring with $O(\log\log N)$ colors [12] (but it worked for 4-uniform hypergraphs).

We reiterate that there is no use of the low-degree long code in this section. The ingredients needed in this section were available circa 1997 after Håstad's work [13], and are similar to those of his result on 4SAT. One simple but useful trick we make use of is to work with the 0-1 indicator vector of the candidate independent set in the soundness analysis (instead of working with k proofs to establish large covering

number). This approach was used in [17] to show a super-constant hardness for coloring 3-colorable 3-uniform hypergraphs.

Let $G = (U, V, E)$ be a LABEL-COVER instance with parameter ℓ as promised in Theorem 5. Let $m = 3\ell$. Our test will use *long code tables* only on the "larger" U side. Specifically, for each $u \in U$ we will have a table $D^{(u)} : \mathcal{F}_m / I^u$ where $I^{(u)}$ is the ideal spanned by the constraints that must be satisfied by the label to u. We won't assume these tables are folded over $\{-1, 1\}$ (this is important so we get a Not-All-Equal-8SAT instance without negations, i.e., an instance of 8-set splitting). Once again, we will extend $D^{(u)}$ to all of $\mathcal{F}_m$ by defining $B^{(u)}(h) = D^{(u)}(h + I^{(u)})$, and assume that $B^{(u)} : \mathcal{F}_m \to \{-1, 1\}$ is folded over $I^{(u)}$.

We now describe our PCP which we call 8-SS-PCP:

1. Pick a random $v \in V$, and independently sample (with replacement) two random neighbors $u, u' \in U$ of v.

   For notational simplicity, denote $B = D^{(u)}$ and $C = D^{(u')}$. Also let $\pi = \pi_{uv} : \mathbb{F}_2^m \to \mathbb{F}_2^\ell$ be the associated projection from the label of u to that of v, and similarly let $\pi' = \pi_{u'v}$ be the projection from the label of u' to that of v.
2. Sample $f \in \mathcal{F}_\ell$, $g_1, g_2, g_3, h_1, h_2, h_3 \in \mathcal{F}_m$ uniformly and independently at random.
3. Sample $g_4 \in \mathcal{F}_m$ as follows: For $y \in \mathbb{F}_2^m$, if $g_1(y) \neq g_2(y)$ then set $g_4(y)$ randomly, else set $g_4(y) = 1 + f(\pi(y)) + g_3(y)$.
4. Sample $h_4 \in \mathcal{F}_m$ as follows: For $y \in \mathbb{F}_2^m$, if $h_1(y) \neq h_2(y)$ then set $h_4(y)$ randomly, else set $h_4(y) = f(\pi'(y)) + h_3(y)$.
5. Accept if $(B(g_1) \neq B(g_2)) \vee (B(g_3) \neq B(g_4)) \vee (C(h_1) \neq C(h_2)) \vee (C(h_3) \neq C(h_4))$.

Perfect completeness. Given a perfectly satisfying assignment for the initial LABEL-COVER, let us assign each $D^{(u)}$ to be the long code of the label for u. If a, a' are the labels assigned to the nodes u, u' chosen by 8-SS-PCP and b the label to v, the above check made by 8-SS-PCP amounts to checking that not all of $\{g_i(a), h_i(a')\}_{i=1}^4$ are equal. To prove this, suppose $g_1(a) = g_2(a)$. By the definition of g_3, this means $g_4(a) = 1 + f(b) + g_3(a)$. If $f(b) = 0$, we would have $g_4(a) \neq g_3(a)$. Similarly, assuming $h_1(a') = h_2(a')$, we would have $h_4(a') \neq h_3(a')$ when $f(b) = 1$. Thus at least one of the pairs $(g_1(a), g_2(a))$, $(g_3(a), g_4(a))$, $(h_1(a'), h_2(a'))$ or $(h_3(a'), h_4(a'))$ are not equal. Note that we would clearly have perfect completeness for the weaker check of not-all-equal.

Covering soundness. For the soundness, we will be prove that for some $\rho = 2^{-\Omega(\ell)}$, every subset consisting of ρ fraction of the vertices in the hypergraph must contain a hyperedge (i.e., one of the 8-query patterns made by 8-SS-PCP). Let A be 0-1 characteristic function of a subset S of fraction ρ of vertices, and let $A^{(u)}$ be the restriction of A to the long code table associated with $u \in U$. Each $A^{(u)}$ will be folded over $I^{(u)}$, and we have $\mathbb{E}_u \mathbb{E}_g[A^{(u)}(g)] = \rho$ for g chosen uniformly at random from $\mathcal{F}_m$.

