On Approximability of Satisfiable k-CSPs: I

Amey Bhangale
University of California, Riverside
Riverside, USA
ameyb@ucr.edu

Subhash Khot[*]
New York University
New York, USA
khot@cs.nyu.edu

Dor Minzer[†]
Massachusetts Institute of Technology
Cambridge, USA
khot@cs.nyu.edu

ABSTRACT

We consider the P-CSP problem for 3-ary predicates P on satisfiable instances. We show that under certain conditions on P and a $(1, s)$ *integrality gap* instance of the P-CSP problem, it can be translated into a dictatorship vs. quasirandomness test with perfect completeness and soundness $s + \varepsilon$, for every constant $\varepsilon > 0$. Compared to Ragahvendra's result [STOC, 2008], we do not lose perfect completeness. This is particularly interesting as this test implies new hardness results on satisfiable constraint satisfaction problems, assuming the Rich 2-to-1 Games Conjecture by Braverman, Khot, and Minzer [ITCS, 2021]. Our result can be seen as the first step of a potentially long-term challenging program of characterizing optimal inapproximability of every satisfiable k-ary CSP.

At the heart of the reduction is our main analytical lemma for a class of 3-ary predicates, which is a generalization of a lemma by Mossel [Geometric and Functional Analysis, 2010]. The lemma and a further generalization of it that we conjecture may be of independent interest.

CCS CONCEPTS

• **Theory of computation → Problems, reductions and completeness**.

KEYWORDS

constraint satisfaction problems, hardness of approximation, non-abelian groups

ACM Reference Format:
Amey Bhangale, Subhash Khot, and Dor Minzer. 2022. On Approximability of Satisfiable k-CSPs: I. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing (STOC '22), June 20–24, 2022, Rome, Italy.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3519935.3520028

1 INTRODUCTION

Constraint satisfaction problems (CSPs) are some of the most fundamental problems in computer science. Given a predicate $P : \Sigma^k \to \{0, 1\}$, for some alphabet Σ, a P-CSP instance consists of a set of variables $x_1, x_2, \ldots, x_n$ and a collection of *local* constraints $C_1, C_2, \ldots, C_m$. Each constraint is of the type $P(x_{i_1}, x_{i_2}, \ldots, x_{i_k})$. The constraints might involve *literals* instead of just the variables.[1] An algorithmic task is to decide if there exists an assignment to the variables that satisfies all the constraints. In a related problem, called the Max-P-CSP problem, the task is to find an assignment to the variables that satisfies the maximum fraction of the constraints. An α-approximation algorithm is a polynomial-time algorithm which always returns an assignment that satisfies at least $\alpha \cdot \textsc{Opt}$ fraction of the constraints, where $\textsc{Opt}$ is the value of the optimum assignment. The focus of the current work is on approximability of fully satisfiable instances.

A systematic study of the complexity of solving CSPs was started by Schaefer in 1978 [23] who showed that for every P over a 2-element alphabet, the problem of checking satisfiability of a P-CSP is either in **P** or is **NP**-complete. A famous Dichotomy Conjecture of Feder and Vardi [10], which was resolved recently in huge breakthroughs by Bulatov and Zhuk independently [8, 25], states that for every P, checking satisfiability of a P-CSP is either in **P** or is **NP**-complete.

However, when it comes to designing optimal approximation algorithms for Max-P-CSP on fully satisfiable instances, the question is wide open. The PCP Theorem [1, 2, 11] proved in the early 90s shows that it is **NP**-hard to approximate many P-CSPs within a constant factor $\alpha < 1$. This was vastly improved in a seminal result by Håstad [13] for certain CSPs. Håstad showed that for many CSPs, it is **NP**-hard to do better than the approximation factor achieved by a random assignment. More specifically, he showed that 3SAT cannot be approximated better than $\frac{7}{8} + \varepsilon$ for any constant $\varepsilon > 0$ in polynomial time unless $\mathbf{P} = \mathbf{NP}$. Note that if we select a random assignment, then it satisfies $\frac{7}{8}$-fraction of the clauses in expectation. The result proved in [13] is stronger than what is stated – even if we know that a given instance is *fully satisfiable*, i.e., there exists an assignment that satisfies all the clauses, it is **NP**-hard to come up with an assignment that satisfies more than $(\frac{7}{8} + \varepsilon)$-fraction of the clauses for any constant $\varepsilon > 0$.

Håstad also showed that it is **NP**-hard to find an assignment to a given 3LIN instance[2] that satisfies more that $(\frac{1}{2} + \varepsilon)$-fraction of the constraints, even if we are guaranteed that there exists an assignment that satisfies $(1 - \varepsilon)$-fraction of the constraints. This is interesting because unlike 3SAT, we can in fact find an assignment that satisfies all the constraints of a given 3LIN instance, if there exists one, in polynomial time. Thus, knowing that a given instance of P-CSP is fully satisfiable, in principle, can be used to design better approximation algorithms for Max-P-CSPs. In this paper, we study the inapproximability of fully satisfiable instances. On the other hand, as we will explain next, if the instance is almost satisfiable, then by Raghavendra's work [21], we know the precise

[*]Research funded by NSF grant CCF-1813438, Simons Investigator Award, and Simons Collaboration on Algorithms and Geometry.
[†]Research supported by a Sloan Research Fellowship.

STOC '22, June 20–24, 2022, Rome, Italy

ACM ISBN 978-1-4503-9264-8/22/06.
https://doi.org/10.1145/3519935.3520028

[1]See Definition 2.7 and Remark 2.8 for more details.
[2]This CSP is over the Boolean domain and constraints are of the type $x_{i_1} \oplus x_{i_2} \oplus x_{i_3} = 1/0$.

approximation threshold for every P-CSP and the optimal algorithm is given by semi-definite programming.

In order to gain better understanding of complexity of approximation algorithms for various optimization problems, Khot [17] in 2002 proposed the Unique Games Conjecture (UGC). Since then, for various optimization problems, we now know the precise approximation factor that one can achieve in polynomial time assuming the UGC. Max-Cut is one of the simplest CSPs in which the constraints are of the type $x \oplus y = 1$. Goemans and Williamson [12] gave a α_{GW}-approximation algorithm for Max-Cut problem where $\alpha_{GW} \approx 0.878$. Surprisingly, [18] showed that the approximation algorithm by Goemans and Williamson is tight assuming the UGC. Their hardness result relied on the 'Majority is Stablest' theorem which was proved in [20].

For general CSPs, Austrin and Mossel [3] gave a very simple sufficient criterion for a predicate P to be *approximation resistant*. A predicate P is called approximation resistant if it is **NP**-hard (or UG-hard) to achieve an approximation algorithm better than the random assignment algorithm. 3SAT and 3LIN predicates described above are examples of approximation resistant predicates. Austrin and Mossel showed that if there exists a distribution supported only on the satisfying assignments in P, which is balanced and pairwise independent, then P is approximation resistant assuming the UGC.

The Max-Cut hardness result was beautifully generalized to *all* constraint satisfaction problems by Raghavendra [21]. More specifically, he showed that for any P-CSP problem, if there exists a (c, s) basic SDP integrality gap instance[3], then it is UG-hard to find an assignment that satisfies $(s + \varepsilon)$ fraction of the constraints, even if the given instance is $(c - \varepsilon)$-satisfiable, for every constant $\varepsilon > 0$. For all $c \in (0, 1]$, let $s(c)$ be the infimum value such that there exists an $(c, s(c))$ integrality gap instance. By definition, the SDP relaxation promises $s(c)$ satisfying assignment on every c-satisfiable instance. Raghavendra gives the rounding algorithm that actually finds the $s(c)$-satisfying assignment. Thus, Raghavendra's result gives a complete answer to the complexity of approximating Max-P-CSP assuming the UGC. However, it does not imply hardness on instances that are fully satisfiable. This is because in translating the integrality gap parameters (c, s) to hardness parameter, there is always a loss of some small constant $\varepsilon > 0$ in the completeness parameter.

The most important building-block in Raghavendra's result (and also in many prior works) is the dictatorship test. A function $f : \Sigma^n \to \Sigma$ is called a dictator function if it depends only on one variable. A dictatorship test is a procedure which queries f at a few (correlated) locations randomly and based on the function values at these locations, it decides if f is a dictator function or *far* from any dictator function (also referred to as *quasirandom* functions). We briefly describe the notion of being *far* from dictator functions here. Influence of a coordinate i in a function f is the probability that for a random input $(x_1, x_2, \ldots, x_n)$, f changes its value if we change the i^{th} coordinate. Note that dictator functions have one coordinate whose influence is 1. A function is called far from dictator functions if for every coordinate i, the influence of the coordinate i in f is small.

[3]See Definition 2.16 for the formal definition.

There are three important properties of the test which are useful in getting hardness of approximation result for Max-P-CSP. The first one is the *completeness parameter* c — this is the probability that the test accepts any dictator function. The second property is the *soundness parameter* s — this is the probability with which the test accepts far from dictator functions. The third property is the decision predicate that the test uses in accepting or rejecting the function f. If the decision predicate is P and the test has completeness c and soundness s, then such a test can be translated into a UG-hardness result for Max-P-CSP with completeness $(c - \varepsilon)$ and soundness $(s + \varepsilon)$, for any constant $\varepsilon > 0$. In other words, it is UG-hard to find an assignment that satisfies $(s + \varepsilon)$ fraction of the constraints, even if the given instance is $(c - \varepsilon)$-satisfiable, for every constant $\varepsilon > 0$.

Raghavendra proved his result by designing a dictatorship test starting with a (c, s) integrality gap instance for Max-P-CSP such that the test has completeness $(c - \varepsilon)$ and soundness $(s + \varepsilon)$, for any constant $\varepsilon > 0$. Therefore, his test loses in the completeness parameter and hence cannot be used in proving hardness result on satisfiable instances. Note that even if the completeness parameter of the test is c, because of the conjectured hardness of Unique Games, one still loses small constant ε in the completeness parameter of the final UG-hardness result.[4] In order to save this loss, Braverman, Khot, and Minzer [7] proposed a Rich 2-to-1 Games Conjecture and if we use this instead of Unique Games, then there is no loss in the completeness parameter. Therefore, it is important that we do not lose anything in the completeness parameter when designing the dictatorship test.

In this work, we initiate a systematic study of completely characterizing the precise approximability of every k-ary CSP on satisfiable instances (recognizing, of course, that the prior works have obtained such a characterization for specific predicates, e.g., 3SAT). In order to answer this challenging question, it was necessary first to understand the complexity of checking satisfiability of CSP which is the famous Dichotomy Conjecture. Now that this conjecture is resolved, we can embark on the study of approximability of satisfiable CSPs.

As with the case with 3SAT and 3LIN, a predicate being linear makes a big difference on the complexity of the CSP. Addressing this issue of linearity is also a challenging aspect in the proof of the Dichotomy Conjecture. In this work, we take the first step by considering special class of non-linear predicates. We show how to convert any $(1, s)$-integrality gap instance of a 3-ary CSP to a dictatorship test with completeness 1 and soundness $s + \varepsilon$, for any constant $\varepsilon > 0$. For our conclusion to hold, we need a few additional properties from the predicate as well as from the integrality gap instance that we describe next.

- *Predicates not satisfying any linear embedding:* Given a predicate $P \colon \Sigma^3 \to \{0, 1\}$, it is said to satisfy a linear equation if there exists an Abelian group $(G, +)$ and 3 embeddings $\sigma : \Sigma \to G$, $\phi : \Sigma \to G$ and $\gamma : \Sigma \to G$ such that the following hold: At least one of the embeddings is non-constant and for every tuple $(x, y, z) \in P^{-1}(1)$, $\sigma(x) + \phi(y) + \gamma(z) = 0$ where 0 is the identity element of G.

[4]Unique Games can hard only on almost satisfiable instances. Therefore, any hardness from Unique Games loses perfect completeness.

- *Semi-rich predicates:* A predicate $P\colon \Sigma^3 \to \{0,1\}$ is called semi-rich if for each $(x,y) \in \Sigma \times \Sigma$, there exists a $z \in \Sigma$ such that $(x,y,z) \in P^{-1}(1)$. Also, for every $(x,z) \in \Sigma \times \Sigma$, there exists a $y \in \Sigma$ such that $(x,y,z) \in P^{-1}(1)$.
- *SDP solution that is semi-rich and that is not linearly embeddable:* An SDP solution for a given P-CSP instance consists of a local distribution for each constraint. We say the SDP solution is semi-rich and is not linearly embeddable if the support of every local distribution is semi-rich and is not linearly embeddable in any Abelian group (See Definitions 2.1, 2.2 and 2.3).

We now state our main theorem.

Theorem 1.1. *Let $P\colon \Sigma^3 \to \{0,1\}$ be any predicate that satisfies the following conditions. (1) P does not satisfy any linear embedding, (2) P is a semi-rich predicate, and (3) there exists an instance of P-CSP that has a $(1,s)$-integrality gap for the basic SDP relaxation and an optimal SDP solution is semi-rich and is not linearly embeddable. Then for every $\varepsilon > 0$, there is a dictatorship test for P that has perfect completeness and soundness $s + \varepsilon$.*

We do not believe that the semi-rich condition is really needed in the theorem, but this is what we could show currently.

In order to focus on designing new dictatorship tests and a new way to analyze the tests, in this work we will not discuss in detail the application of this towards getting the conditional **NP**-hardness results. However, for completeness, we mention the following important corollary of our main theorem. This follows from a recent work by Braverman, Khot, Lifshitz and Minzer [6].

Corollary 1.2. *For a predicate P satisfying the conditions from Theorem 1.1, assuming the Rich-2-to-1 Games Conjecture with perfect completeness, for every constant $\varepsilon > 0$, it is **NP**-hard to find an assignment to a Max-P-CSP instance with value $s + \varepsilon$ even if the instance is fully satisfiable.*[5]

Note that in Corollary 1.2, the first condition from the hypothesis is *necessary* for such a statement to hold. This can be seen from the Max-3LIN problem on an Abelian group G.[6] This predicate has a linear embedding as well as there exists an instance with a SDP integrality gap of $(1, \frac{1}{|G|} + \varepsilon)$, for every constant $\varepsilon > 0$. However, if the instance is satisfiable, then one can find the satisfying assignment in polynomial time using Gaussian elimination.

It might be instructive to consider a couple of examples of predicates that satisfy the first two conditions.

(1) Linear equations over a quasirandom group: Fix any group $(G,\cdot)$ such that any non-trivial irreducible representation of G has dimension greater than 1. Consider the predicate $P_G : G^3 \to \{0,1\}$ where $P_G^{-1}(1) = \{(x,y,z) \mid x \cdot y \cdot z = 1_G\}$, where 1_G is the identity element. The fact that G does not have any non-trivial representation of dimension 1 implies that P does not satisfy any linear embedding. Also, it is easily observed that the predicate is semi-rich.

(2) Arithmetic progression over a quasirandom group: For a similar group as above, consider a predicate $P_{AP} : G^3 \to \{0,1\}$ where $P_{AP}^{-1}(1) = \{(x, x\cdot g, x \cdot g^2) \mid x, g \in G\}$. It can be shown that this predicate does not satisfy any linear embedding. To see that P_{AP} is semi-rich, we need to permute the coordinates. Note that permuting the coordinates of a predicate does not really change the complexity of the corresponding CSP problem. By the change of variables $x \cdot g = h$ we can write $P_{AP}^{-1}(1) = \{(h \cdot g^{-1}, h, h\cdot g) \mid h, g \in G\}$. We can permute the coordinates to get the following predicate $\tilde{P}_{AP}^{-1}(1) = \{(h, h\cdot g^{-1}, h \cdot g) \mid h, g \in G\}$. Now, it is easily observed that the predicate is semi-rich.

Remark 1.3. *A dictatorship test with optimal parameters (in fact, the optimal **NP**-hardness result for satisfiable instances) for the predicate P_G was shown by Bhangale and Khot [4]. Our main theorem gives new results for the predicate P_{AP} (and many more). The predicate P_{AP} is fundamentally different from P_G as it does not support any pairwise-independent distribution, whereas P_G does.*

1.1 Related Work

Many hardness results on satisfiable CSPs are known for specific CSPs. In this section, we state these results. Here, $\varepsilon > 0$ is an arbitrary small constant. Håstad [13] in his seminal result showed that for every $k \geqslant 3$, k-SAT is **NP**-hard to approximate within a factor of $1 - 1/2^k + \varepsilon$, even if the instance is satisfiable. Håstad and Khot [14] proved that Boolean CSPs on k variables are **NP**-hard to approximate within ratio $\frac{2^{O(k^{1/2})}}{2^k}$. For every prime p, they also showed the hardness result for CSPs over an alphabet of size p, where the hardness factor is $\frac{p^{O(k^{1/2})}}{p^k}$. Huang [15] improved the result for Boolean CSPs to the factor $\frac{2^{\tilde{O}(k^{1/3})}}{2^k}$. Brakensiek and Guruswami [5] formulated a problem called the 'V Label Cover' to improve these results on satisfiable k-ary CSPs. Towards this, assuming the hardness of the V Label Cover, they showed that there is an absolute constant c_0 such that for $k \geqslant 3$, given a satisfiable instance of Boolean k-CSP, it is hard to find an assignment satisfying more than $c_0 k^2/2^k$ fraction of the constraints. These results are non-trivial only for large values of k.

Towards getting an improved hardness result for Boolean satisfiable 3-CSPs, Håstad [16] showed that the predicate NTW[7] is **NP**-hard to approximate within a factor of $5/8 + \varepsilon$. For larger alphabet, Engebretsen and Holmerin [9] showed that 3-ary CSPs over an alphabet of size q is **NP**-hard to approximate within a factor of $\frac{1}{q} + \frac{1}{q^2} + \varepsilon$. Tang [24] showed a conditional result with the hardness factor $\frac{1}{q} + \frac{1}{q^2} - \frac{1}{q^3} + \varepsilon$.[8] Very recently, the first two authors [4] improved these results for 3-ary CSPs where where they showed that it is **NP**-hard to approximate satisfiable 3-ary CSPs over an alphabet of size q to within a factor of $\frac{1}{q} + \varepsilon$, for infinitely many q.

1.2 Techniques

For a given predicate P, we are interested in finding the maximum α_P such that (1) there exists an approximation algorithm that satisfies at least α_P fraction of the constraints on satisfiable instances,

[5] A notion of negation or literals or a mix of predicates is necessary for this result and Theorem 1.1 to hold (See Remark 2.8).

[6] Here, the predicate is $\{(x,y,z) \mid x + y + z = 0\}$ where 0 is the identity element in G.

[7] The satisfying assignments for the NTW (Not-Two-Ones) predicate are all 3 bit strings such that the number of 1s in them is not *two*.

[8] The theorem in [9] holds for every $q \geqslant 3$, and the theorem in [24] holds for every $q \geqslant 4$.

and (2) for all $\varepsilon > 0$, it is hard to find $(\alpha_P+\varepsilon)$-satisfying assignments on satisfiable instances.

In order to answer the above question, the starting point is the Dichotomy Theorem which gives a full characterization of predicates for which the corresponding CSP is **NP**-complete or is in **P** (i.e., deciding if $\alpha_P = 1$ or $\alpha_P < 1$). The characterization is based on whether a certain non-trivial polymorphism exists for a given predicate. For a given predicate $P : \Sigma^k \to \{0, 1\}$, a function $f : \Sigma^n \to \Sigma$ is called a polymorphism if for every $k \times n$ matrix constructed by letting every column to be an arbitrary satisfying assignment to P and letting $x_1, x_2, \ldots, x_k \in \Sigma^n$ be the rows of the matrix, it is the case that $(f(x_1), f(x_2), \ldots, f(x_k))$ is also a satisfying assignment to P. It is easy to see that a dictator function, i.e., $f(x) = x_i$ for some $1 \leqslant i \leqslant n$ is always a polymorphism, and any other polymorphism is called a non-trivial polymorphism. The Dichotomy Theorem states that for a predicate P, checking satisfiability of P-CSP is in **P** if there exists a non-trivial polymorphism; otherwise, it is **NP**-complete (ignoring some subtle issues).

Dictatorship test. Similar to polymorphisms, dictatorship tests form the back-bone of proving hardness of approximating Max-P-CSPs. Here we formally define the dictatorship test for a given predicate.

Definition 1.4. *A dictatorship test for a predicate $P : \Sigma^k \to \{0, 1\}$ can query a function $f : \Sigma^n \to \Sigma$. The test picks a random $k\times n$ matrix by letting every column to be a random satisfying assignment to P (i.e., in $P^{-1}(1)$, with some fixed distribution μ on $P^{-1}(1)$) and letting $x_1, x_2, \ldots, x_k \in \Sigma^n$ be the rows of the matrix. The test accepts if $(f(x_1), f(x_2), \ldots, f(x_k))$ is also a satisfying assignment to P.*

Here again, it is obvious that if f is a dictator function, then the test accepts with probability 1. If solving P-CSP is **NP**-complete then it has no non-trivial polymorphisms according to the Dichotomy Theorem. Therefore, the question here is to determine the maximum probability the test accepts a function f if f is *far* from being a dictator function. If such a test exists where the maximum probability of acceptance for far from dictator functions is at most α_P, then using the Rich 2-to-1 Conjecture of Braverman, Khot, and Minzer [7], one gets an **NP**-hardness of approximating P-CSP on satisfiable instances to within a factor of α_P.

We now describe the dictatorship test that we design for a large class of predicates. The starting point is an instance ϕ of P-CSP and let the value (i.e., maximum fraction of the constraints that can be satisfied by an assignment) of this instance be s. The distribution μ in the test depends on the SDP solution for ϕ and we only consider instances whose SDP value is 1. The SDP solution consists of vectors as well as local distribution for each constraint. Since the SDP value is 1, all these local distributions are supported on the satisfying assignments to P. Let μ_i be the local distribution corresponding to the i^{th} constraint of the instance. The test is as follows. Here $\varepsilon > 0$ is a small constant independent of n.

If f is a dictator function, then the test accepts with probability 1. This follows because for every i, the distribution μ_i is supported on the satsifying assignments to P and therefore every column of the matrix is from $P^{-1}(1)$. A challenging task is to compute the acceptance probability when f is far from dictator functions.

Given $f : \Sigma^n \to \Sigma$,

(1) Select a constraint from ϕ according to the weights of the constraints. Let i be the selected constraint.
(2) Construct a $k \times n$ matrix by setting each column of the matrix independently according to the following distribution: sample the column using μ_i.
(3) Check if $P(f(x_1), f(x_2), \ldots, f(x_k) = 1$.

Figure 1: Dictatorship test with completeness 1.

This test is a slight modification of Raghavendra's test [21]. In his test, in Step (2) with probability ε, a random sample is chosen from Σ^k. This uniform *noise* has an effect of killing all the *high-degree* monomials of f and hence the analysis boils down to only considering the low-degree functions. At this point, one can apply the invariance principle for low-degree functions from Mossel [19] and can replace the inputs with correlated Gaussians. Finally, the expression involving the Gaussians is interpreted as a rounding algorithm that rounds the SDP solution to an integral solution and the value is upper-bounded by the integral value of the instance which is s. Thus, the soundness of the test essentially matches the value of the integral solution. However, because of the uniform noise, the dictator functions will no longer pass the test with probability 1 and hence this test will not give hardness results on satisfiable instances.

Coming back to our test, we cannot add uniform noise as we want to maintain the completeness of the test to be 1. However, this introduces a few challenges in the analysis of the test. The main challenge is to show that the local distribution is enough to kill the high-degree part of f. This in general is not true. Specifically, if the predicate satisfies a linear equation, then this distribution is not enough to kill the high-degree part (see the counterexample in Remark 1.8). This is where we need the predicate (and the local distributions) to *not* satisfy any linear equation. In this case, we use our main analytical lemma, that we will discuss later, to show that the high-degree part of f contributes little to the test acceptance probability. However, we additionally need the predicate and the SDP solution to be semi-rich.

Finally, similar to Raghavendra's analysis, we use the low degree-part of f in the rounding algorithm and relate the performance of the algorithm to the test acceptance probability. This shows that if f is far from dictator function, then the acceptance probability of the test is upper bounded by the value of the assignment returned by the rounding procedure, which is always upper-bounded by s.

Main analytical lemma. Analyzing the acceptance probability of the test is a challenging task in general. One begins by thinking of the function f as a real valued function, e.g. as an indicator of the event that it takes a specific symbol in Σ as its value. Skipping some details, one needs to analyze expectations of the form

$$\mathop{\mathbf{E}}_{x_1,x_2,\ldots,x_k\sim\mu^{\otimes n}}\left[\prod_{i=1}^{k} f(x_i)\right],$$

here $x_1, x_2, \ldots, x_k$ are distributed as discussed in Definition 1.4. As the low-degree part of f corresponds to the SDPs from the algorithmic side, in order to prove our main theorem, we need to show that when f is a high-degree function, then this expectation is small. Our main analytical lemma shows that this is indeed the case. Following is the informal statement of the lemma (for a formal statement, see Lemma 2.6).

Lemma 1.5 (Informal). *Let P be any 3-ary predicate that is semi-rich and does not satisfy any linear embedding. Let μ be any distribution that is fully supported on $P^{-1}(1)$. Then for any high-degree bounded function f, we have*

$$\left|\mathop{\mathbf{E}}_{x_1,x_2,x_3\sim\mu^{\otimes n}}[f(x_1)f(x_2)f(x_3)]\right| \leqslant \delta,$$

where $\delta \to 0$ as the degree of the function increases.

We note that a high-degree function has $\mathbf{E}[f] \approx 0$. This lemma is proved in Section 3 and it is evident that the proof of this lemma is quite involved. We believe that the semi-rich condition is not needed for the conclusion to hold. Generalizing the lemma for k-ary predicates and proving it without the semi-rich condition is a fascinating analytical question for future work.

The lemma is a generalization of Lemma 6.2 by Mossel [19]. That lemma states that if the distribution μ is *connected* then the expectation is small. The connectedness condition can be stated as follows: For every pair of assignments (a, b, c) and (a', b', c') in $P^{-1}(1)$, there is a way to convert the first assignment to the second by replacing only once coordinate at a time such that every intermediate triple is in $P^{-1}(1)$. The predicate P_G that was mentioned earlier where $P_G^{-1}(1) = \{(x, y, z) \mid x \cdot y \cdot z = 1_G\}$ for some non-Abelian group does not satisfy the connectedness condition, as changing one coordinate from any satisfying assignment gives a triple which is outside of $P_G^{-1}(1)$. This predicate, however, does not satisfy any linear embedding if G does not have any non-trivial representation of dimension 1. P_G is also semi-rich and hence we can apply the above analytical lemma for P_G.

The proof of the above lemma for the predicate P_G is implicit in the work of Bhangale and Khot [4]. Given this fact, our high-level strategy to prove the lemma is as follows. We modify the underlying distribution μ and the predicate P so that the modified predicate can be viewed as a set of equations over some non-Abelian group. We do this by carefully adding more satisfying triplets to the predicate. During the modifications, we maintain the properties of the original predicate (i.e., semi-richness and not having any linear embedding) as well as make sure that the expectation does not change by much. Since the original predicate does not satisfy any linear embedding, the group must be non-Abelian and also lacks any non-trivial representation of dimension 1. Therefore, the final expectation must be small. This shows that the earlier expectation is also small.