The probability that all 8 queries fall inside S, which is the fraction of hyperedges inside S, is given by

$$\delta \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{v,u,u'} \Big[\mathop{\mathbb{E}}_{f,g_i,h_i} [B(g_1)B(g_2)B(g_3)B(g_4)C(h_1)C(h_2)C(h_3)C(h_4)] \Big] \tag{32}$$

where we denote $B = A^{(u)}$ and $C = A^{(u')}$ for notational simplicity. Let us expand the inner expectation

over the functions f, g_i, h_i for a fixed v, u, u' using Fourier analysis as

$$\sum_{\beta_i,\gamma_i}\Big(\prod_{i=1}^{4}(\widehat{B}(\beta_i)\widehat{C}(\gamma_i))\,\mathbb{E}\Big[\prod_{i=1}^{4}\chi_{\beta_i}(g_i)\prod_{i=1}^{4}\chi_{\gamma_i}(h_i)\Big]\Big) \tag{33}$$

summed over $\beta_i, \gamma_i \in \mathcal{F}_m$ for $1 \leqslant i \leqslant 4$. In what follows, we will also equivalently treat β_i, γ_i as subsets of $\mathbb{F}_2^m$ (equal to the support of the respective functions). We will now argue that the only nonzero terms in the above sum are when $\beta_3 = \beta_4$, $\beta_1 = \beta_2 \subseteq \beta_3$ (and similarly for the γ_i's). Indeed, if $y \in \beta_3 \setminus \beta_4$, then $\mathbb{E}\Big[\prod_{i=1}^{4}\chi_{\beta_i}(g_i)\prod_{i=1}^{4}\chi_{\gamma_i}(h_i)\Big]$ has a factor of $\mathbb{E}[(-1)^{g_3(y)}]$ which is 0 (as the value $g_3(y)$ is independent of everything else in the product). A similar argument holds if $\beta_4 \setminus \beta_3 \neq \emptyset$. Thus we must have $\beta_3 = \beta_4 = \beta$ (say) for the expectation in (33) to be nonzero. If there exists a $y \in \beta_1 \setminus \beta$ (resp. $\beta_2 \setminus \beta$), then the distribution of $g_1(y)$ (resp. $g_2(y)$) is independent of the rest of the values in the product, again making $\mathbb{E}\big[\prod_{i=1}^{4}\chi_{\beta_i}(g_i)\prod_{i=1}^{4}\chi_{\gamma_i}(h_i)\big] = 0$. Thus we must have $\beta_1, \beta_2 \subseteq \beta$. Finally, if there exists $y \in \beta_1 \setminus \beta_2$, then $g_1(y)$ is still uniformly distributed given the values of g_i on β_i for $i = 2, 3, 4$, and so once again the relevant expectation will be 0. Using similar arguments for the γ_i's, we can simplify (33) as

$$\sum_{\substack{\beta_1\subseteq\beta\\ \gamma_1\subseteq\gamma}}\widehat{B}(\beta_1)^2\widehat{B}(\beta)^2\widehat{C}(\gamma_1)^2\widehat{C}(\gamma)^2\,\mathbb{E}\Big[\chi_{\beta_1}(g_1+g_2)\chi_{\beta}(g_3+g_4)\chi_{\gamma_1}(h_1+h_2)\chi_{\gamma}(h_3+h_4)\Big]\,. \tag{34}$$

Now, if $g_1(y) \neq g_2(y)$ for some $y \in \beta$, then $g_4(y)$ is independent of $g_3(y)$, making $\mathbb{E}[(-1)^{g_3(y)+g_4(y)}] = 0$, and the inner expectation in (34) above is 0 as well. Thus, the expectation vanishes unless $g_1(y) = g_2(y)$ for all $y \in \beta$ and $h_1(z) = h_2(z)$ for all $z \in \gamma$. These requirements on g_1, g_2, h_2, h_2 are met with probability $2^{-\mathsf{wt}(\beta)} \cdot 2^{-\mathsf{wt}(\gamma)}$, and in this case, we have

$$\chi_\beta(g_3+g_4) = \chi_\beta(1+f\circ\pi) = (-1)^{\mathsf{wt}(\beta)}\chi_{\pi_2(\beta)}(f)\,,$$

and $\chi_\gamma(h_3+h_4) = \chi_{\pi_2'(\gamma)}(f)$. Here $\pi_2(\beta) : \mathbb{F}_2^\ell \to \mathbb{F}_2$ is defined as before to be $\pi_2(\beta)(x) = \sum_{y\in\pi^{-1}(x)}\beta(y)$ (and similarly for $\pi_2'(\gamma)$). Therefore, (34) simplifies further to

$$\sum_{\substack{\beta_1\subseteq\beta\\ \gamma_1\subseteq\gamma}}\widehat{B}(\beta_1)^2\widehat{B}(\beta)^2\widehat{C}(\gamma_1)^2\widehat{C}(\gamma)^2(-1)^{\mathsf{wt}(\beta)}2^{-\mathsf{wt}(\beta)}2^{-\mathsf{wt}(\gamma)}\,\mathop{\mathbb{E}}_{f}[\chi_{\pi_2(\beta)}(f)\chi_{\pi_2'(\gamma)}(f)]\,.$$