1.3 Conclusion and Future Work

Our work leaves open many interesting problems. One obvious open problem is to extend our main theorem for other class of predicates. We could prove our analytical lemma for 3-ary semi-rich predicates. However, we believe that this semi-richness condition is not necessary for the conclusion to hold. One obvious open question is to extend our main theorem to other 3-ary predicates that are not semi-rich. More ambitiously, we put forth the following conjecture for general k-ary predicates. One can naturally extend the definition of 3-ary predicates not satisfying any linear equation to k-ary predicates as follows.

Definition 1.6. *Let $P : \Sigma^k \to \{0, 1\}$ be any k-ary predicate such that the support on each coordinate is full. We say P satisfies a linear embedding if there exists an Abelian group $(G, +)$ and mappings $\sigma_i : \Sigma \to G$ such that*

- *$\sum_{i=1}^k \sigma_i(x_i) = 0$ for every $(x_1, x_2, \ldots, x_k) \in P^{-1}(1)$, where 0 is the identity element of G.*
- *one of the mappings $\{\sigma_i\}_{i=1}^k$ is non-constant.*

Otherwise, we say P does not satisfy any linear embedding.

With this definition, we conjecture the following.

Conjecture 1.7 (Informal). *Let P be any k-ary predicate that does not satisfy any linear embedding. Let μ be any distribution that is fully supported on $P^{-1}(1)$. Given k functions $f_1, f_2, \ldots, f_k : \Sigma^n \to [-1, 1]$, such that one of the f_is is a high-degree function, then we have*

$$\left|\mathop{\mathbf{E}}_{x_1,x_2,\ldots,x_k\sim\mu^{\otimes n}}[f_1(x_1)f_2(x_2)\cdots f_k(x_k)]\right| \leqslant \delta,$$

where $\delta \to 0$ as the degree of the function increases.

Remark 1.8. *We note that if the predicate satisfies a linear equation, then the conclusion does not hold. To see this, suppose P satisfies a linear equation over an Abelian group G given by the embeddings $\{\sigma_i\}_{i=1}^k$. Let χ be any non-trivial character of G and define $f_i(x_i) = \prod_{j=1}^n \chi(\sigma_i((x_i)_j))$. Now,*

$$\begin{aligned} f_1(x_1)f_2(x_2)\cdots f_k(x_k) &= \prod_{i=1}^k\prod_{j=1}^n \chi(\sigma_i((x_i)_j)) \\ &= \prod_{j=1}^n\prod_{i=1}^k \chi(\sigma_i((x_i)_j)) \\ &= \prod_{j=1}^n \chi\left(\sum_{i=1}^k \sigma_i((x_i)_j)\right), \end{aligned}$$

where the last equality is because of the multiplicativity of the character χ. For every j, we have $\sum_{i=1}^k \sigma_i((x_i)_j) = 0$ and hence the product becomes 1 as $\chi(0) = 1$. Moreover, for large n, since one of the embeddings is non-constant, one of the f_is is a high-degree function.

With a positive answer to the conjecture, we may be able to make progress on predicates P that do not satisfy any linear equation. On the other hand, if P does satisfy a certain linear equation, then a hybrid algorithm that solves the SDP as well as the system of linear equations might give an optimal algorithm for satisfiable CSPs. We leave these as open problems for future work.

1.4 Organization

In Section 2 we state our main dictatorship test for 3-ary CSPs satisfying conditions from Theorem 1.1. We start with preliminaries in Section 2.1 where we define constraint satisfaction problems, functions on product spaces and state the invariance principle. We also state our main analytical lemma that we use in our dictatorship test analysis in this section. In Section 2.2 we define the basic SDP relaxation for a Max-P-CSP. In Section 2.3, we state our dictatorship test and prove the completeness and soundness analysis of the test. We give an overview of the proof of our main analytical lemma in Section 3. The proof consists of a series of steps which we prove in the full version of the paper.

2 FROM INTEGRALITY GAP TO DICTATORSHIP TEST

In this section, we show that if a P-CSP instance has a $(1, s)$ integrality gap for the basic SDP relaxation, then there is a dictatorship test with completeness 1 and soundness $s + \varepsilon$ for any $\varepsilon > 0$, if the predicate and the SDP solution satisfy certain conditions.

2.1 Preliminaries

The focus of this paper is on special type of predicates that do not satisfy any linear equation and that are semi-rich. We define these two properties next. We define these properties for more general predicates having different alphabet for each location, although in our dictatorship test we only consider predicates of the type $P : \Sigma^3 \to \{0, 1\}$.

Definition 2.1. *Let Σ, Φ, Γ be finite alphabets. Let $H \subseteq \Sigma \times \Phi \times \Gamma$ and $\Sigma' \subseteq \Sigma, \Phi' \subseteq \Phi$ and $\Gamma' \subseteq \Gamma$ be the subsets on which H is supported. We say H can be linearly embedded in an Abelian group if there is an Abelian group $(G, +)$ and maps $\sigma \colon \Sigma' \to G$, $\phi \colon \Phi' \to G$, $\gamma \colon \Gamma' \to G$ such that*

1. *for all $(x, y, z) \in H$ it holds that $\sigma(x) + \phi(y) + \gamma(z) = 0$;*
2. *at least one of σ, ϕ, γ is non-constant.*

Otherwise, we say H cannot be embedded linearly into an Abelian group, or simply H does not satisfy any linear equation.

Definition 2.2. *Let Σ, Φ, Γ be finite alphabets. Let $H \subseteq \Sigma \times \Phi \times \Gamma$ and $\Sigma' \subseteq \Sigma, \Phi' \subseteq \Phi$ and $\Gamma' \subseteq \Gamma$ be the subsets on which H is supported. We say H is* semi-rich *if the following two properties hold.*

1. *For all $(x, y) \in \Sigma' \times \Phi'$, there exists $z \in \Gamma'$ such that $(x, y, z) \in H$.*
2. *For all $(x, z) \in \Sigma' \times \Gamma'$, there exists $y \in \Phi'$ such that $(x, y, z) \in H$.*

Note that in the definition of semi-rich one of the three coordinates is special. However, the location of the special coordinate does not matter as we can permute the coordinates and study the modified subset instead.

We now define the predicates that have these two properties.

Definition 2.3. *A predicate $P \colon \Sigma \times \Phi \times \Gamma \to \{0, 1\}$ is said to be linearly embedded into an Abelian group if and only if $P^{-1}(1)$ can be linearly embedded in an Abelian group. P is called semi-rich if $P^{-1}(1)$ is semi-rich.*

Let (Ω, μ) be a probability space. Define the inner product on this space by $\langle f, g\rangle_\mu := \mathbf{E}_{x\in\mu}[f(x)g(x)]$. We will use the notation $\|f\|_{p;\mu} := \mathbf{E}_{x\in\mu}[|f(x)|^p]^{1/p}$ to denote the p^{th} norm of f. In order to state our main analytical lemma, we need the following definition of the noise operator.

Definition 2.4. *Let Φ be a finite alphabet, and ν be a measure on Φ. For a parameter $\rho \in [0, 1]$, we define the ρ-correlated distribution with respect to ν as follows. For any $y \in \Phi$, the distribution of inputs that are ρ-correlated with y is denoted by $\mathbf{y}' \sim \mathrm{T}_\rho y$ and is defined by taking $\mathbf{y}' = y$ with probability ρ, and otherwise sampling $\mathbf{y}' \sim \nu$.*

As is often the case, we also view T_ρ as an operator on functions, mapping $L^2(\Phi, \nu)$ to $L^2(\Phi, \nu)$ by

$$(\mathrm{T}_\rho g)(y) = \mathop{\mathbf{E}}_{\mathbf{y}'\sim \mathrm{T}_\rho y}\left[g(\mathbf{y}')\right].$$

We then tensorize this operator, i.e., consider $\mathrm{T}_\rho^{\otimes n}$ which acts on functions on n-variables, i.e. on $L^2(\Phi^n, \nu^{\otimes n})$. When clear from context, we drop the $\otimes n$ superscript from notation.

Definition 2.5. $\mathrm{Stab}_\rho^\nu(g) = \langle g, \mathrm{T}_\rho g\rangle_{\nu^{\otimes n}}$. *We drop the superscript ν from $\mathrm{Stab}_\rho^\nu(g)$, if it is clear from the context.*

Let $m = |\Phi|$ and write the multilinear expansion of g with respect to ν, i.e., $g(y) = \sum_{\sigma\in\{0,1,\dots,m-1\}^n} \widehat{g}(\boldsymbol{\sigma})\boldsymbol{\ell}_{\boldsymbol{\sigma}}(y)$, where $\boldsymbol{\ell}_0 \equiv 1$ is the trivial character.[9] Then $\mathrm{Stab}_\rho^\nu(g) = \sum_{\sigma\in\{0,1,\dots,m-1\}^n} \rho^{|\boldsymbol{\sigma}|}\widehat{g}(\boldsymbol{\sigma})^2$, where $|\boldsymbol{\sigma}|$ is the number of non-zero entries in $\boldsymbol{\sigma}$. Thus, if g has small weight on $\widehat{g}(\boldsymbol{\sigma})$ where $|\boldsymbol{\sigma}|$ is small then $\mathrm{Stab}_\rho^\nu(g)$ is small. Thus, we use the notion of small stability of a function as a proxy for high-degreeness of the function.

We now state the main analytical lemma that we use in the analysis of our dictatorship test. The lemma is proved in the next section (Section 3).

Lemma 2.6. *For all $m \in \mathbb{N}$, $\varepsilon, \alpha > 0$ there exist $\xi > 0$ and $\delta > 0$ such that the following holds. Suppose μ is a distribution over $\Sigma\times\Phi\times\Gamma$ whose support (a) is semi-rich, and (b) cannot be embedded in an Abelian group. Further suppose that $|\Sigma|, |\Phi|, |\Gamma| \leqslant m$, each atom in μ has probability at least α and marginals of μ on Σ, Φ and Γ have full support. If $f \colon \Sigma^n \to [-1, 1]$, $g \colon \Phi^n \to [-1, 1]$, $h \colon \Gamma^n \to [-1, 1]$ are functions such that*

either $\mathrm{Stab}_{1-\xi}(f) \leqslant \delta$, *or* $\mathrm{Stab}_{1-\xi}(g) \leqslant \delta$, *or* $\mathrm{Stab}_{1-\xi}(h) \leqslant \delta$,

then $\left|\mathbf{E}_{\mathbf{x},\mathbf{y},\mathbf{z}\sim\mu^{\otimes n}}\left[f(\mathbf{x})g(\mathbf{y})h(\mathbf{z})\right]\right| \leqslant \varepsilon$.

2.1.1 Constraint Satisfaction Problems.

We will use the notation $[R]$ to denote the set $\{1, 2, \dots, R\}$. In our dictatorship test analysis, we are going to need a few lemmas from Raghavendra's thesis [22] as black-box. Therefore, we try to use the same notations from his thesis. Our analytical lemma (Lemma 2.6) that we prove in the next section works only for the 3-ary CSPs. However, in this section, we work with general k-ary P-CSPs. If we have the analogous analytical lemma for any k-ary CSP, then the test designed in this section can be combined with it to get a result for k-ary CSPs.

Raghavendra considered CSPs with mixed predicates. In this work, we consider CSPs with one predicate $P : \Sigma^k \to \{0, 1\}$ (or possibly mix of predicates with the same template P, as described below). We formally define the P-CSP instance in the following definition.

Definition 2.7. *For a given predicate $P : \Sigma^k \to \{0, 1\}$, a P-CSP instance is given by $\mathfrak{I} = (\mathcal{V}, \mathcal{P})$ where*

- *$\mathcal{V}$ is the set of variables.*

[9] This is formally defined in the Section 2.1.2 below.

- *$\mathcal{P}$ is a probability distribution on the payoff functions $P' : \Sigma^{\mathcal{V}} \to \{0,1\}$ of type,*

$$P'(\boldsymbol{y}) = P(y_{i_1}, y_{i_2}, \ldots, y_{i_k}),$$

for some $i_1, i_2, \ldots, i_k \in \mathcal{V}$.

Remark 2.8. *A P-CSP instance actually consists of a mix of payoffs on the same template P. In the Boolean CSP, these mix of payoffs are formed by using literals (or negations). Here are few examples to illustrate this for Boolean CSPs as well as for general CSPs.*

1. *In 3SAT, the template predicate P is $P : \{0,1\}^3 \to \{0,1\}$ where $P(x,y,z) = 0$ iff $x = y = z = 0$. However, a 3SAT instance contains 8 different payoffs, one for each literal pattern.*
2. *In 3LIN, the template predicate $P : \{0,1\}^3 \to \{0,1\}$ is such that $P(x,y,z) = 1$ iff $x \oplus y \oplus z = 1$. In this case, the instance also contains constraints of type $x \oplus y \oplus z = 0$.*
3. *In 3LIN equations over a non Abelian group $(G, \cdot)$, the predicate is $P : G^3 \to \{0,1\}$ such that $P(x,y,z) = 1$ iff $x \cdot y \cdot z = 1_G$, where 1_G is the identity element of G. The instance contains constraints of type $x \cdot y \cdot z = g$ for some $g \in G$.*

Without the mix of payoffs, certain P-CSPs are trivial; for instance, the all 1 assignment would satisfy every 3SAT and 3LIN instance. Therefore we allow the use of such mix of payoffs in our instances. Note that for certain predicates, like $3NAE : \{0,1\}^3 \to \{0,1\}$, defined as $3NAE(x,y,z) = 1$ iff x, y, z are not all the same, instances without any mix of payoffs are non-trivial to solve.

For a payoff P', the set of indices $i_1, i_2, \ldots, i_k \in \mathcal{V}$ on which it depends is denoted by $\mathcal{V}(P')$. Let $\mathsf{supp}(\mathcal{P})$ be the set of payoffs in $\mathfrak{I}$. Given a P-CSP instance $\mathfrak{I}$, the objective is to find an assignment $\boldsymbol{y} \in \Sigma^{\mathcal{V}}$ that maximizes the value of the instance which is defined as follows:

$$\mathbf{val}(\boldsymbol{y}) = \mathop{\mathbf{E}}_{P' \sim \mathcal{P}}[P'(\boldsymbol{y})].$$

The optimum value of the instance $\mathfrak{I} = (\mathcal{V}, \mathcal{P})$ is defined as:

$$\textsc{Opt}(\mathfrak{I}) = \max_{\boldsymbol{y} \in \Sigma^{\mathcal{V}}} \mathbf{val}(\boldsymbol{y}).$$

Let $\blacktriangle(\Sigma)$ be the set of probability distributions on Σ.

2.1.2 Functions on Product Spaces. Let (Ω, μ) be a probability space with $|\Omega| = q$ and μ has full support on Ω. Define the inner product between two functions $f, g : \Omega \to \mathbb{R}$ on this space as follows: $\langle f, g \rangle = \mathbf{E}_{x \sim \mu}[f(x)g(x)]$.

Definition 2.9. *An orthonormal ensemble consists of a basis of real orthonormal random variables $\mathcal{L} = \{\ell_0 \equiv \mathbf{1}, \ell_1, \ldots, \ell_{q-1}\}$, where $\mathbf{1}$ is the constant 1 function.*

Henceforth, we will sometimes refer to orthonormal ensembles as just ensembles. For an ensemble $\mathcal{L} = \{\ell_0 \equiv \mathbf{1}, \ell_1, \ldots, \ell_{q-1}\}$ of random variables, we will use $\mathcal{L}^R$ to denote the ensemble obtained by taking R independent copies of $\mathcal{L}$. Furthermore, $\mathcal{L}^{(i)} = \{\ell_0^{(i)}, \ell_1^{(i)}, \ldots, \ell_{q-1}^{(i)}\}$ will denote the i^{th} independent copy of $\mathcal{L}$.

Fix an ensemble $\mathcal{L} = \{\ell_0 \equiv \mathbf{1}, \ell_1, \ldots, \ell_{q-1}\}$ that forms a basis for $L^2(\Omega)$. Given such a basis for $L^2(\Omega)$, it induces a basis for the space $L^2(\Omega^R)$, given by the random variables

$$\left\{ \ell_{\boldsymbol{\sigma}} := \prod_{i=1}^{R} \ell_{\sigma_i}^{(i)} \;\middle|\; \boldsymbol{\sigma} \in \{0, 1, \ldots, q-1\}^R \right\}.$$

Therefore, any function $\mathcal{F} : \Omega^R \to \mathbb{R}$ has a multilinear expansion

$$\mathcal{F}(\boldsymbol{z}) = \sum_{\boldsymbol{\sigma} \in \{0,1,\ldots,q-1\}^R} \hat{\mathcal{F}}(\boldsymbol{\sigma}) \ell_{\boldsymbol{\sigma}}(\boldsymbol{z}),$$

where $\ell_{\boldsymbol{\sigma}}(\boldsymbol{z}) = \prod_{i=1}^{R} \ell_{\sigma_i}(z_i)$.

Definition 2.10. *A multi-index $\boldsymbol{\sigma}$ is a vector $(\sigma_1, \sigma_2, \ldots \sigma_R) \in \{0, 1, \ldots, q-1\}^R$ and the degree of $\boldsymbol{\sigma}$ is denoted by $|\boldsymbol{\sigma}|$ which is equal to $|\boldsymbol{\sigma}| = |\{i \in [R] \mid \sigma_i \neq 0\}|$. Given a set of indeterminates $\boldsymbol{\mathcal{X}} = \{x_j^{(i)} | j \in \{0, 1, \ldots, q-1\}, i \in [R]\}$ and a multi-index $\boldsymbol{\sigma}$, define the monomial $x_{\boldsymbol{\sigma}}$ as*

$$x_{\boldsymbol{\sigma}} = \prod_{i=1}^{R} x_{\sigma_i}^{(i)}.$$

The degree of the monomial is given by $|\boldsymbol{\sigma}|$. A multilinear polynomial over such indeterminates is given by

$$F(\boldsymbol{x}) = \sum_{\boldsymbol{\sigma} \in \{0,1,\ldots,q-1\}^R} \hat{F}_{\boldsymbol{\sigma}} x_{\boldsymbol{\sigma}}.$$

Given any function $\mathcal{F} : \Omega^R \to \mathbb{R}$, with the multilinear expansion $\mathcal{F}(\boldsymbol{z}) = \sum_{\sigma \in \{0,1,\ldots,q-1\}^R} \hat{\mathcal{F}}(\sigma) \ell_{\sigma}(\boldsymbol{z})$ with respect to the orthonormal ensemble $\mathcal{L} = \{\ell_0 \equiv \mathbf{1}, \ell_1, \ldots, \ell_{q-1}\}$, we define a corresponding *formal polynomial* in the indeterminates $\boldsymbol{\mathcal{X}} = \{x_j^{(i)} | j \in \{0, 1, \ldots, q-1\}, i \in [R]\}$, as follows:

$$F(\boldsymbol{x}) = \sum_{\boldsymbol{\sigma}} \hat{\mathcal{F}}(\boldsymbol{\sigma}) x_{\boldsymbol{\sigma}}.$$

We will always use the symbol $\mathcal{F}$ to denote real-valued function on a product probability space Ω^R. Further $F(\boldsymbol{x})$ will denote the formal multilinear polynomial corresponding to $\mathcal{F}$. Hence $F(\mathcal{L}^R)$ is a random variable obtained by substituting the random variables $\mathcal{L}^R$ in place of $\boldsymbol{x}$. For instance, the following equation holds in this notation:

$$\mathop{\mathbf{E}}_{\boldsymbol{z} \in \Omega^R}[\mathcal{F}(\boldsymbol{z})] = \mathbf{E}[F(\mathcal{L}^R)].$$

We now define the notion of the influence of a variable.

Definition 2.11. *For a function $\mathcal{F} : \Omega^R \to \mathbb{R}$ over the space $(\Omega^R, \mu^{\otimes R})$, the influence of the j^{th} coordinate is given by:*

$$\mathbf{Inf}_j[\mathcal{F}; \mu^{\otimes R}] = \mathop{\mathbf{E}}_{\boldsymbol{z}^{(-j)} \in \Omega^{R-1}}[\mathbf{Var}_{\boldsymbol{z}^{(j)} \in \Omega}[\mathcal{F}(\boldsymbol{z})]],$$

where $\boldsymbol{z}^{(-j)}$ is a string missing the j^{th} coordinate.

We have the following proposition that relates the average value and the variance of a function to its Fourier coefficients.

Proposition 2.12. *For a function $\mathcal{F} : \Omega^R \to \mathbb{R}$ over the space $(\Omega^R, \mu^{\otimes R})$, if $\mathcal{F}(\boldsymbol{z}) = \sum_{\boldsymbol{\sigma}} \hat{\mathcal{F}}(\boldsymbol{\sigma}) \ell_{\boldsymbol{\sigma}}(\boldsymbol{z})$ with respect to an orthonormal ensemble $\mathcal{L}$ of (Ω, μ), then $\mathbf{E}_{\boldsymbol{z} \in \Omega^R}[\mathcal{F}(\boldsymbol{z})] = \hat{\mathcal{F}}_{\mathbf{0}}$ and $\mathbf{Var}[\mathcal{F}] = \sum_{\boldsymbol{\sigma} \neq \mathbf{0}} \hat{\mathcal{F}}_{\boldsymbol{\sigma}}^2$.*

We also define the degree $\geqslant D$ weight of a function $\mathcal{F}$ as follows:

$$W^{\geqslant D}[\mathcal{F}; \mu^{\otimes R}] = \sum_{\boldsymbol{\sigma} : |\boldsymbol{\sigma}| \geqslant D} \hat{\mathcal{F}}_{\boldsymbol{\sigma}}^2.$$

Another way of writing a function on a probability space as sum of orthogonal functions is called the Efron-Stein decomposition.

Definition 2.13. *Let (Ω, μ) be a probability space and $(\Omega^R, \mu^{\otimes R})$ be the corresponding product space. For a function $f : \Omega^R \to \mathbb{R}$, the Efron-Stein decomposition of f with respect to the product space is given by*

$$f(z_1, \cdots, z_R) = \sum_{\beta \subseteq [R]} f_\beta(z),$$

where f_β depends only on z_i for $i \in \beta$ and for all $\beta' \not\supseteq \beta, \boldsymbol{a} \in \Omega^{\beta'}$, $\mathbf{E}_{z \in \mu^{\otimes R}} \left[f_\beta(z) \mid z|_{\beta'} = \boldsymbol{a} \right] = 0$.

We have the following facts about the Efron-Stein decomposition of functions.

Fact 2.14. *If $f(z) = \sum_{\beta \subseteq [R]} f_\beta(z)$ and $g(z) = \sum_{\beta \subseteq [R]} g_\beta(z)$ are the Efron-Stein decompositions of f and g respectively w.r.t. the product space $(\Omega^R, \mu^{\otimes R})$, then*

(1) *$T_\rho f(z) = \sum_{\beta \subseteq [R]} \rho^{|\beta|} f_\beta(z)$ and*
(2) *$\langle f, g \rangle_{\mu^{\otimes R}} = \sum_{\beta \subseteq [R]} \langle f_\beta, g_\beta \rangle_{\mu^{\otimes R}}$.*

2.1.3 Vector valued functions. We will always use the symbol $\mathcal{F} = (\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_q)$ to denote a vector-valued function on a product probability space Ω^R. Further, $\boldsymbol{F}(\boldsymbol{x}) = (F_1, F_2, \ldots, F_q)$ will denote the formal multilinear polynomial corresponding to $\mathcal{F}$.

The notions of influence and degree $\geqslant D$ weight can be extended to the vector valued functions using the following definitions.

$$\mathbf{Inf}_i[\mathcal{F}; \mu^{\otimes R}] = \sum_{j=1}^{q} \mathbf{Inf}_i[\mathcal{F}_j; \mu^{\otimes R}]$$

and

$$W^{\geqslant D}[\mathcal{F}; \mu^{\otimes R}] = \sum_{j=1}^{q} W^{\geqslant D}[\mathcal{F}_j; \mu^{\otimes R}].$$

2.1.4 Invariance Principle. Define functions $f_{[0,1]} : \mathbb{R} \to \mathbb{R}$ and $\xi : \mathbb{R}^q \to \mathbb{R}$ as follows:

$$f_{[0,1]}(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \leqslant x \leqslant 1, \\ 1 & \text{if } x > 1. \end{cases}$$

$$\xi(\boldsymbol{a}) = \sum_{j=1}^{q} (f_{[0,1]}(a_j) - a_j)^2.$$

A crucial step in the analysis of the dictatorship test is to replace the discrete inputs with correlated Gaussians. The following theorem from Mossel [19] states that one can do this provided the functions do not have influential coordinates and the functions are low-degree.

Theorem 2.15 ([19]). *Fix $0 < \alpha \leqslant 1/2$ and $d \in \mathbb{N}$. Let (Ω, μ), $|\Omega| = m$, be a finite probability space such that every atom has probability at least α. Let $\mathcal{L}^{(r)} = \{\ell_0^{(r)} \equiv \mathbf{1}, \ell_1^{(r)}, \ldots, \ell_{m-1}^{(r)}\}$ be an orthonormal ensemble of random variables over Ω and $\mathcal{G}^{(r)} = \{g_0^{(r)} \equiv \mathbf{1}, g_1^{(r)}, \ldots, g_{m-1}^{(r)}\}$ be an orthonormal ensemble of Gaussian random variables.*

Let $\boldsymbol{F} = (F_1, F_2, \ldots, F_d)$ denote a vector valued multilinear polynomial on Ω^R. If $\mathbf{Inf}_i[\boldsymbol{F}; \mu^{\otimes R}] \leqslant \tau$ for all $i \in [R]$, $W^{\geqslant D}[\boldsymbol{F}; \mu^{\otimes R}] \leqslant \delta$ and $\mathbf{Var}[F_j] \leqslant 1$ for all $j \in \{1, \ldots, d\}$, then the following holds.

(1) *For every function $\psi : \mathbb{R}^d \to \mathbb{R}$ that is thrice differentiable with all its partial derivatives up to order 3 bounded uniformly by C_0,*

$$\left| \mathbf{E}[\psi(\boldsymbol{F}(\mathcal{L}^R))] - \mathbf{E}[\psi(\boldsymbol{F}(\mathcal{G}^R))] \right| \leqslant O\left(D\sqrt{\tau} \left(8\alpha^{-1/2} \right)^D \right) + O(\sqrt{\delta}).$$

(2) *For the function ξ defined above,*

$$\left| \mathbf{E}[\xi(\boldsymbol{F}(\mathcal{L}^R))] - \mathbf{E}[\xi(\boldsymbol{F}(\mathcal{G}^R))] \right| \leqslant O\left(\sqrt{\tau} \left(10\alpha^{-1/2} \right)^D \right)^{2/3} + O(\sqrt{\delta}).$$

In both the cases, the $O(.)$ hides the constant C_0.