As $\mathbb{E}_f[\chi_\alpha(f)\chi_{\alpha'}(f)] = 0$ when $\alpha \neq \alpha'$ and 1 otherwise, we can simplify (32) and obtain

$$\delta = \mathop{\mathbb{E}}_{v,u,u'}\Big[\sum_{\substack{\beta_1\subseteq\beta;\ \gamma_1\subseteq\gamma\\ \pi_2(\beta)=\pi_2'(\gamma)}}\widehat{B}(\beta_1)^2\widehat{B}(\beta)^2\widehat{C}(\gamma_1)^2\widehat{C}(\gamma)^2(-1)^{\mathsf{wt}(\beta)}2^{-\mathsf{wt}(\beta)}2^{-\mathsf{wt}(\gamma)}\Big]\,. \tag{35}$$

As $\widehat{B}(0) = \mathbb{E}_g[B(g)]$ and $\widehat{C}(0) = \mathbb{E}_h[C(h)]$, the term with $\beta = \gamma = 0$ above equals $\big(\mathbb{E}_g[B(g)]\big)^4\big(\mathbb{E}_g[C(h)]\big)^4$. Taking expectation over v, u, u', and using the regularity of the instance, the terms with $\beta = \gamma = 0$ contribute at least $\big(\mathbb{E}_{u,g}[A^{(u)}(g)]\big)^8 = \rho^8$ to (35). Our goal is to prove that the other terms have a very small contribution.

The terms with $\mathsf{wt}(\beta)$ even in (35) are positive and so can be ignored in any lower bound on δ. When $\mathsf{wt}(\beta)$ is odd, $\mathsf{wt}(\pi_2(\beta))$ is also odd, and in particular $\pi_2(\beta) \neq 0$. Using these facts and $\sum_{\beta_1\subseteq\beta}\widehat{B}(\beta_1)^2 \leqslant 1$,

$\sum_{\gamma_1 \subseteq \gamma} \widehat{C}(\gamma_1)^2 \leqslant 1$ in (35), we get the lower bound:

$$\begin{aligned}\delta &\geqslant \rho^8 - \mathop{\mathbb{E}}_{v,u,u'}\Bigg[\sum_{\substack{\beta,\gamma \\ \pi_2(\beta)=\pi_2'(\gamma)\neq 0}} \widehat{B}(\beta)^2\widehat{C}(\gamma)^2\, 2^{-\mathsf{wt}(\beta)}2^{-\mathsf{wt}(\gamma)}\Bigg] \\ &\geqslant \rho^8 - 2^{\ell} - \mathop{\mathbb{E}}_{v,u,u'}\Bigg[\sum_{\substack{\beta,\gamma;\mathsf{wt}(\beta),\mathsf{wt}(\gamma)<\ell \\ \pi_2(\beta)=\pi_2'(\gamma)\neq 0}} \widehat{B}(\beta)^2\widehat{C}(\gamma)^2\Bigg]. \qquad (36)\end{aligned}$$

An argument similar to Lemma 15 shows that the expectation in (36) is at most $2^{-\Omega(\ell)}$. Therefore, $\delta > 0$ when $\rho \geqslant 1/2^k$ for some $k = \Theta(\ell)$. In other words, the hypergraph consisting of the query patterns of 8-SS-PCP does not have an independent set of density $2^{-\Omega(\ell)}$.

Parameter choices. Picking $\ell = \lfloor \log\log n \rfloor/4$, the size of the instance produced will be $N = n^{O(\ell)}2^{2^{3\ell}} \leqslant n^{O(\log\log n)}$. When the LABEL-COVER instance is satisfiable, the hypergraph will be 2-colorable, and in the soundness case, the hypergraph will contain no independent set of size $N/(\log n)^{\Omega(1)}$. Therefore we can conclude the following result.

Theorem 31 (Restating Theorem 4)**.** *Assume that* NP *does not admit $n^{O(\log\log n)}$ time algorithms. There is an absolute constant $c > 0$ such that the following holds. Given a 8-uniform hypergraph on N vertices, there is no polynomial time algorithm to distinguish between the following two cases:*

- *The hypergraph can be colored with 2 colors so that every hyperedge is bichromatic.*
- *The hypergraph does not have an independent set with $N/(\log N)^c$ vertices, and in particular any coloring of the vertices with $(\log N)^c$ colors will have a monochromatic hyperedge.*

ECCC ISSN 1433-8092

http://eccc.hpi-web.de