Proof. The theorem follows from Theorem 4.1 from [19]. Truncate the polynomial $\boldsymbol{F}$ to degree D to get a polynomial $\boldsymbol{L}$. Using Theorem 4.1 of Mossel [19], we have

$$\begin{aligned} \left| \mathbf{E}[\psi(\boldsymbol{L}(\mathcal{L}^R))] - \mathbf{E}[\psi(\boldsymbol{L}(\mathcal{G}^R))] \right| &\leqslant 2DC_0 d^3 (8\alpha^{-1/2})^D \tau^{1/2} \\ &= O\left(D\sqrt{\tau} \left(8\alpha^{-1/2} \right)^D \right). \end{aligned}$$

Since ψ is a smooth functional,

$$\begin{aligned} \left| \mathbf{E}[\psi(\boldsymbol{L}(\mathcal{L}^R))] - \mathbf{E}[\psi(\boldsymbol{F}(\mathcal{L}^R))] \right| &\leqslant C_0 \|\boldsymbol{L}(\mathcal{L}^R) - \boldsymbol{F}(\mathcal{L}^R)\| \\ &= C_0 \left(W^{\geqslant D}[\boldsymbol{F}; \mu^{\otimes R}] \right)^{1/2} \\ &\leqslant C_0 \sqrt{\delta}. \end{aligned}$$

Similarly, we get

$$\left| \mathbf{E}[\psi(\boldsymbol{L}(\mathcal{G}^R))] - \mathbf{E}[\psi(\boldsymbol{F}(\mathcal{G}^R))] \right| \leqslant C_0 \sqrt{\delta}.$$

Combining the three inequalities, we get the required bound for (1).

The second item follows from Theorem 3.19 from [20]. Here again, let $\boldsymbol{L}$ be the low-degree part of $\boldsymbol{F}$ truncated at degree D and let $\boldsymbol{H} = \boldsymbol{F} - \boldsymbol{L}$. Using Theorem 3.19 of [20],

$$\left| \mathbf{E}[\xi(\boldsymbol{L}(\mathcal{L}^R))] - \mathbf{E}[\xi(\boldsymbol{L}(\mathcal{G}^R))] \right| \leqslant O\left(\sqrt{\tau} \left(10\alpha^{-1/2} \right)^D \right)^{2/3}.$$

Using Lemma 3.24 from [20],

$$\begin{aligned} &\left| \mathbf{E}[\xi(\boldsymbol{F}(\mathcal{L}^R))] - \mathbf{E}[\xi(\boldsymbol{L}(\mathcal{L}^R))] \right| \\ &\leqslant 2\,\mathbf{E}[\boldsymbol{L}(\mathcal{L}^R)\boldsymbol{H}(\mathcal{L}^R)] + \mathbf{E}[\boldsymbol{H}(\mathcal{L}^R)^2] \\ &\leqslant 2\sqrt{\mathbf{E}[\boldsymbol{L}(\mathcal{L}^R)^2]}\sqrt{\mathbf{E}[\boldsymbol{H}(\mathcal{L}^R)^2]} + \mathbf{E}[\boldsymbol{H}(\mathcal{L}^R)^2] \\ &\leqslant 2\sqrt{\mathbf{E}[\boldsymbol{H}(\mathcal{L}^R)^2]} + \mathbf{E}[\boldsymbol{H}(\mathcal{L}^R)^2] \leqslant 2\sqrt{\delta} + \delta \leqslant 3\sqrt{\delta}, \end{aligned}$$

where the second step follows from the Cauchy-Schwarz inequality. Similarly, we get,

$$\left| \mathbf{E}[\xi(\boldsymbol{F}(\mathcal{G}^R))] - \mathbf{E}[\xi(\boldsymbol{L}(\mathcal{G}^R))] \right| \leqslant 3\sqrt{\delta},$$

and the claim follows. ◻

2.2 SDP Relaxation

Given an instance $\mathfrak{I} = (\mathcal{V}, \mathcal{P})$, the basic semi-definite programming relaxation of the instance is given in Figure 2. It consists of vectors $\{\boldsymbol{b}_{i,a}\}_{i\in\mathcal{V},a\in\Sigma}$, distributions $\{\mu_{P'}\}_{P'\in\mathsf{supp}(\mathcal{P})}$ over the local assignments (i.e., on $\Sigma^{\mathcal{V}(P')}$) and a unit vector $\boldsymbol{b}_0$. Let $\mathbf{val}(\boldsymbol{V}, \boldsymbol{\mu})$ be the objective value of the solution $(\boldsymbol{V}, \boldsymbol{\mu})$.

maximize

$$\mathop{\mathbf{E}}_{P'\sim\mathcal{P}} \mathop{\mathbf{E}}_{x\in\mu_{P'}} [P'(x)] \quad (1)$$

subject to

$$\langle \boldsymbol{b}_{i,a}, \boldsymbol{b}_{j,b}\rangle \quad (2)$$

$$= \Pr_{x\sim\mu_{P'}}[x_i = a, x_j = b] \qquad P' \in \mathsf{supp}(\mathcal{P}),\ i,j \in \mathcal{V}(P'),\ a,b\in\Sigma \quad (3)$$

$$\langle \boldsymbol{b}_{i,a}, \boldsymbol{b}_0\rangle = \|\boldsymbol{b}_{i,a}\|_2^2 \qquad \forall i \in \mathcal{V}, a \in \Sigma \quad (4)$$

$$\|\boldsymbol{b}_0\|_2^2 = 1 \quad (5)$$

$$\mu_{P'} \in \blacktriangle(\Sigma^{\mathcal{V}(P')}) \qquad P' \in \mathsf{supp}(\mathcal{P}) \quad (6)$$

Figure 2: Basic SDP relaxation of a P-CSP instance $\mathfrak{I} = (\mathcal{V}, \mathcal{P})$.

Following is a definition of $(1, s)$ integrality gap instance.

Definition 2.16. *An instance $\mathfrak{I} = (\mathcal{V}, \mathcal{P})$ is a $(1, s)$ SDP integrality gap instance if the optimal value of the instance is at most s and the optimal value of the basic SDP relaxation for $\mathfrak{I}$ is 1.*

For our dictatorship test to work, we require that the support of every local distribution $\mu_{P'}$ is semi-rich and it is not linearly embeddable in any Abelian group. Henceforth, we will assume that the SDP solution satisfies this property.

2.3 Dictatorship Test

In this section, we study the dictatorship test for P-CSP instances over a k-ary predicate P. Throughout this section, when $k = 3$, we restrict ourselves to the predicates P that are semi-rich and that do not satisfy any linear equation.

Let $\mathfrak{I} = (\mathcal{V}, \mathcal{P})$ be an instance of P-CSP, where $P : \Sigma^k \to \{0, 1\}$ and $|\Sigma| = q$. We will fix an arbitrary mapping from Σ to $\{1, 2, \ldots, q\}$, denoted by $\varsigma : \Sigma \to \{1, 2, \ldots, q\}$.

Let $(\boldsymbol{V}, \boldsymbol{\mu})$ be a solution for the basic SDP relaxation of $\mathfrak{I}$ which is semi-rich and which does not satisfy any linear equation. For each $s \in \mathcal{V}$, let $\Omega_s = (\Sigma, \mu_s)$ be a probability space with atoms in Σ where the probability of $a \in \Sigma$ is $\|\boldsymbol{b}_{s,a}\|_2^2$. We assume that Ω_s has full support for every $s \in \mathcal{V}$. However, our proof works even when the support is a subset of Σ.

A function $F : \Sigma^R \to \Sigma$ is called a dictator function if $F(\boldsymbol{z}) = \boldsymbol{z}^{(i)}$ for some $i \in [R]$. In Figure 3, we give the dictatorship test $\mathbf{Dict}_{\boldsymbol{V},\boldsymbol{\mu}}$ for functions $F : \Sigma^R \to \Sigma$.

Remark 2.17. *There is one main difference between our test and the dictatorship test given in [22]. In [22], in Step 2 (Figure 3), uniformly random noise is added from Σ^k. This step loses the perfect completeness of the dictatorship test.*

(1) Sample a payoff $P' \sim \mathcal{P}$. Let $\mathcal{V}(P') = \{s_1, s_2, \ldots, s_k\}$.
(2) Sample $\boldsymbol{z}_{P'} = \{\boldsymbol{z}_{s_1}, \boldsymbol{z}_{s_2}, \ldots, \boldsymbol{z}_{s_k}\}$ from the product distribution $\mu_{P'}^{\otimes R}$, i.e., independently for each $i \in [R]$, $(z_{s_1}^{(i)}, z_{s_2}^{(i)}, \ldots, z_{s_k}^{(i)}) \sim \mu_{P'}$.
(3) Query the function values $F(\boldsymbol{z}_{s_1}), F(\boldsymbol{z}_{s_2}), \ldots, F(\boldsymbol{z}_{s_k})$.
(4) Accept iff $P'(F(\boldsymbol{z}_{s_1}), F(\boldsymbol{z}_{s_2}), \ldots, F(\boldsymbol{z}_{s_k})) = 1$.

Figure 3: SDP integrality gap to a dictatorship test $\mathbf{Dict}_{\boldsymbol{V},\boldsymbol{\mu}}$.

2.3.1 Completeness Analysis. The completeness of the test is defined as follows,

$$\text{Completeness}(\mathbf{Dict}_{\boldsymbol{V},\boldsymbol{\mu}}) = \min_{\substack{i\in[R],\\ F \text{ is the } i^{th} \text{ dictator}}} \mathbf{Pr}[F \text{ passes } \mathbf{Dict}_{\boldsymbol{V},\boldsymbol{\mu}}].$$

If the function is a dictator function, then the test accepts with probability 1. The simple claim is proven below.

Lemma 2.18. *If $\mathbf{val}(\boldsymbol{V}, \boldsymbol{\mu}) = 1$ then*

$$\text{Completeness}(\mathbf{Dict}_{\boldsymbol{V},\boldsymbol{\mu}}) = 1.$$

Proof. Consider a dictator function $F(\boldsymbol{z}) = \boldsymbol{z}^{(j)}$ for some $j \in [R]$. In this case, $(F(\boldsymbol{z}_{s_1}), F(\boldsymbol{z}_{s_2}), \ldots, F(\boldsymbol{z}_{s_k})) = (z_{s_1}^{(j)}, z_{s_2}^{(j)}, \ldots, z_{s_k}^{(j)})$. When the payoff $P' \sim \mathcal{P}$ is selected, then $(z_{s_1}^{(j)}, z_{s_2}^{(j)}, \ldots, z_{s_k}^{(j)})$ is distributed according to $\mu_{P'}$. As the SDP value is 1, the distribution $\mu_{P'}$ is fully supported on $P'^{-1}(1)$ and hence the test passes with probability 1. □

2.3.2 Soundness Analysis. We now move to prove the soundness analysis of the test. Here we formally define the functions which are far from dictator functions (also known as quasirandom functions). Let $\Delta_q := \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_q\}$ where $\boldsymbol{e}_j$ is the j^{th} basis vector of $\mathbb{R}^q$.

Definition 2.19. *For a function $F : \Sigma^R \to \Sigma$, the corresponding Δ_q-representation is a function $\mathcal{F} : \Sigma^R \to \Delta_q$ given by*

$$\mathcal{F}(\boldsymbol{z}) = \boldsymbol{e}_{\varsigma(F(\boldsymbol{z}))}.$$

Therefore, in this setting F is a dictator function if $\mathcal{F}(\boldsymbol{z}) = \boldsymbol{e}_{\varsigma(\boldsymbol{z}^{(i)})}$ for some $i \in [R]$. Any function $\mathcal{F} : \Sigma^R \to \blacktriangle_q$ can be interpreted as a distribution on functions $\mathcal{F}' : \Sigma^R \to \Delta_q$ as follows: For each $\boldsymbol{z} \in \Sigma^R$, set the value of $\mathcal{F}'(\boldsymbol{z})$ independently as

$$\mathcal{F}'(\boldsymbol{z}) = \boldsymbol{e}_j \quad \text{with probability } \mathcal{F}(\boldsymbol{z})_j \text{ for all } j \in \{1, 2, \ldots, q\}.$$

Thus, for each $\boldsymbol{z} \in \Sigma^R$, we have $\mathcal{F}(\boldsymbol{z}) = \mathbf{E}[\mathcal{F}'(\boldsymbol{z})]$.

Fix a function $\mathcal{F} : \Sigma^R \to \blacktriangle_q$. For each $s \in \mathcal{V}$, let $\mathcal{F}_s$ denote the function $\mathcal{F}$ interpreted as a function on the product probability space $(\Sigma^R, \mu_s^{\otimes R})$.

Definition 2.20. *A function $\mathcal{F} : \Sigma^R \to \blacktriangle_q$ is said to be (τ, δ)-quasirandom if for each $s \in \mathcal{V}$, it holds that*

$$\max_{1\leqslant i\leqslant R} \mathbf{Inf}_i[T_{1-\delta}\mathcal{F}_s; \mu_s^{\otimes R}] \leqslant \tau,$$

where $\mathbf{Inf}_i[\mathcal{F}_s; \mu_s^{\otimes R}] = \sum_{j=1}^q \mathbf{Inf}_i[\mathcal{F}_{s,j}; \mu_s^{\otimes R}]$ and $\mathcal{F}_{s,j}$ is the map $\mathcal{F}_s$ restricted to the j^{th}-coordinate of $\blacktriangle_q$.

The domain of payoff P' can be extended from Σ^k to $\blacktriangle_q^k$. To see this, by the abuse of notation, first define a Δ_q-representation of a payoff $P' : \Sigma^k \to \{0,1\}$ as $P' : \Delta_q^k \to \{0,1\}$ where

$$P'(\boldsymbol{e}_{a_1}, \boldsymbol{e}_{a_2}, \ldots, \boldsymbol{e}_{a_k}) = P'(\varsigma^{-1}(a_1), \varsigma^{-1}(a_2), \ldots, \varsigma^{-1}(a_k)),$$

for all $(a_1, a_2, \ldots, a_k) \in \{1, 2, \ldots, q\}^k$.

The function P' can be extended to the domain $\blacktriangle_q^k$ by its multilinear extension. Again, by abusing the notation, define the extension P' as:

$$P'(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k) = \sum_{\sigma \in \Sigma^k} P'(\sigma) \prod_{i=1}^{k} x_{i,\varsigma(\sigma_i)}, \tag{7}$$

for all $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \blacktriangle_q$.

Define the soundness of the test as:

$$\text{Soundness}(\mathbf{Dict}_{\boldsymbol{V},\boldsymbol{\mu}}) = \sup_{\substack{\mathcal{F}:\Sigma^R \to \blacktriangle_q \\ \mathcal{F} \text{ is } (\tau,\delta)\text{–quasirandom w.r.t.}(\boldsymbol{V},\boldsymbol{\mu})}} \mathbf{Pr}[\mathcal{F} \text{ passes } \mathbf{Dict}_{\boldsymbol{V},\boldsymbol{\mu}}].$$

Extending P' to $\mathbb{R}^{qk}$: We will extend the payoff function P' further to a real valued function on $(\mathbb{R}^q)^k$, by plugging the real values in the expansion of P' given in the Equation (7). This extension of P' is smooth in the following sense:

1. All the partial derivatives of P' up to order 3 are uniformly bounded by $C_0(q,k)$.
2. P' is a Lipschitz function with Lipschitz constant $C_0(q,k)$, i.e.,$\forall \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}, \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k\} \in (\mathbb{R}^q)^k$,

$$|P'(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) - P'(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k)| \leqslant C_0(q,k) \sum_{i=1}^{k} \|\boldsymbol{x}_i - \boldsymbol{y}_i\|_2.$$

Setting of parameters. Let $\xi > 0$ be the parameter from Lemma 2.6. Let $\delta > 0$ be a sufficiently small constant. Set $\eta \in (0,1)$ to be the smallest constant such that for all $\ell \geqslant 0$,

$$(1-\xi)^\ell (1-(1-\delta)^\ell)^2 \leqslant \eta.$$

Note that as $\delta \to 0$, $\eta(\delta) \to 0$. We will denote the smallest non-zero probability of an atom in the SDP local distribution by α. As the SDP instance is finite, we can assume that $\alpha > 0$ independent of R.

Local and Global Ensembles. Fix a given SDP solution $(\boldsymbol{V}, \boldsymbol{\mu})$ with value 1. We define the following local and global orthonormal ensembles of random variables for every $s \in \mathcal{V}$ as follows.

- **Local Integral Ensembles $\mathcal{L}$:** The Local Integral Ensemble $\mathcal{L} = \{\boldsymbol{\ell}_s \mid s \in \mathcal{V}\}$ for a variable $s \in \mathcal{V}$, $\boldsymbol{\ell}_s = \{\ell_{s,0} \equiv \mathbf{1}, \ell_{s,1}, \ldots, \ell_{s,q-1}\}$ is a set of random variables that are orthonormal ensembles for the space Ω_s.

We also define the following global ensembles of random variables:

- **Global Gaussian Ensembles $\mathcal{G}$:** The Global Gaussian Ensembles $\mathcal{G} = \{\boldsymbol{g}_s \mid s \in \mathcal{V}\}$ are generated by setting $\boldsymbol{g}_s = \{g_{s,0} \equiv \mathbf{1}, g_{s,1}, \ldots, g_{s,q-1}\}$ where

$$g_{s,c} = \sum_{\omega \in \Sigma} \ell_{s,c}(\omega) \langle \boldsymbol{b}_{s,\omega}, \boldsymbol{\zeta} \rangle, \quad \forall c \in \{1, \ldots, q-1\},$$

and $\boldsymbol{\zeta}$ is a normal Gaussian random vector of appropriate dimension.

Input: An SDP solution $(\boldsymbol{V}, \boldsymbol{\mu})$.
Setup: For each $s \in \mathcal{V}$, the probability space $\Omega_s = (\Sigma, \mu_s)$ consists of atoms in Σ with the distribution $\mu_s(a) = \|b_{s,a}\|^2$. Let $\mathcal{F}_s$ denote the function obtained by interpreting the function $\mathcal{F} : \Sigma^R \to \blacktriangle_q$ as a function over Ω_s^R. Let $\mathcal{H}_s = T_{1-\delta}\mathcal{F}_s$ for all $s \in \mathcal{V}$. Let $\boldsymbol{F}_s, \boldsymbol{H}_s$ denote the multilinear polynomials corresponding to functions $\mathcal{F}_s, \mathcal{H}_s$ respectively.
Rounding Scheme:
Step I: Sample R Gaussian vectors $\boldsymbol{\zeta}^{(1)}, \boldsymbol{\zeta}^{(2)}, \ldots, \boldsymbol{\zeta}^{(R)}$ with the same dimension as $\boldsymbol{V}$.
Step II: For each $s \in \mathcal{V}$, do the following:

(1) For each $j \in [R]$, let $g_{s,0}^{(j)} \equiv \mathbf{1}$ and for $c \in \{1, \ldots, q-1\}$, set
$$g_{s,c}^{(j)} = \sum_{\omega \in \Sigma} \ell_{s,c}(\omega) \langle \boldsymbol{b}_{s,\omega}, \boldsymbol{\zeta}^{(j)} \rangle.$$
Let $\boldsymbol{g}_s^{(j)} = (g_{s,0}^{(j)} \equiv \mathbf{1}, g_{s,1}^{(j)}, \ldots, g_{s,q-1}^{(j)})$ and $\boldsymbol{g}_s = (\boldsymbol{g}_s^{(1)}, \boldsymbol{g}_s^{(2)}, \ldots, \boldsymbol{g}_s^{(R)})$.

(2) Evaluate the multilinear polynomial $\boldsymbol{H}_s$ with $\boldsymbol{g}_s$ as inputs to obtain $\boldsymbol{p}_s \in \mathbb{R}^q$, i.e., $\boldsymbol{p}_s = \boldsymbol{H}_s(\boldsymbol{g}_s)$.

(3) Round $\boldsymbol{p}_s$ to $\boldsymbol{p}_s^*$.
$$\boldsymbol{p}_s^* = \mathbf{Scale}(f_{[0,1]}((\boldsymbol{p}_s)_1), f_{[0,1]}((\boldsymbol{p}_s)_2), \ldots, f_{[0,1]}((\boldsymbol{p}_s)_q)),$$
where
$$f_{[0,1]}(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \leqslant x \leqslant 1, \\ 1 & \text{if } x > 1, \end{cases}$$
and
$$\mathbf{Scale}(x_1, x_2, \ldots, x_q) = \begin{cases} \frac{1}{\sum_i x_i}(x_1, \ldots, x_q) & \text{if } \sum_i x_i \neq 0, \\ (1, 0, 0, \ldots, 0) & \text{if } \sum_i x_i = 0. \end{cases}$$

(4) Assign the variable $s \in \mathcal{V}$ a value $a \in \Sigma$ with probability $(\boldsymbol{p}_s^*)_{\varsigma^{-1}(a)}$.

Step III: Output the assignment from Step II.

Figure 4: Rounding Scheme Round$_{\mathcal{F}}$.

The following lemma states that the local integral ensemble and the global Gaussian ensemble have matching first and second moments. We need this to apply the invariance principle in our analysis below.

Lemma 2.21. *For every $s \in \mathcal{V}$, $\boldsymbol{g}_s$ is an orthonormal ensemble w.r.t. the space Ω_s. Also, for any payoff $P' \in \mathcal{P}$, the global ensembles $\mathcal{G}$ match the following moments of the local integral ensembles $\mathcal{L}$:*

$$\mathop{\mathbf{E}}_{\boldsymbol{\zeta}}[g_{s,c} \cdot g_{s',c'}] = \mathop{\mathbf{E}}_{(\omega,\omega') \sim \mu_{P'}|(s,s')}[\ell_{s,c}(\omega).\ell_{s',c'}(\omega')]$$

for all $c, c' \in \{1, \ldots, q-1\}, s, s' \in \mathcal{V}(P')$, where $\mu_{P'}|(s,s')$ is the marginal distribution of $\mu_{P'}$ on the coordinates of s, s'.

Proof. For any $s, s' \in \mathcal{V}$ and $c, c' \in \{1, \ldots, q-1\}$, we have

$$\begin{aligned}\mathbf{E}[g_{s,c} \cdot g_{s',c'}] &= \mathbf{E}\left[\sum_{\omega\in\Sigma} \ell_{s,c}(\omega)\langle \boldsymbol{b}_{s,\omega}, \boldsymbol{\zeta}\rangle \sum_{\omega'\in\Sigma} \ell_{s',c'}(\omega')\langle \boldsymbol{b}_{s',\omega'}, \boldsymbol{\zeta}\rangle\right] \\ &= \sum_{\omega,\omega'\in\Sigma} \ell_{s,c}(\omega)\ell_{s',c'}(\omega')\,\mathbf{E}\left[\langle \boldsymbol{b}_{s,\omega}, \boldsymbol{\zeta}\rangle\langle \boldsymbol{b}_{s',\omega'}, \boldsymbol{\zeta}\rangle\right] \\ &= \sum_{\omega,\omega'\in\Sigma} \ell_{s,c}(\omega)\ell_{s',c'}(\omega')\langle \boldsymbol{b}_{s,\omega}, \boldsymbol{b}_{s',\omega'}\rangle. \qquad (8)\end{aligned}$$

Now, when $s = s'$, for $\omega \neq \omega'$, $\langle \boldsymbol{b}_{s,\omega}, \boldsymbol{b}_{s',\omega'}\rangle = 0$ because of the SDP constraints (3). Therefore, in this case

$$\begin{aligned}\mathbf{E}[g_{s,c} \cdot g_{s,c'}] &= \sum_{\omega\in\Sigma} \ell_{s,c}(\omega)\ell_{s,c'}(\omega)\|\boldsymbol{b}_{s,\omega}\|_2^2 \\ &= \mathop{\mathbf{E}}_{\omega\sim\mu_s}[\ell_{s,c}(\omega)\ell_{s,c'}(\omega)] = \langle \ell_{s,c}, \ell_{s,c'}\rangle_{\mu_s},\end{aligned}$$

which is 1 when $c = c'$ and 0 otherwise. This shows the orthonormality of $\boldsymbol{g}_s$. Coming back to the Equation (8), again by the SDP constraints (3), the inner-product $\langle \boldsymbol{b}_{s,\omega}, \boldsymbol{b}_{s',\omega'}\rangle$ is precisely the probability of (ω, ω') according to the distribution $\mu_{P'}|(s, s')$ for any payoff P' containing s and s'. This proves the lemma. □

Let $\mathbf{Round}_{\mathcal{F}}(\boldsymbol{V}, \boldsymbol{\mu})$ be the expected value of the assignment returned by the rounding algorithm in Figure 4. In this section, we prove the following soundness lemma.

Lemma 2.22. *Let $k = 3$ and assume that the SDP solution is semi-rich and does not satisfy any linear equation. Then, for any (τ, δ)-quasirandom function $\mathcal{F}$,*

$$\text{Soundness}(\mathbf{Dict}_{\boldsymbol{V},\boldsymbol{\mu}}) \leqslant \mathbf{Round}_{\mathcal{F}}(\boldsymbol{V}, \boldsymbol{\mu}) + o_{\delta,\tau}(1).$$

The notation $o_{\delta,\tau}(1)$ means that it goes to 0 as $\delta \to 0$ and $\tau \to 0$. Therefore, in this case the acceptance probability of the test is upper bounded by the integral value of the given instance. This shows that if there exists an $(1, s)$ integrality gap instance of Max-P-CSP, then there exists a dictatorship test with completeness 1 and soundness $s + \varepsilon$ for any constant $\varepsilon > 0$.

Remark 2.23. *If we can extend our main analytical lemma to other predicates, then we can remove the condition on the predicate from Lemma 2.22.*

The acceptance probability of the test for a given function $\mathcal{F}$ is given by:

$$\begin{aligned}&\mathbf{Pr}[\mathcal{F} \text{ passes } \mathbf{Dict}_{\boldsymbol{V},\boldsymbol{\mu}}] \\ &= \mathop{\mathbf{E}}_{P'\sim\mathcal{P}} \mathop{\mathbf{E}}_{\boldsymbol{z}_{P'}}[P'(\mathcal{F}_{s_1}(\boldsymbol{z}_{s_1}), \mathcal{F}_{s_2}(\boldsymbol{z}_{s_2}), \ldots, \mathcal{F}_{s_k}(\boldsymbol{z}_{s_k}))].\end{aligned}$$

We will prove a series of claims which will help us relate the probability to $\mathbf{Round}_{\mathcal{F}}(\boldsymbol{V}, \boldsymbol{\mu})$. We begin with the following claim which shows that we can replace $\mathcal{F}$ with its *noisy* version $T_{1-\delta}\mathcal{F}$. Here, we use the main analytical lemma (Lemma 2.6).

Claim 2.24 (Changing $\mathcal{F}$ to $\mathcal{H}$). *Let $k = 3$ and assume that the SDP solution is semi-rich and does not satisfy any linear equation. Then for every $P' \in \mathcal{P}$,*

$$\left|\begin{array}{c}\mathbf{E}_{\boldsymbol{z}_{P'}}[P'(\mathcal{F}_{s_1}(\boldsymbol{z}_{s_1}), \mathcal{F}_{s_2}(\boldsymbol{z}_{s_2}), \ldots, \mathcal{F}_{s_k}(\boldsymbol{z}_{s_k}))] - \\ \mathbf{E}_{\boldsymbol{z}_{P'}}[P'(\mathcal{H}_{s_1}(\boldsymbol{z}_{s_1}), \mathcal{H}_{s_2}(\boldsymbol{z}_{s_2}), \ldots, \mathcal{H}_{s_k}(\boldsymbol{z}_{s_k}))]\end{array}\right| \leqslant \eta(\delta).$$

Proof. Consider the following expression.

$$P'(\mathcal{F}_{s_1}(\boldsymbol{z}_{s_1}), \mathcal{F}_{s_2}(\boldsymbol{z}_{s_2}), \ldots, \mathcal{F}_{s_k}(\boldsymbol{z}_{s_k})) = \sum_{\sigma\in\Sigma^k} P'(\sigma)\prod_{j=1}^{k} \mathcal{F}_{s_j,\sigma_j}(\boldsymbol{z}_{s_j}).$$

We will show that for all $P' \in \mathcal{P}$ and $\sigma \in \Sigma^k$,

$$\Gamma := \left|\mathop{\mathbf{E}}_{\boldsymbol{z}_{P'}}\left[\prod_{j=1}^{k} \mathcal{F}_{s_j,\sigma_j}(\boldsymbol{z}_{s_j})\right] - \mathop{\mathbf{E}}_{\boldsymbol{z}_{P'}}\left[\prod_{j=1}^{k} \mathcal{H}_{s_j,\sigma_j}(\boldsymbol{z}_{s_j})\right]\right| \leqslant \eta.$$

Let us define $\Gamma_{j'}$ for $j' = 1, \ldots, k$ as follows:

$$\Gamma_{j'} := \left|\begin{array}{c}\mathbf{E}_{\boldsymbol{z}_{P'}}\left[\prod_{j=1}^{j'-1} \mathcal{H}_{s_j,\sigma_j}(\boldsymbol{z}_{s_j}) \prod_{j=j'}^{k} \mathcal{F}_{s_j,\sigma_j}(\boldsymbol{z}_{s_j})\right] - \\ \mathbf{E}_{\boldsymbol{z}_{P'}}\left[\prod_{j=1}^{j'} \mathcal{H}_{s_j,\sigma_j}(\boldsymbol{z}_{s_j}) \prod_{j=j'+1}^{k} \mathcal{F}_{s_j,\sigma_j}(\boldsymbol{z}_{s_j})\right]\end{array}\right|.$$

By triangle inequality, $\Gamma \leqslant \sum_{j'} \Gamma_{j'}$.

$$\begin{aligned}\Gamma_{j'} &= \left|\begin{array}{c}\mathbf{E}_{\boldsymbol{z}_{P'}}\left[\prod_{j=1}^{j'-1} \mathcal{H}_{s_j,\sigma_j}(\boldsymbol{z}_{s_j}) \prod_{j=j'}^{k} \mathcal{F}_{s_j,\sigma_j}(\boldsymbol{z}_{s_j})\right] - \\ \mathbf{E}_{\boldsymbol{z}_{P'}}\left[\prod_{j=1}^{j'} \mathcal{H}_{s_j,\sigma_j}(\boldsymbol{z}_{s_j}) \prod_{j=j'+1}^{k} \mathcal{F}_{s_j,\sigma_j}(\boldsymbol{z}_{s_j})\right]\end{array}\right| \\ &= \left|\mathop{\mathbf{E}}_{\boldsymbol{z}_{P'}}\left[\begin{array}{c}(\mathcal{F}_{s_{j'},\sigma_{j'}}(\boldsymbol{z}_{s_{j'}}) - \mathcal{H}_{s_{j'},\sigma_{j'}}(\boldsymbol{z}_{s_{j'}}))\cdot \\ \prod_{j=1}^{j'-1} \mathcal{H}_{s_j,\sigma_j}(\boldsymbol{z}_{s_j}) \prod_{j=j'+1}^{k} \mathcal{F}_{s_j,\sigma_j}(\boldsymbol{z}_{s_j})\end{array}\right]\right| \\ &= \left|\mathop{\mathbf{E}}_{\boldsymbol{z}_{P'}}\left[\begin{array}{c}(\mathbf{Id} - T_{1-\delta})\mathcal{F}_{s_{j'},\sigma_{j'}}(\boldsymbol{z}_{s_{j'}}) \cdot \prod_{j=1}^{j'-1} \mathcal{H}_{s_j,\sigma_j}(\boldsymbol{z}_{s_j}) \\ \prod_{j=j'+1}^{k} \mathcal{F}_{s_j,\sigma_j}(\boldsymbol{z}_{s_j})\end{array}\right]\right|.\end{aligned}$$

Here, $\mathbf{Id}$ is the identity operator. Now, the function $Q := (\mathbf{Id} - T_{1-\delta})\mathcal{F}_{s_{j'},\sigma_{j'}}(\boldsymbol{z}_{s_{j'}})$ is a function $Q : \Sigma^R \to [0, 1]$ that satisfies the property of being a 'high-degree' function: Using the Efron-Stein decomposition of Q and using Fact 2.14, we have

$$\langle Q, T_{1-\xi}Q\rangle_{\mu_{s_{j'}}^{\otimes R}} = \sum_{S\subseteq[R]} (1-\xi)^{|S|}(1-(1-\delta)^{|S|})^2\|(\mathcal{F}_{s_{j'},\sigma_{j'}})_S\|_2^2.$$

Now, $(1-\xi)^\ell(1-(1-\delta)^\ell)^2 \leqslant \eta$ for every $\ell \geqslant 0$. Therefore,

$$\langle Q, T_{1-\xi}Q\rangle_{\mu_{s_{j'}}^{\otimes R}} \leqslant \eta \sum_{S\subseteq[R]} \|(\mathcal{F}_{s_{j'},\sigma_{j'}})_S\|_2^2 \leqslant \eta(\delta).$$

Hence, the product inside the expectation satisfies the hypothesis of Lemma 2.6, with $\mathrm{Stab}_{1-\xi}(Q) \leqslant \eta(\delta)$. Applying the lemma, we conclude that $\Gamma_{j'} \leqslant \eta(\delta)/k$, where $\eta(\delta) \to 0$ as $\delta \to 0$. Therefore, $\Gamma \leqslant \sum_{j'} \Gamma_{j'} \leqslant \eta(\delta)$. □

We now switch to the multilinear polynomials. By definition, we have

$$\begin{aligned}&\mathop{\mathbf{E}}_{\boldsymbol{z}_{P'}}[P'(\mathcal{H}_{s_1}(\boldsymbol{z}_{s_1}), \mathcal{H}_{s_2}(\boldsymbol{z}_{s_2}), \ldots, \mathcal{H}_{s_k}(\boldsymbol{z}_{s_k}))] \\ &= \mathop{\mathbf{E}}_{\mathcal{L}_{P'}^R}[P'(\boldsymbol{H}_{s_1}(\boldsymbol{\ell}_{s_1}), \boldsymbol{H}_{s_2}(\boldsymbol{\ell}_{s_2}), \ldots, \boldsymbol{H}_{s_k}(\boldsymbol{\ell}_{s_k}))].\end{aligned}$$

Here, $\mathcal{L}_{P'}$ is the joint distribution of the local ensembles based on the distribution $\mu_{P'}$. We now apply the Invariance Principle to replace the Integral Ensembles with the Gaussian Ensembles.

Claim 2.25. *(Moving to the global Gaussian ensembles) Using the invariance principle, for every $P' \in \mathcal{P}$, we have*

$$\left|\begin{array}{c}\mathbf{E}_{\mathcal{L}_{P'}^R}[P'(\boldsymbol{H}_{s_1}(\boldsymbol{\ell}_{s_1}), \boldsymbol{H}_{s_2}(\boldsymbol{\ell}_{s_2}), \ldots, \boldsymbol{H}_{s_k}(\boldsymbol{\ell}_{s_k}))] - \\ \mathbf{E}_{\mathcal{G}_{P'}^R}[P'(\boldsymbol{H}_{s_1}(\boldsymbol{g}_{s_1}), \boldsymbol{H}_{s_2}(\boldsymbol{g}_{s_2}), \ldots, \boldsymbol{H}_{s_k}(\boldsymbol{g}_{s_k}))]\end{array}\right| \leqslant \tau^{O_{\delta,\alpha}(1)}.$$

Proof. This claim follows directly from the Invariance Principle, i.e., from Theorem 2.15, and using Lemma 2.21. Here, the maximum influence of the functions is at most τ and any non-zero probability of an atom is at least α. Also, for $D = O(\log_{1-\delta}\tau)$, the degree $\geqslant D$ weight of the functions $\boldsymbol{H}_s$ is at most $O(\tau)$. This is as follows.

$$W^{\geqslant D}[\mathcal{H}_{s,i};\mu^{\otimes R}] = \sum_{\sigma:|\sigma|\geqslant D}(\hat{\mathcal{H}_{s,i}})^2_{\sigma} = \sum_{\sigma:|\sigma|\geqslant D}(1-\delta)^{|\sigma|}(\hat{\mathcal{F}_{s,i}})^2_{\sigma} \leqslant (1-\delta)^D \leqslant \tau.$$

Therefore, $W^{\geqslant D}[\mathcal{H}_s;\mu^{\otimes R}] = \sum_{j=1}^{q} W^{\geqslant D}[\mathcal{H}_{s,j};\mu^{\otimes R}] \leqslant q\cdot\tau = O(\tau)$. □

The final claim shows that, as far as the multilinear polynomial evaluations are concerned, the rounding step (Step II (3)) does not change the expectation by much if the function $\mathcal{F}$ is a quasirandom function.

Claim 2.26. *(Analyzing the loss due to truncation and scaling) For every payoff $P' \in \mathcal{P}$,*

$$\left|\begin{array}{c}\mathbf{E}_{\mathcal{G}^R}[P'(\boldsymbol{H}_{s_1}(\boldsymbol{g}_{s_1})^{\star},\boldsymbol{H}_{s_2}(\boldsymbol{g}_{s_2})^{\star},\ldots,\boldsymbol{H}_{s_k}(\boldsymbol{g}_{s_k})^{\star})]- \\ \mathbf{E}_{\mathcal{G}^R}[P'(\boldsymbol{H}_{s_1}(\boldsymbol{g}_{s_1}),\boldsymbol{H}_{s_2}(\boldsymbol{g}_{s_2}),\ldots,\boldsymbol{H}_{s_k}(\boldsymbol{g}_{s_k}))]\end{array}\right| \leqslant \tau^{O_{\delta,\alpha}(1)}.$$

Proof. $\mathcal{H}_{s_j} = T_{1-\delta}\mathcal{F}_{s_j}$ is over the domain Σ^R and has the range $\blacktriangle_q$. The difference between the first and the second expression (rounding error because of scaling and truncation) is bounded by $O(C_0,q)\cdot\sum_{s\in\mathcal{V}(P')}\mathbf{E}[\boldsymbol{\xi}(\boldsymbol{H}_s(\boldsymbol{g}_s))]$ [22, Claim 7.4.2], where $\boldsymbol{\xi}(\boldsymbol{a}) = \sum_j (f_{[0,1]}(a_j)-a_j)^2$ and C_0 is an absolute constant from the smoothness property of the payoff P'. We know that $\mathbf{E}[\boldsymbol{\xi}(\boldsymbol{H}_s(\boldsymbol{\ell}_s))] = 0$, as $\boldsymbol{H}_s(\boldsymbol{\ell}_s)\in\blacktriangle_q$. Now, we can apply the invariance principle to conclude

$$\left|\mathop{\mathbf{E}}_{\mathcal{G}^R}[\boldsymbol{\xi}(\boldsymbol{H}_s(\boldsymbol{g}_s))] - \mathop{\mathbf{E}}_{\mathcal{L}^R_{P'}}[\boldsymbol{\xi}(\boldsymbol{H}_s(\boldsymbol{\ell}_s))]\right| \leqslant \tau^{O_{\delta,\alpha}(1)}.$$

As $\mathbf{E}[\boldsymbol{\xi}(\boldsymbol{H}_s(\boldsymbol{\ell}_s))] = 0$, the claim follows. □

Proof of Lemma 2.22. We are now ready to prove the soundness of the test: The value returned by the rounding scheme is

$$\mathbf{Round}_{\mathcal{F}}(\boldsymbol{V},\boldsymbol{\mu}) = \mathop{\mathbf{E}}_{P'\in\mathcal{P}}\mathop{\mathbf{E}}_{\mathcal{G}^R}\left[P'(\boldsymbol{H}_{s_1}(\boldsymbol{g}_{s_1})^{\star},\boldsymbol{H}_{s_2}(\boldsymbol{g}_{s_2})^{\star},\ldots,\boldsymbol{H}_{s_k}(\boldsymbol{g}_{s_k})^{\star})\right]$$

and the soundness of the test is given by the following expression:

$$\mathbf{Pr}[\mathcal{F}\text{ passes }\mathbf{Dict}_{\boldsymbol{V},\boldsymbol{\mu}}] = \mathop{\mathbf{E}}_{P'\sim\mathcal{P}}\mathop{\mathbf{E}}_{\boldsymbol{z}_{P'}}[P'(\mathcal{F}_{s_1}(\boldsymbol{z}_{s_1}),\mathcal{F}_{s_1}(\boldsymbol{z}_{s_2}),\ldots,\mathcal{F}_{s_k}(\boldsymbol{z}_{s_k}))].$$

For $k=3$, using the Claims 2.24, 2.25, 2.26 that we proved earlier, we can relate the two quantities as follows:

$$\mathbf{Pr}[\mathcal{F}\text{ passes }\mathbf{Dict}_{\boldsymbol{V},\boldsymbol{\mu}}] \leqslant \mathbf{Round}_{\mathcal{F}}(\boldsymbol{V},\boldsymbol{\mu}) + \eta(\delta) + \tau^{O_{\delta,\alpha}(1)}.$$

Now, $\eta(\delta)\to 0$ as $\delta\to 0$. Therefore, we get

$$\mathbf{Pr}[\mathcal{F}\text{ passes }\mathbf{Dict}_{\boldsymbol{V},\boldsymbol{\mu}}] \leqslant \mathbf{Round}_{\mathcal{F}}(\boldsymbol{V},\boldsymbol{\mu}) + o_{\delta,\tau}(1),$$

as required. □

3 THE MAIN ANALYTICAL LEMMA

In this section, we give an overview of the proof our main analytical lemma (Lemma 2.6). The reader is referred to the full-version of the paper for the complete proof .

We begin by addressing a more specialized case, in which the requirement of semi-rich support of the distribution is replaced with the stronger condition that the support of the distribution is a *union of matchings*:

Definition 3.1. *We say a set $S \subseteq \Sigma\times\Phi\times\Gamma$ is a union of matchings if there exists $\Sigma' \subseteq \Sigma$ and a collection of matchings $M_x \subseteq \Phi\times\Gamma$, one for each $x\in\Sigma'$, such that*

$$S = \bigcup_{x\in\Sigma'}\{x\}\times M_x.$$

The version of Lemma 2.6 for union of matchings is Lemma 3.2 stated below; another difference is that below we introduce some asymmetry in the roles of f, g and h, and we need the stability of either g or h to be small. In the full-version of the paper, we explain the slight adaptations that allow our argument to go through in the case of semi-rich support, and then explain how to generalize the statement to the case the stability of f is small (thereby establishing Lemma 2.6).

Lemma 3.2. *For all $m\in\mathbb{N}$, $\varepsilon,\alpha > 0$ there exist $\xi > 0$ and $\delta > 0$ such that the following holds. Suppose μ is a distribution over $\Sigma\times\Phi\times\Gamma$ whose support (a) is a union of matchings, and (b) cannot be embedded in an Abelian group. Further suppose that $|\Sigma|,|\Phi|,|\Gamma| \leqslant m$ and each atom in μ has probability at least α. Then, if $f\colon\Sigma^n\to[-1,1]$ $g\colon\Phi^n\to[-1,1]$, $h\colon\Gamma^n\to[-1,1]$ are functions such that*

- $\mathsf{Stab}_{1-\xi}(g)\leqslant\delta$ *or* $\mathsf{Stab}_{1-\xi}(h)\leqslant\delta$.

Then $\left|\mathbf{E}_{\mathbf{x},\mathbf{y},\mathbf{z}\sim\mu^{\otimes n}}\left[f(\mathbf{x})g(\mathbf{y})h(\mathbf{z})\right]\right| \leqslant \varepsilon$.

As the roles of g and h will be interchangeable in our arguments, without loss of generality we shall focus on the case that $\mathsf{Stab}_{1-\xi}(g)\leqslant\delta$ throughout this section. Before proceeding to the formal argument, we begin with a quick overview of the proof that outlines the main components involved.

Proof overview. The proof of Lemma 3.2 consists of several steps. We think of $\mathrm{supp}(\mu)$ as a graph between Φ and Γ, wherein edges are labeled by elements of Σ in the natural way. Our initial premise is that for each $x\in\Sigma$, the collection of edges labeled by x forms a matching, and we perform several steps in order to improve the structure we have on that graph (by possibly increasing the size of the alphabet Σ).

(1) Let $T_x\in\{0,1\}^{\Phi\times\Gamma}$ be the permutation matrix corresponding to the matching labeled by x. First, we show that by moving to a different distribution μ' satisfying similar properties to μ', we may assume that not only the edges of T_x lie in the graph of μ', but rather also the edges of $T_{x_1}T^t_{x_2}T_{x_3}$ for any $x_1,x_2,x_3\in\Sigma$. In other words, we may compose various matchings and "insert" them into the support of our distribution. Performing this step $\ell = O_m(1)$ times, we get that as the graph of μ is connected, we would end up with the complete bipartite graph between Φ and Γ. We now move on to a similar looking expectation to the one in the main lemma but for μ', which is a distribution over $\Sigma^{\ell}\times\Phi\times\Gamma$.

(2) We next reduce the size of the alphabet Σ^ℓ to be smaller. Note that for each $\vec{x} \in \Sigma^\ell$, the edges in the graph of μ' labeled by $\vec{x}$ form a matching. We show that if for $\vec{x}, \vec{x}'$ these matchings are not edge disjoint, then we may glue together the symbols $\vec{x}, \vec{x}'$ and modify the distribution μ' and the functions f, g, h (in a way that preserves their various properties) so that the expectation does not drop too much. The edges of the new symbols will consist of the union of the edges of the old symbols, and the new alphabet for x is $\Sigma' \subseteq \Sigma^\ell$
We note that in such operation, if the matchings corresponding to $\vec{x}, \vec{x}'$ were not identical, then the edges corresponding to the new symbol will not form a matching. We show that in that case, one may further do identification of symbols in Φ and Γ that preserve the properties of the distributions and the functions, and keeps the expectation high. Performing such identification steps sufficiently many times, one returns to the case wherein for each $x \in \Sigma'$ the edges corresponding to x form a matching. We note that each time we perform such step, the alphabet of y or z drops by at least 1, so in total we will have at most $2m$ such steps.

(3) We thus reach new alphabets Σ', Φ', Γ'. We consider further operations of composing three x-matchings, i.e. moving from Σ' to Σ'^3. We say that this move is worthwhile if doing it, and then the subsequent identifications, the alphabets Φ', Γ' will shrink further. As long as performing this move is worthwhile, we do so and otherwise we proceed to the next step.

(4) After performing $O_m(1)$ steps as in the previous item, we reach to the state wherein the alphabets are Σ'', Φ'' and Γ'', and it is no longer worthwhile to execute the previous step. This means that for every $(x_1, x_2, x_3) \in \Sigma''^3$ and $(x_4, x_5, x_6) \in \Sigma''^3$, the permutations $T_{x_1} T_{x_2}^t T_{x_3}$ and $T_{x_4} T_{x_5}^t T_{x_6}$ are either identical, or are edge disjoint (otherwise we would be able to execute the previous step once more). We use this structure in order to identify a non-Abelian group structure. More specifically, we construct a group $(G, \cdot)$ that has no representations of dimension 1 (besides the trivial representation), such that our expectation is

$$\mathop{\mathbb{E}}_{(g_1,g_2,g_3): g_3 = g_1 g_2} \left[f'(g_1) g'(g_2) h'(g_3) \right].$$

Here, f', g', h' are really the same as the functions f, g, h we have, except that they interpret their input as elements from G. We argue that the fact that g' is highly noise sensitive implies that almost all of the mass of g' (with respect to the representation theoretic Fourier decomposition over G) lies on the high degrees. We use this fact along with basic Fourier analysis in order to give an upper bound on the expectation above that vanishes as $\xi, \delta \to 0$ (uniformly in n), and hence finish the proof.

REFERENCES

[1] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. 1998. Proof Verification and the Hardness of Approximation Problems. *J. ACM* 45, 3 (May 1998), 501–555. https://doi.org/10.1145/278298.278306 (Preliminary version in *33rd FOCS*, 1992).

[2] Sanjeev Arora and Shmuel Safra. 1998. Probabilistic Checking of Proofs: A New Characterization of NP. *J. ACM* 45, 1 (Jan. 1998), 70–122. https://doi.org/10.1145/273865.273901 (Preliminary version in *33rd FOCS*, 1992).

[3] Per Austrin and Elchanan Mossel. 2009. Approximation resistant predicates from pairwise independence. *Computational Complexity* 18, 2 (2009), 249–271. https://doi.org/10.1007/s00037-009-0272-6

[4] Amey Bhangale and Subhash Khot. 2021. Optimal Inapproximability of Satisfiable K-LIN over Non-Abelian Groups. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing (STOC)* (Virtual, Italy). Association for Computing Machinery, New York, NY, USA, 1615–1628. https://doi.org/10.1145/3406325.3451003

[5] Joshua Brakensiek and Venkatesan Guruswami. 2021. The Quest for Strong Inapproximability Results with Perfect Completeness. *ACM Trans. Algorithms* 17, 3, Article 27 (jul 2021), 35 pages. https://doi.org/10.1145/3459668

[6] Mark Braverman, Subhash Khot, Noam Lifshitz, and Dor Minzer. 2022. An Invariance Principle for the Multi-slice, with Applications. In *62nd Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, Denver, CO, USA, 228–236. https://doi.org/10.1109/FOCS52979.2021.00030

[7] Mark Braverman, Subhash Khot, and Dor Minzer. 2021. On Rich 2-to-1 Games. In *12th Innovations in Theoretical Computer Science Conference (ITCS) (Leibniz International Proceedings in Informatics (LIPIcs), Vol. 185)*. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 27:1–27:20. https://doi.org/10.4230/LIPIcs.ITCS.2021.27

[8] Andrei A. Bulatov. 2017. A Dichotomy Theorem for Nonuniform CSPs. In *58th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, Berkeley, CA, USA, 319–330. https://doi.org/10.1109/FOCS.2017.37

[9] Lars Engebretsen and Jonas Holmerin. 2005. Three-query PCPs with perfect completeness over non-Boolean domains. *Random Structures & Algorithms* 27, 1 (2005), 46–75. https://doi.org/10.1002/rsa.20050

[10] Tomás Feder and Moshe Y. Vardi. 1998. The Computational Structure of Monotone Monadic SNP and Constraint Satisfaction: A Study through Datalog and Group Theory. *SIAM J. Comput.* 28, 1 (1998), 57–104. https://doi.org/10.1137/S0097539794266766

[11] Uriel Feige, Shafi Goldwasser, Laszlo Lovász, Shmuel Safra, and Mario Szegedy. 1996. Interactive proofs and the hardness of approximating cliques. *Journal of the ACM (JACM)* 43, 2 (1996), 268–292. https://doi.org/10.1145/226643.226652

[12] Michel X. Goemans and David P. Williamson. 1995. Improved Approximation Algorithms for Maximum Cut and Satisfiability Problems Using Semidefinite Programming. *J. ACM* 42, 6 (1995), 1115–1145. https://doi.org/10.1145/227683.227684

[13] Johan Håstad. 2001. Some optimal inapproximability results. *J. ACM* 48, 4 (July 2001), 798–859. https://doi.org/10.1145/502090.502098 (Preliminary version in *29th STOC*, 1997).

[14] Johan Håstad and Subhash Khot. 2005. Query Efficient PCPs with Perfect Completeness. *Theory of Computing* 1, 7 (2005), 119–148. https://doi.org/10.4086/toc.2005.v001a007

[15] Sangxia Huang. 2014. Approximation Resistance on Satisfiable Instances for Sparse Predicates. *Theory of Computing* 10, 14 (2014), 359–388. https://doi.org/10.4086/toc.2014.v010a014

[16] Johan Håstad. 2014. On the NP-Hardness of Max-Not-2. *SIAM J. Comput.* 43, 1 (2014), 179–193. https://doi.org/10.1137/120882718

[17] Subhash Khot. 2002. On the power of unique 2-prover 1-round games. In *Proceedings of the 34th Annual ACM symposium on Theory of computing (STOC)*. ACM, Association for Computing Machinery, New York, NY, USA, 767–775. https://doi.org/10.1145/509907.510017

[18] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. 2007. Optimal Inapproximability Results for MAX-CUT and Other 2-Variable CSPs? *SIAM J. Comput.* 37, 1 (2007), 319–357. https://doi.org/10.1137/S0097539705447372

[19] Elchanan Mossel. 2010. Gaussian bounds for noise correlation of functions. *Geometric and Functional Analysis* 19, 6 (2010), 1713–1756. https://doi.org/10.1007/s00039-010-0047-x

[20] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. 2005. Noise stability of functions with low influences: invariance and optimality. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. IEEE, IEEE, Pittsburgh, PA, USA, 21–30. https://doi.org/10.1109/SFCS.2005.53

[21] Prasad Raghavendra. 2008. Optimal algorithms and inapproximability results for every CSP?. In *Proceedings of the 14th annual ACM symposium on Theory of computing (STOC)*. Association for Computing Machinery, New York, NY, USA, 245–254. https://doi.org/10.1145/1374376.1374414

[22] Prasad Raghavendra. 2009. *Approximating np-hard problems efficient algorithms and their limits.* Ph.D. thesis, University of Washington, Seattle, WA, USA.

[23] Thomas J. Schaefer. 1978. The Complexity of Satisfiability Problems. In *Proceedings of the 10th Annual ACM Symposium on Theory of Computing (STOC)* (San Diego, California, USA). Association for Computing Machinery, New York, NY, USA, 216–226. https://doi.org/10.1145/800133.804350

[24] Linqing Tang. 2009. Conditional hardness of approximating satisfiable Max 3CSP-q. In *International Symposium on Algorithms and Computation*. Springer, Springer Berlin Heidelberg, Berlin, Heidelberg, 923–932. https://doi.org/10.1007/978-3-642-10631-6_93

[25] Dmitriy Zhuk. 2020. A Proof of the CSP Dichotomy Conjecture. *J. ACM* 67, 5, Article 30 (Aug. 2020), 78 pages. https://doi.org/10.1145/3402029